Semantic Models of Performance Indicators: A Systematic Survey

CLAUDIA DIAMANTINI, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy
TARIQUE KHAN, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy
DOMENICO POTENA, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy
EMANUELE STORTI, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy

Performance indicators and metrics are essential management tools. They provide synthetic objective measures to monitor the progress of a process, set objectives, and assess deviations, enabling effective decision-making. They can also be used for communication purposes, facilitating the sharing of objectives and results or improving the awareness of certain phenomena, thus motivating more responsible and sustainable behaviors. Given their strategic role, it is of paramount importance, as well as challenging, to guarantee that the intended meaning of an indicator is fully shared among stakeholders and that its implementation is aligned with the definition provided by decision makers, as this is a precondition for data quality and trustworthiness of the information system. Formal models, such as ontologies, have been long investigated in the literature to address the issues. This article proposes a comprehensive survey on semantic approaches aimed to specify conceptual definitions of indicators and metrics, illustrating also the advantages of these formal approaches in relevant use cases and application domains.

CCS Concepts: • **General and reference → Surveys and overviews**; **Metrics**; • **Information systems → Ontologies;**

Additional Key Words and Phrases: KPI, performance indicators, indicators, semantic models, challenges

ACM Reference Format:
Claudia Diamantini, Tarique Khan, Domenico Potena, and Emanuele Storti. 2025. Semantic Models of Performance Indicators: A Systematic Survey. *ACM Comput. Surv.* 57, 8, Article 202 (March 2025), 37 pages. https://doi.org/10.1145/3719291

Authors' Contact Information: Claudia Diamantini, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy; e-mail: c.diamantini@univpm.it; Tarique Khan, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy; e-mail: t.khan@pm.univpm.it; Domenico Potena, Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy; e-mail: d.potena@univpm.it; Emanuele Storti (Corresponding author), Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy; e-mail: e.storti@univpm.it.

ACM 0360-0300/2025/03-ART202
https://doi.org/10.1145/3719291

1 Introduction

Performance indicators (PIs) and metrics provide a synthetic view of the progress of an action, the level of achievement of a goal, or the performance of a process within an organization. The adoption of few, synthetic indicators lies at the core of performance measurement and decision-making across diverse sectors, allowing organizations to track progress, identify areas for improvement, and make informed strategic decisions. From corporate and supply chain management (e.g., lead time, throughput, inventory turnover, delivery time, defect rates) to healthcare (e.g., average hospital stay, bed turnover, average patient wait time), education (e.g., graduation rate, exam pass rate, student-to-faculty ratio), ICT (e.g., system uptime, latency and packet loss), finance (e.g., net profit margin, cash flow, return on investment), and the non-profit and public sector (most notably, the recent release of sustainability goals and related targets and indicators by the United Nations[1]), indicators play a pivotal role. Given their relation to strategic decision-making, it is of paramount importance for corporates, supply chains, and networked organizations to provide a well-founded means for sharing and integrating indicators (consider, for instance, the case of the United Nations, which collects indicators from different countries to measure the achievement of its sustainability goals). To address such issues, semantic technologies have proven to be a powerful tool for injecting intelligence (i.e., reasoning ability, flexibility) into computer systems as they facilitate the sharing and integration of resources. These technologies indeed allow semantically enriched representations to provide meaning and context to data, improving performance in a variety of applications, like information retrieval, data access and query answering, data source integration and interoperability, and knowledge sharing. For this reason, a plethora of artifacts like ontologies and knowledge graphs have been proposed to model real-world entities and properties in the most disparate variety of domains, from life science and medicine to physical sciences, social sciences, economy, and engineering (for a partial list, see catalogs like LOV[2] or OLS[3]).

A huge amount of work, witnessed by several ontologies and semantic models, has been developed by the research community with the aim to represent P and metrics in a formal way, in order to support more advanced monitoring and analytical applications, reasoning, sharing, and reuse of indicator definitions. In the literature, a good number of articles, in the form of catalogs, reviews, and surveys on semantic approaches, have been published in the context of several application fields, e.g., [27, 37, 41, 64, 76]. To the best of our knowledge, however, a similar work has not yet been attempted with the goal to collect, categorize, and analyze semantic models for performance indicators.

In this article, we propose a survey of the research conducted in the past two decades on semantic models of performance indicators. We adopted a principled methodology to select the relevant literature from scientific libraries, according to the designed research questions. Results have been analyzed according to a classification scheme that includes a set of dimensions considered relevant to the subject. In particular, we analyze the expressiveness of semantic models in terms of the covered concepts, also digging into the way core concepts and relations are represented in different articles, the formalization degree, and related reasoning capabilities supported by the model. Furthermore, we recognized the different application domains and goals for which the models have been developed. In particular, we discovered that the level of generality greatly varies from very specific artifacts built for modeling narrow domains to general-purpose models that can be applied to a variety of situations and contexts. Finally, we address technical aspects related to sharing and reusability by discussing the existence of implementations and

[1] https://sdgs.un.org/goals
[2] https://lov.linkeddata.es/dataset/lov/
[3] https://www.ebi.ac.uk/ols4

specifications. As a result, the survey provides a complete overview of the landscape of existing research on semantic representations of indicators and related products that we believe represents an invaluable resource for researchers and practitioners who want to adopt semantic technologies to manage, share, and integrate indicators in modern collaborative networked scenarios.

The rest of the study is structured as follows: Section 2 introduces the notions of indicator and measure, the challenges related to their management from both business and technical perspectives, and semantic models. Section 3 focuses on the state of the art on surveys related to the current topic. Section 4 discusses the methodology in detail, involving the planning of the study and its implementation. Results of the study are reported in Section 5. Finally, Section 6 summarizes the contribution of the work and proposes general guidelines for selecting the most appropriate approach to modeling PIs based on a given scenario. Additionally, it discusses the limitation of the work and open challenges.

2 Background

2.1 Indicators and Measures

In the context of this work, we use the term *indicator* to refer to any synthetic, quantitative measure used to monitor a certain phenomenon inside an organization. In business, indicators are commonly used to evaluate the progress of a process, to set objectives, and to assess deviations. Depending on the user community and the domain, different terminologies are adopted, e.g., metric, PI, **Key Performance Indicators (KPI)**, or **Process Performance Indicator (PPI)**. Although sometimes used as synonyms, there are some differences. In the present work we propose these definitions:

— Metric: A measure to quantitatively assess a phenomenon, e.g., CPU utilization.
— Performance Indicator: A measure that assumes a meaning in relation to the achievement of a goal or objective, typically associated to a positive trend and/or a threshold. For instance, in order to justify the investment in an IT system, optimal CPU utilization is crucial, and a threshold of 70% or higher is desirable. Establishing predefined ranges for acceptable values is also common. An overloaded resource is indicated when CPU utilization exceeds 80%, suggesting that a favorable operational range lies between 70% and 80%.
— Key Performance Indicator: A performance indicator tied to a strategic goal, e.g., Return on Investment.
— Process Performance Indicator: A performance indicator specifically tied to a process, e.g., Lead Time.

Regardless of the specific nuance of meaning, indicators are essential management tools, as they enable effective decision-making and allow to verify the impact produced by decisions, facilitate the sharing of objectives and results, and help to improve awareness and motivate more responsible and sustainable behaviors.

In organizational scenarios, traditional solutions for modeling **Performance Measurement and Monitoring (PMM)** systems were used to take merely an economic and financial perspective, until more structured approaches, such as Balanced Scorecards [40], introduced also non-strictly financial aspects in performance management [89]. In this context, indicators have been introduced as a way to model PMMs in a more structured manner and with a more comprehensive scope than traditional ones. Since then, practically every sector has developed a set of potentially relevant PIs for the most disparate goals. New trends have also recently emerged, for instance, modeling PIs to monitor innovation, sustainability, and environmental aspects with a more holistic approach (e.g., [36, 42]). This is also driven by increasingly competitive economies and more flexible organizational practices that require the ability to address challenges related to more dynamic scenarios

Table 1. Delivery Performance to Original Supplier Commit Date

Definition: The percentage of orders that are fulfilled on or before the internal Commit date, used as a measure of internal scheduling systems' effectiveness. Delivery measurements are based on the date a complete order is shipped or the ship-to date of a complete order. A complete order has all items on the order delivered in the quantities requested. An order must be complete to be considered fulfilled. Multiple line items on a single order with different planned delivery dates constitute multiple orders, and multiple planned delivery dates on a single line item also constitute multiple orders. **Calculation**: Delivery Performance to Original Supplier Commit Date = [Total number of orders delivered in full and on time to the scheduled commit date] / [Total number of orders delivered]

From https://scor.ascm.org/performance/reliability/RL.2.6.

and sustainability issues. In [42], a total of 29 indicators related to green IT infrastructures are reported (among which is the so-called "data center energy productivity (DCeP)" indicator, defined as the number of bytes that are processed per kWh of electric energy). The United Nations has devised a global indicator framework composed of 231 unique indicators to monitor the achievement of 17 established sustainability goals,[4] among which are "Proportion of total government spending on essential services" (education, health, and social protection) and "Tons of material recycled."

Let us notice from these examples two key features of indicators. On the one hand, (1) they are complex *derived* measures obtained by combining different data through a calculation formula; e.g., the DCeP is the *ratio* of bytes processed to total energy consumed. On the other hand, (2) they are *synthetic* measures, aggregated using sum, mean, or other operators, often segmented in relevant subgroups (e.g., time spent on unpaid domestic work segmented by sex, age, and location). Due to the first feature, the number of different ways to combine the available data in informative indicators easily becomes overwhelming. This can lead to information overload [13], making it crucial for managers to adopt a structured approach to selecting relevant PIs and designing a minimal and complete performance management system [26, 31]. As another issue, managers must clarify the meaning of an indicator to effectively communicate performance goals with stakeholders and ensure alignment with IT for implementation into usable PIs. Many catalogs, published as reference models, dictionaries, or libraries, have been proposed by researchers and public or private bodies to support the standardization and organization of numerous indicators. For instance, the **Supply Chain Reference Model (SCOR)**[5] includes the specification of relevant indicators. Table 1 shows the description of the indicator called "Delivery Performance to Original Supplier Commit Date." The level of detail of the specification is not uniform across the model. Examples of libraries are the UN indicator library,[6] the indicator libraries of the Canadian Institute for Health Information[7] and of the National Institute for Health and Care Excellence,[8] the collaborative (now dismissed) KPI Library project,[9] the OKRify library,[10] and the proprietary KPI Mega Library facility.[11]

From a technical perspective, the synthetic nature of indicators presents challenges associated with efficiently calculating aggregate values from large volumes of data and effectively supporting analytical functions. Data warehousing and **Online Analytical Processing (OLAP)** technologies have been conceived for this purpose [14]. OLAP systems organize data using a multidimensional

[4] https://unstats.un.org/sdgs/indicators/indicators-list/
[5] https://scor.ascm.org
[6] https://www.unepfi.org/impact/impact-radar-mappings/indicator-library/
[7] https://www.cihi.ca/en/access-data-and-reports/indicator-library
[8] https://www.nice.org.uk/Standards-and-Indicators/index
[9] http://kpilibrary.com
[10] https://okrify.com/kpi-library/
[11] https://www.kpimegalibrary.com

model, structuring it into data cubes, where each cell contains a (possibly materialized) aggregated measure indexed by a tuple of *dimension* values (e.g., time, geography, gender, product type). Dimensions are arranged in *hierarchies of levels* to support multiple aggregation granularities (e.g., a dimension "time" can include day, month, semester, and year as levels). This structure enables high-level operations like *drill-down* (viewing data at finer levels) and *roll-up* (aggregating data at coarser levels), facilitating flexible data analysis. Data Warehouse systems are general data management architectures designed to meet OLAP applications requirements, including historical data storage, integration of several internal and external data sources, data consolidation, and metadata management. In a Data Warehouse several measures can be collected together in a *fact*, providing different quantitative insights on a subject of analysis (e.g., quantity sold and total revenue for sales analysis).

Metadata, literally "data about data," describe the structure, meaning, relationships, and usage of stored data. They are crucial for data understanding, integration, quality management, and query optimization. To enhance interoperability, metadata are frequently represented through semantic technologies, employing standards and frameworks for formalized, structured data description.

2.2 Semantic Models

In the scientific literature, semantic technologies encompass methods, standards, and principles that enhance data representation, retrieval, and interpretation. These technologies enable structured, interconnected knowledge representation; promote data interoperability; and facilitate semantic understanding across various domains. They play a pivotal role in realizing the vision of data spaces that are not just accessible to humans but also comprehensible to machines, enabling more intelligent and efficient data utilization. These technologies are rooted in the vision of the Semantic Web [6], initially proposed by Tim Berners-Lee and his colleagues, which is an evolution of the World Wide Web, built upon the principles of encoding data in a way that is not only human readable but also machine understandable. Key technologies underpinning the Semantic Web include the **Resource Description Framework (RDF)** and ontologies, which provide a foundation for representing, connecting, and reasoning about data in a structured and semantically meaningful manner. RDF [65] is a fundamental data model that represents data as triples (subject–predicate–object) and serves as a universal format for linking and encoding data on the web, enabling the creation of knowledge graphs. RDFS, or RDF Schema [9], an extension of RDF, provides a basic vocabulary to describe data structure and semantics, introducing classes, properties, and relationships for defining simple ontologies. More complex ontological schemas can be developed by using the **Web Ontology Language (OWL)** [82], which supports the definition of concepts, relationships, and axioms within a specific domain, providing a shared vocabulary for knowledge representation and reasoning. Ontologies play a pivotal role in enabling systems to understand and process data semantically.

The application of these languages has led to significant advancements in the last 25 years in the fields of data management, knowledge representation, data integration, and query answering. The use of semantic technologies has enhanced data interoperability and facilitated more intelligent information retrieval systems. In the context of performance management, semantic technologies have been used since their foundation to model indicators, offering a more flexible and insightful approach to analyzing performance data. As discussed in the rest of the survey, semantic models are capable of supporting a higher level of interoperability, providing a common framework for expressing indicators and facilitating seamless data exchange and integration. The graph representation enabled by RDF enables richer metadata and semantic links between indicators and related entities, improving the understanding of relationships and dependencies for more comprehensive and accurate analysis.

By adhering to established standards, these languages promote consistency and reusability of KPI definitions while aligning with Linked Data principles to foster the creation of a global, interconnected web of data. This approach allows organizations to link performance indicators with external datasets, enabling deeper and more contextualized insights. Additionally, the formal schema representation supports SPARQL querying and reasoning to uncover hidden relations and dependencies.

3 Related Literature

As also discussed in Section 2.1, several frameworks of KPIs have been developed by institutional bodies and organizations with the purpose of supporting performance monitoring and data exchange within specific fields. However, these frameworks primarily focus on vertical application scenarios and often neglect the development of semantic models for indicators. As an initial phase of our research, we thus conducted an analysis of relevant literature, specifically examining reviews, surveys, and mapping studies related to the subject of performance indicator modeling.

Several articles investigated the use of indicators in various application domains. These articles share the objective of collecting specific sets of KPIs and describe their characteristics. We briefly discuss them to document what properties of performance indicators are typically considered as relevant. The authors of [32] carried out a systematic mapping study that proposes a classification scheme for KPIs in the software ecosystem domain by extracting information on the types of selected contributions. Among them, there are papers dealing with models of KPIs; assessment objectives, entities, and attributes used for measuring KPIs; and application scenarios. In [1], the authors review the existing KPIs proposed by reports/legislations/research that measure the performance and success in achieving goals in smart buildings. The article aims to collect and categorize existing KPIs. For some of them, the mathematical expression for their calculation, if agreed upon within the community, is provided. In [12], a survey on indicators for energy storage is proposed, starting from both the scientific literature and documents produced by public bodies and private associations. However, the work is mostly focused on collecting KPIs produced by different sources and comparing specific monitored values related to the domain at hand, instead of discussing modeling issues. In [81], the authors perform a categorization of business process performance measurements, proposing 11 performance perspectives, obtained by analyzing 76 documents. The article also documents a list of 140 KPIs, categorizing them into the proposed perspectives. In the domain of product–service system design, a comparative review is proposed by [51] by classifying KPIs into four general representative classes according to the different aspects to be measured. A taxonomy of specific measures is designed by [74], delineated according to the involved processes, what they measure, and whether they are quantitative or qualitative. Besides collecting and analyzing indicators for specific domains, however, none of the mentioned approaches extends the analysis also to representation models.

Some articles include a review of related literature, although they are not generally meant to provide a comprehensive evaluation of the field and therefore have a quite limited scope with respect to the present survey. Nonetheless, some of the dimensions considered for the analysis are relevant also for our work. Among them, the authors of [90] perform a comparison of proposals with regard to the modeling of KPIs and relative objectives, their mutual relationship, and assisted KPI derivation (reasoning functionalities). Five relevant characteristics for an appropriate representation of KPIs are identified in [68], including (1) proficiency in computational tractability, (2) clear and accurate syntax, (3) clear and accurate semantics, (4) stakeholder understandability, and (5) extensibility. A comparison among some models is proposed in [45], mostly focusing on expressions to calculate KPIs and related implementations (e.g., MDX). Furthermore, in [49] the authors aim to propose a unified specification for KPIs within the field of **Enterprise Architecture**

(EA) management. Fifteen approaches are compared against several dimensions, including the goal and the type of metric calculation. In [44] some KPI meta-models and dependency models are compared, along with a KPI analysis framework. Among the dimensions covered by the analysis, some proposed dimensions are also relevant for our work, including content-related aspects (e.g., whether the notion of goal or measurability is considered), the representation of dependencies among indicators, and whether the language of the model enables reasoning or allows semantic annotations. The comparison, however, is not meant to comprehensively review the state of the art and is focused neither on representation aspects, like in our work, nor on semantic models.

Comprehensive surveys are much less investigated, apart from a few examples focusing on models for KPIs with different goals and extents (e.g., [25, 48, 80]). In [48] a bibliometric analysis is proposed along with a classification model based on the method, the object and extent of analysis, and the level of granularity. In total, 23 models are analyzed from 2005 to 2015, 5 of which are semantic models that are described in more detail, considering an assessment of the methodological approach, reuse, reasoning functionalities, aim, and expressiveness. In the business model domain, a comprehensive survey [80] is proposed with the goal to support managers in selecting relevant approaches in the KPI management. Authors consider different types of contributions, including articles defining KPI models, methods (practices), and instantiations (i.e., implementations or prototype systems). In the work, the term “models” not only refers to the aspect of representation, as in our work, but also includes specific definitions of KPIs, catalogs of KPIs, and frameworks. Dealing with the business domain, further analysis dimensions include the type of support given by the approach within the management life-cycle phase (possibly including support to KPI calculation), the reference stage of use (whether the approach is used before or after the realization of the business model), the type of KPI considered (qualitative or quantitative), and the context of the approach (general or specific). The work most related to ours is [25], which proposes a survey aimed to derive a unified taxonomy capturing the overall characteristics of KPIs in existing work. In particular, it aims at enhancing the understanding of KPI management, helping users to select the most suitable solution for their requirements.

Starting from a motivation similar to ours, namely the lack of a common terminology and systematization of indicator frameworks, and a similar domain-independent approach, the survey analyzes 43 papers published in the last 10 years. The perspective focuses on the broader KPI management from specification to maintenance and includes simple taxonomies and categorizations along with semantic models such as ontologies. While some research questions are in common with our work, particularly regarding model expressivity, our study spans a broader 20-year period, focusing specifically on semantic models for KPIs. We also explore further dimensions, including document statistics, application domains, model goals, formalization, and technical aspects like serialization formats and model documentation.

From the analysis of the scientific literature, to the best of our knowledge, no review has been produced, in the form of a survey or systematic mapping study, with the purpose of organizing the work published on semantic models for performance indicators. Consequently, our current undertaking is driven by the necessity and opportunity to document and categorize the research conducted in the past two decades on this subject.

4 Methodology of the Study

The goal of this survey consists in the identification and classification of the semantic approaches aimed to define a representation model for metrics or performance indicators, their purpose, and their practical usability. This study has been developed following the steps identified in state-of-the-art methodological guidelines [43, 59, 77]. We discuss such steps in the following, by grouping them into two phases, planning and conducting the study, as also summarized in Figure 1.

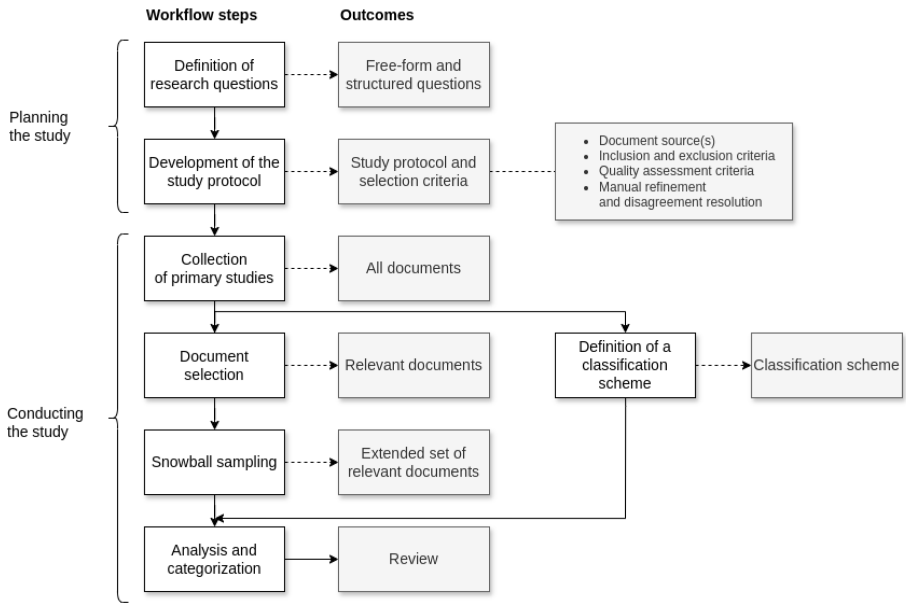


Fig. 1. Methodology workflow.

The essential process steps of the *planning phase* include the definition of the research questions and the development of the study protocol. The latter normally starts with the identification of the documents, i.e., the set of primary studies on which the analysis is performed. They can be identified in several ways, typically by means of a search on scientific databases or by browsing the relevant literature. The approach taken in this work consists in (1) a database search followed by (2) snowball sampling. *Database search* is the most adopted approach in reviews and mapping studies, since it is more structured and enables reproducibility of results to a larger extent. It consists in choosing a set of relevant repositories and executing a query string to retrieve the documents. However, database search is not always capable of identifying all the relevant literature. For this reason, we performed *snowball sampling* as a further step. This refers to using the list of references of a paper or the citations to the paper to identify additional papers that are worthy of interest. According to results reported by [59] and [87], conducting a snowball search can be highly beneficial in identifying relevant sources that may belong to different research communities or may not fall within the chosen set of keywords. Further steps in the planning phase include the definition of criteria for exclusion of documents and for assessing their quality. Finally, a manual screening of relevant metadata of selected documents is performed by multiple reviewers, which can require the resolution of possible disagreement among them.

The second phase, aimed at *conducting the study,* consists in the actual gathering of the documents from sources, the application of exclusion criteria, the manual refinement in order to obtain a validated list of documents, and the definition of a classification scheme. The definition of this scheme, which is then used to perform the final analysis, has followed a methodology inspired by Nickerson et al. [54].

4.1 Definition of the Research Questions

The definition of the research questions stems from the identified research goal and from a *deductive* approach [54] based on the authors' experience on the topic, existing surveys, and reviews on the topic (as reported in Section 3). At first, the goal involves a focus on "semantic approaches" for the representation of performance indicators. As such, the survey will investigate how information is represented in the proposed models and the specific role of semantics. Furthermore, the survey will also investigate the "purpose" of such models along with their "practical usability." In the study, we refer to the following **research questions (RQs)**:

— RQ1: What are the main approaches for semantic modeling of indicators in the literature?
— RQ2: What is the expressivity of such models in terms of represented concepts and level of formalization?
— RQ3: What advantage can formal approaches and reasoning capabilities offer?
— RQ4: What are the objectives of the models, and in which application domains are such approaches utilized?
— RQ5: What is the level of practical usability and reusability of such models?

4.2 Development of the Study Protocol

The study protocol involves the definition of the source on which to perform the database search and the criteria for selecting and assessing the quality of the retrieved documents.

4.2.1 Document Sources. In this work, we relied on Scopus[12] and Web of Science[13] as sources of primary documents. To define the query string, the authors of [43] suggest to consider keywords stemming from research questions by analyzing them in terms of population, intervention, comparison, and outcome of the studies. To obtain a broader overview of the field, here we avoided terms that could have narrowed the search results too much. As a result, we considered two query conditions on title, abstract, and keywords, the former broader and the second more focused. The query conditions were defined as follows and submitted focusing on documents published up until 2023:

Q1) `("KPI" OR "indicator") AND ("model" OR "representation") AND ("semantic" OR "ontology")`
Q2) `("performance indicator") AND ("ontology" OR "metamodel")`

Starting from the documents collected by executing the queries, study selection was performed by excluding a number of documents according to a set of criteria, followed by a quality assessment as reported below. Finally, a manual refinement was needed to filter out documents that were not relevant for the goals at hand.

4.2.2 Exclusion Criteria. Exclusion criteria are applied to selected documents to refine and better focus the scope of the research. The following criteria were applied:

— The research is not within the focus area.
— The documents are outside the field of computer science. This constraint is implemented, e.g., in Scopus, by adding the condition "`LIMIT-TO (SUBJAREA, "COMP")`" to the query string.
— The document is an editorial. This is implemented, e.g., in Scopus, by adding the condition "`EXCLUDE (DOCTYPE, "ed")`" to the query string.
— The document belongs to gray literature.
— The study duplicates other studies.
— The document is not written in English. This is checked by adding a specific constraint to the query string, e.g., "`LIMIT-TO(LANGUAGE, ‘‘English’’)`" in Scopus.

[12] https://www.scopus.com/
[13] https://www.webofscience.com/

4.2.3 Quality Assessment Criteria. The set of selected papers was further refined by considering some criteria concerning impact and quality, which take into account (1) citation metrics as a measure of the impact of the publication and (2) publication venue. In particular, the following three conditions for selection were applied. Papers satisfying at least one criterion were selected:

— Average number of citations per year: Documents were sorted in descending order based on the number of citations they have received. Papers belonging to the first quartile were selected.
— Journal rank: The publication venue for journal papers was evaluated by considering the **SCImago Journal Rank (SJR)**,[14] a numeric value representing the average number of weighted citations received during a selected year per document published in a given journal during the previous 3 years, as indexed by Scopus and Web of Science. Papers belonging to quartiles 1 or 2 were selected for this study. Taking a conservative approach, if the journal name for a paper did not match any entry in SJR, it was selected for further analysis.
— Conference rank: The publication venue for conference papers was evaluated by using the CORE Conference DB,[15] which ranks conferences from A* (top conferences) to C. We selected documents published in conferences proceedings ranked from A* to B. Papers published in conferences that are not reported in CORE were selected for further analysis.

Regarding recent documents from the year 2023, due to the unreliability of using the number of citations as a quality criterion given their recent nature, all documents were chosen for manual refinement.

4.2.4 Workflow. Once documents were collected from the sources, they were evaluated against the inclusion and exclusion criteria, duplicated documents were removed, and a following quality assessment was performed. Then, titles, abstracts, and keywords (whenever available) were considered for a manual evaluation aimed at refining the study selection. This step was needed in order to filter out spurious documents that had been selected because they incidentally include the search terms although not dealing with the research topic. The process was performed by two authors, who classified each paper as "relevant" (score 1.0) if the paper was clearly aligned with the research goal, "partially relevant" (score 0.1) if it was only partially relevant with the research goal, and "not relevant" (score 0.0) if the document was out of scope. In some cases, a full-text reading was performed to disambiguate specific cases when the evaluation of title, abstract, and keywords was unclear. As a selection criterion the following disagreement solution was considered: only documents with a score greater than 1.0 were selected, meaning that they had been evaluated as at least relevant by one researcher and partially relevant by the other, without any "not relevant" evaluation.

The filtered list of articles was considered to apply snowball sampling by looking for relevant papers in the reference list or papers citing some selected documents. The quality of such additional documents was directly manually assessed before their selection in the list. Finally, documents were manually analyzed in order to confirm their relevance to the research goal and to detect relations between documents, such as *inclusion*, if a document is an extended abstract of another, or *extensions*, if a document is an earlier version of another from the same authors. Included documents were left out from the analysis, while earlier versions were retained only in case the described model differed from the one presented in the extended version.

A classification scheme including a number of dimensions of analysis was manually drafted following a methodology inspired by [54], as discussed in Section 4.4, and then exploited to analyze and categorize the papers.

[14] https://www.scimagojr.com/
[15] http://portal.core.edu.au/conf-ranks/

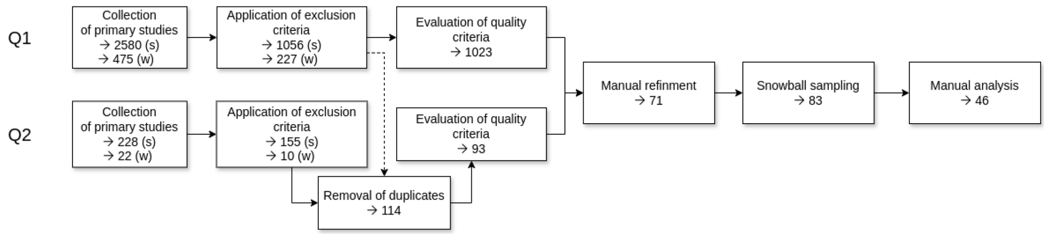


Fig. 2. Extraction of documents from Scopus (s) and Web of Science (w) and their selection.

Details on the number of papers included and excluded at each step are reported in the next subsection.

4.3 Collection of Primary Sources and Document Selection

In this subsection, we describe in detail the data extraction workflow for the two queries reported above. As summarized in Figure 2, queries Q1 and Q2 were submitted to Scopus and Web of Science,[16] resulting in a number of results equal, respectively, to 2,580 (Scopus) and 475 (Web of Science) for Q1, and 228 (Scopus) and 22 (Web of Science) for Q2. As for Q1, by considering exclusion criteria, we reduced the list to 1,056 (Scopus) and 227 (Web of Science) documents. The other 260 documents were filtered out by applying quality criteria and removal of duplicated documents, resulting in 1,023 papers. As for query Q2, the list was reduced to 155 (Scopus) and 10 (Web of Science) documents by taking into account exclusion criteria. In order to speed up the process, we then removed duplicated documents that had been previously retrieved by Q1, resulting in 114 documents. Finally, 93 papers satisfied quality criteria.

The two lists of selected papers were then merged, resulting in 1,116 documents. The manual refinement involved assigning a score based on the relevance of the paper to the research goal. The resulting set consisted of 71 documents. During this step, a number of additional references were collected through snowball sampling both by looking at the reference list of particularly interesting papers and by checking the list of papers citing them. As a result, a further 12 papers were included in the list, summing up to 83 documents.

Finally, the manual analysis allowed us to remove documents that were not aligned to the research goal and identify included and extended documents, resulting in a total of 46 articles that were selected for the study.

4.4 Definition of a Classification Scheme

The classification scheme is the set of dimensions, related features, and corresponding values that have been used to describe the selected documents. As mentioned in the previous subsection, the scheme was defined following a methodology inspired by [54], based on an iterative approach aimed to identify and structure the most relevant dimensions. In our case, the process started from the identified research questions, from which we drafted the corresponding dimensions of analysis. The iterative step was aimed to structure the dimensions in features and corresponding values. We followed both a *deductive* approach, by analyzing existing taxonomies (see Section 3), and an *inductive* approach, based on the analysis of the selected documents, which allowed us to identify relevant aspects to consider. This helped in iteratively constructing and refining the features included in the dimensions. The stop conditions included *subjective* and *objective* aspects.

[16] Queries were submitted in September 2024. All numbers reported in the subsection refer to documents available at that time on the platform.

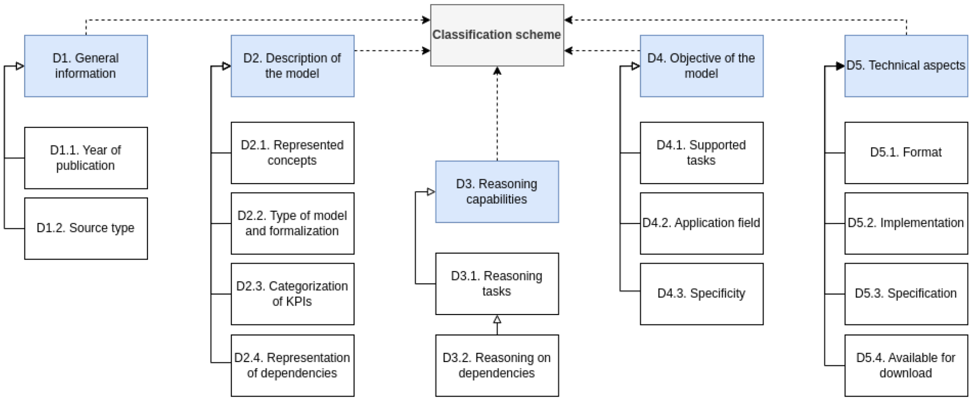


Fig. 3. Classification scheme.

Among the former we considered conciseness, robustness, comprehensiveness, extendibility, and explainability. Conversely, objective stop conditions are such that dimensions are mutually exclusive and include collectively exhaustive features. All documents have been examined, and no changes are made in the last iteration.

The classification scheme includes 15 features arranged in five dimensions as summarized in Figure 3. We report dimensions in the following, specifying the possible set of values for each feature.

A first set of features is aimed to describe *General information (D1)* on editorial aspects of the papers, including year and source of publication:

— *D1.1. Year of publication*: {2003|. . .|2023}
— *D1.2. Source type*: {article | conference paper}

This, and the whole survey itself, provides answers to RQ1.

The following features are related to information on the *Description of the model (D2)*, identifying what concepts are represented, the degree of formalization adopted, whether a taxonomy of PIs (or a similar categorization) is included or not, and whether the model also represents dependencies among indicators:

— *D2.1. Represented concepts*: {indicator, dimension, business objective, unit of measurement, goal, formula, aggregation, other} (multiple options)
— *D2.2. Type of model and formalization*: {ontology | taxonomy | metamodel | other}
— *D2.3. Categorization of KPIs*: {full taxonomy | partial categorization | no categorization}
— *D2.4. Representation of dependencies*: {no dependency | dependency relations | mathematical expressions (formulas)}

Collectively, these features address the complex and multi-faceted topic of RQ2.

Somewhat related to it, RQ3 is addressed by a further set of features, which investigate the availability of *Reasoning capabilities (D3)* and whether they are applied on dependency relations among indicators:

— *D3.1. Reasoning*: {no reasoning | classic reasoning | advanced reasoning}
— *D3.2. Reasoning on dependencies*: {no | partially | yes}

Some features are identified to describe the *Objective of the model (D4)*, in terms of its level of specificity, application field, and the specific usage for which the model has been proposed:

— *D4.1. Supported tasks*: {management, documentation, validation, monitoring, advanced analysis, other} (multiple options)

— *D4.2. Application field*:* (not constrained)
— *D4.3. Specificity*: {general-purpose | application-oriented | mixed}

These features aim to provide answers to RQ4.

Finally, a set of features focuses on *Technical aspects (D5)* of sharing and reusability, including the format, implementation details, specifications, and download links, providing answers to RQ5:

— *D5.1. Format*: {OWL | RDFS | other format | unspecified}
— *D5.2. Implementation*: {not reported | not available | partial | yes}
— *D5.3. Specification*:{not reported | partial | full}
— *D5.4. Available for download*: {no | yes (upon request) | yes (but currently unreachable) | yes}

4.5 Validity Evaluation

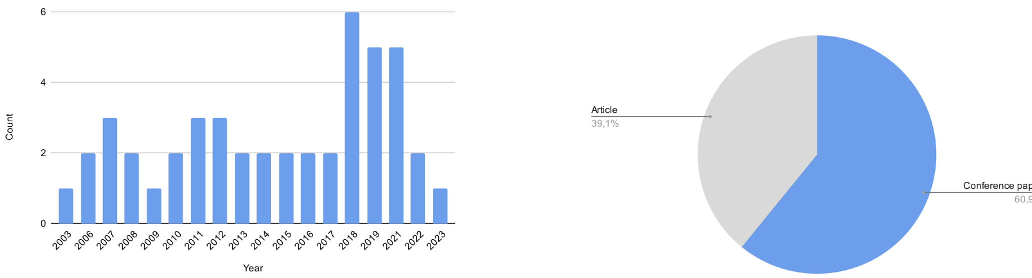
The methodological process described in the previous subsections can be evaluated in order to assess the validity along multiple perspectives [59].

4.5.1 Theoretical Validity. It involves the capability to correctly capture what was intended and to avoid confounding factors and biases. In order to reduce these possible sources of issues, the steps involving forms of subjective evaluation, namely the definition of the search queries and the manual refinement, were conducted by more than one researcher. This also limited the risk of including not relevant documents, which would have been more likely to occur by solely relying on an automated extraction and reporting process. Furthermore, to reduce the threat of overlooking relevant papers that may not have been retrieved by the queries, snowball sampling was performed. This resulted in an increased number of articles that were added to the selected list. During the analysis and classification steps, researcher bias also could be a threat. For such a reason, one researcher was in charge of the extraction, while the other two analyzed the extracted articles. Finally, a researcher who was not previously involved in prior work on the topic contributed to snowball sampling. This helped in reducing the selection bias from the same community, although a complete removal of the threat is generally considered not to be feasible, also according to shared guidelines (e.g., [8]).

4.5.2 Interpretative Validity. The interpretation of the review's outcome is valid if the conclusions are reasonable given the data. In this case, although some of the authors of the present document co-authored papers on the object of this survey, the adoption of objective criteria in some of the steps, e.g., definition of queries, exclusion criteria, helped in reducing the interpretative bias. Furthermore, the experience of such co-authors on the object of the study provided help in the interpretation of the data.

4.5.3 Repeatability. In order to guarantee repeatability, detailed reporting of the process has been provided in terms of the workflow followed for this survey. Repeatability was also aided by the use of guidelines.

4.5.4 Validity of the Classification Scheme. Different perspectives can be adopted to validate the classification scheme. On the one hand, subjective stop criteria for its definition [54] include a number of aspects, namely conciseness, robustness, comprehensiveness, extendibility, and explainability. In particular, the scheme contains a *concise* number of dimensions (i.e., five), features (from 2 to 4, for a total of 15), and possible values for constrained features (from three to eight), in accordance with the guidelines. These numbers ensure the scheme's clarity and ease of application. As for *robustness*, the scheme clearly distinguishes various aspects related to KPI modeling. Moreover, by taking into account a significant number of relevant documents, the scheme addresses a *comprehensive* set of aspects involved in semantic models for KPIs. Although the work focuses on



(a) (b)

Fig. 4. (a) D1.1 Year of publication. (b) D1.2 Source type.

formal aspects, objectives, and applicability of the model, *extensions* to the analysis could introduce new dimensions or features. Finally, being derived from research questions, each dimension is defined in a *self-explanatory* manner, which enhances its clarity and potential use.

We also compare our classification scheme with those used in the surveys most similar to ours [25, 48, 80]. In particular, in [48] the proposed dimensions include (1) a bibliometric analysis, which is mostly aligned to our dimension D1; (2) the broadness and the level of granularity of the analysis, which is partially covered by D2.1; (3) the methods and methodological approach, in terms of representation language and type of evaluation, addressed by D5; and (4) the level of formalization and of complexity of the models, which is discussed by D2.2 and D2.4. Differently from [80], we focus only on semantic models that are a subset of the considered types of approach. Also, the support offered during different KPI management lifecycle phases/stages is partially out of our scope, although the "operationalization" phase (including formulas or relations among KPIs) and the "measurement" phases are respectively covered by D2.4 and D4.1. Finally, the context (specific vs. general) is covered by D4.3. In [25], a comprehensive analysis is conducted for the development of the taxonomy, based on [54]. While their scope is broader than our focus on semantic models, their KPI analysis, which includes perspectives and context, partly relates to D2 and D4.2. As for the features considered in the specification of KPIs, they are partially covered by D2.1 and D2.4. Although they focus on a larger set of proposals and not only semantic modeling, formal specification is covered by D2.2 and D5. Finally, their dimension about characteristics of the KPI management approach is partially covered by D5 (for syntax, format, and implementation) and D3 (reasoning). In conclusion, our scheme is mostly aligned with their core dimensions regarding what is represented and its purpose. On the other hand, it uniquely emphasizes formal representation models, reasoning capabilities, specific supported tasks, and aspects of reuse and sharing.

5 Results

In this section, we summarize the results obtained by analysis of the selected documents along the identified classification scheme.

5.1 General Information

The time span of the papers (D1.1) ranges from 2003 to 2023, with an increase in the average number of papers per year over time, going from one document (e.g., 2003) to six documents (2018). The distribution is shown in Figure 4(a). As such, the earliest publications on the topic are temporally aligned with the rise of the research interest in the "Semantic web," while the trend in recent years (2018–2021) may witness the increasing interests of semantic models for KPIs in a variety of domains. As for the type of source (D1.2), as shown in Figure 4(b), the majority of

Table 2. Represented Concepts and Level of Formalization (D2)

Document	Dimension	Aggregation	Goal	Process	Formalization	Categorization	Dependencies
[17, 56]	-	✓	✓	✓	✓	-	✓✓
[78]	✓	-	-	-	✓	✓	-
[10]	✓	-	✓	-	✓	-	✓
[39]	-	-	✓	✓	-	✓	-
[85]	-	✓	-	✓	-	-	✓✓
[73]	-	-	-	-	✓	✓	✓
[19, 20, 67]	✓	✓	✓	✓✓	✓	✓	✓✓
[28]	✓	✓	✓	✓✓	-	✓	✓✓
[29]	✓	✓	✓	✓✓	✓	✓	✓✓
[16]	✓	-	✓	-	✓	-	-
[71]	✓	✓	-	-	✓	-	-
[52, 66]	✓	-	-	-	✓	-	✓✓
[2]	-	-	-	✓	✓	-	✓
[72]	✓	✓	✓	-	-	✓✓	✓✓
[61, 62]	✓	✓	✓✓✓	✓	✓✓	-	✓
[3, 7]	✓✓	✓	-	-	✓	✓	✓✓
[38]	-	-	✓	-	✓	-	-
[70]	✓	-	✓	-	✓	✓✓	-
[18]	✓	-	✓✓✓	-	✓	-	✓
[46]	✓	-	✓✓	-	✓	✓	✓✓✓
[60]	✓	✓	✓	-	✓	-	✓✓
[47]	-	-	-	-	-	✓✓	✓✓
[33, 84]	✓	-	-	-	✓	✓✓	✓✓
[34]	✓	✓	-	-	✓	✓✓	✓✓
[57]	-	-	-	-	✓	-	✓
[91]	✓	✓	-	-	✓	-	✓✓✓
[53]	✓✓	✓	-	-	✓	-	-
[63]	✓✓✓	✓✓	-	-	✓	-	-
[58]	-	✓	-	✓	✓✓	-	✓✓
[88]	✓✓	✓	-	-	✓	-	✓✓✓
[44]	-	-	✓✓✓	✓	✓	-	✓✓✓
[30]	✓✓✓	✓	-	-	✓	-	-
[79]	✓	-	-	✓	✓	-	-
[22]	-	✓	-	-	✓	-	✓✓✓
[23]	✓✓	✓	✓✓	-	✓	-	✓✓✓
[5, 21, 24]	✓✓	✓	✓✓	-	✓	-	✓✓✓
[4]	-	-	✓	-	✓	✓✓	✓

Documents are marked with a dash (-) if the concept is not modeled, or with one to three ticks (✓) depending on how detailed the concept is modeled.

papers (60.9%) were published in conference proceedings, while more than one-third of them were published as articles in journals (39.1%).

5.2 Description of the Model

Several models have been proposed over the years for the semantic representation of indicators. In the following, we compare the approaches with respect to several aspects regarding the representation model. These aspects encompass the conceptual elements employed to describe indicators (D2.1), the degree of formalization within the model (D2.2), the presence of a categorization for indicators (D2.3), and their approach to representing dependencies among indicators or computation formulas (D2.4). Overall, this provides an idea of the level of expressiveness of the models, which is qualitatively summarized in Table 2 and graphically represented in Figure 5, while a detailed discussion for each feature is reported in the following.

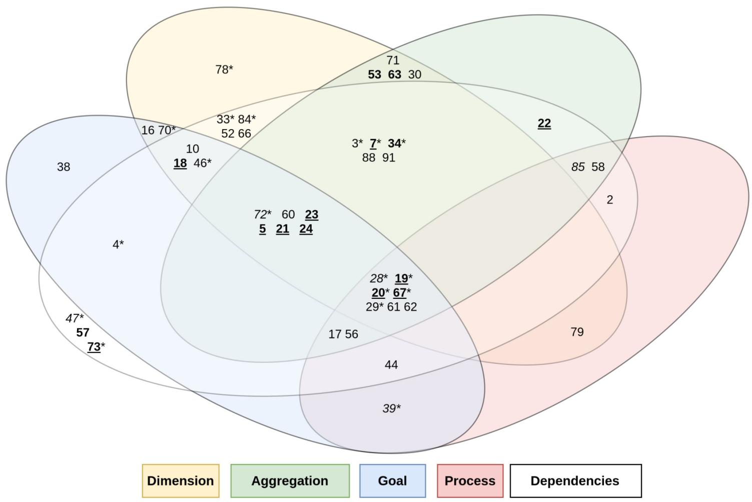


Fig. 5. Overview of represented concepts in the selected articles (dimensions D2.1, D2.4). Articles not providing a formalization are in italics (D2.2), while those categorizing KPIs are marked with an asterisk (D2.3). Articles exploiting reasoning tasks are shown in bold (D3.1), while those specifically using reasoning for mathematical manipulation and dependency analysis are underlined (D3.2).

5.2.1 Represented Concepts. Most of the articles focus on the notion of indicator, although some of them make a distinction between an *indicator* and a *metric*. For example, in [56] (and previous work thereof [17]), a metric is defined through the measurement or calculation method and the measurement scale, while an indicator is a new mapping obtained from the interpretation of the metric's measured value of an attribute into a new variable. For instance, authors exemplify the approach in the framework of e-learning activities, where a metric is the Task Completeness Ratio (TCR). TCR has a formula defined as $TCR = \#CT/\#PT$, where both $\#CT$ and $\#PT$ are metrics ($\#CT$ = "number of completed tasks," $\#PT$ = "number of proposed tasks"). In this example the scale is of type ratio, and with real values. On the other hand, the indicator Task Completeness Performance Level (TC_PL) is introduced, whose specification (elementary model) is $TC\_PL = TCR * 100\%$ if $X_{min} \leq TCR \leq 1$, and $TC\_PL = 0\%$ if $TCR \leq X_{min}$, where X_{min} is some agreed-upon lower threshold (e.g., 0.45). In this sense, the indicator interprets the value of the metric as a satisfactory or non-satisfactory task completion percentage. Also in [85], metrics can be composed using arithmetic and logical operators, while a KPI consists of an aggregate metric definition, which computes the actual value in a certain analysis period. In [44], instead, a KPI is distinct from the notion of metric, which can be composite and provided with a formula, although a relation exists between the two. Some other articles make a distinction between measures, or simple indicators, and indicators, further distinguishing between KPIs and **Key Result Indicators (KRIs)**, which evaluate the actual satisfaction of a goal and have a set time to meet their target, e.g., in KPIOWL [18], following the modeling approach by Horkoff et al. [35].

In most approaches, indicators are described through their name and a possible textual description. As detailed in the following, further properties are defined in the majority of the articles concerning the notion of *dimensions* along which an indicator can be measured and its *aggregation* functions. In several cases, the *goal* of the monitoring is represented, as well as a *process* perspective providing an alignment between indicators and process activities.

Dimension. The formalization of dimensions within the model exhibits a range of approaches, each tailored to the specific context. For instance, in [10, 91], dimensions are not explicitly formalized as a class of the model but only appear in the definition of the measurement method or formula. To provide an example, in [10] the GDP measurement method is defined by referring to concepts *Year* or *TimeInterval* of the SUMO upper ontology [55]. In some approaches, dimensions are tailored for the specific domain. For instance, within the domain of smart cities, in [16] the city dimension allows to specify a number of entities related to the indicator (e.g., geographical areas, buildings), while in [34], a class of indicators is related to time, place, and city service in the context of a city indicator ontology. As for the industrial domain, in [66] a KPI is related to the industrial equipment that is being measured and is characterized by an attribute specifying the scope, the group in the organization, and the production methodology, while in EM-KPI [46] the dimension, besides the time, is the infrastructural element for which the KPI is defined. Following the representation of process-centric performance indicators, [60] relates PPIs to activities or products. Similarly, WfMetricOnto [79] establishes a connection between metrics and workflow elements, with varying granularities, ranging from the generic workflow down to individual activities and invoked applications. In [61, 62] dimensions are identified by the agent/role and the process associated to the indicator. In PPINOT [19, 20] as well, dimensions are not explicitly represented, but they are embedded in the definition of indicators, e.g., for what concerns the time period and the process. The model is then extended with the resource dimension in [67]. A similar approach for the target time is adopted by [18].

A subset of articles explicitly defines specific dimensions within the model, e.g., in [70, 71] (for the geographic dimension) or in NPIOnto [72], which includes the scope of the KPI (regional/national/global) and the category (economic, social, environmental issue) among others.

In the context of OLAP and multidimensional analysis of indicators, some articles employ formal definition for dimensional hierarchies. For instance, in [53] a graph of acyclic relations is used to formally define dimensions within a Datalog-based Multidimensional Ontology. QB4OLAP [30] extends the RDF Data Cube [83] with classes and properties to fully represent dimensional hierarchies. In EIAW [88], dimensions and their hierarchies can be flexibly defined from domain terminology. Likewise, in KPIOnto [23] dimensions are defined together with a hierarchy of levels and corresponding members adopting the OWL-DL language. The approach is echoed in the multidimensional ontology proposed in [3]. Similarly, [63] adopts the OWL-DL language to formalize a number of properties related to dimensions.

Comparing the different modeling strategies, we observe that the major difference lies in the extent to which the reification design pattern is applied. In [88], a dimension is a virtual property whose range is an enumeration of nodes organized as a tree-like hierarchy. This does not support the representation of multiple hierarchies. All other models represent dimensions and levels as distinct, related classes. In [23, 53] a reflexive association between levels (called *ancestor* in [53] and *rollup* in [23]) is introduced to represent the hierarchical structure. While [23] links the *Level* and *Dimension* classes directly (by the *inDimension* association), [53] defines a class *Hierarchy* as *composed* by levels,[17] with a one-to-many association with the class *Dimension*. Thus, the model separates the notion of dimension from the definition of its structure, allowing to reuse the same hierarchy to define different dimensions. In contrast, in [23] a dimension is intrinsically defined by its structure. A different approach is taken in [63], where the rollup association is reified in a *RollUp* class. In this approach, a class *Hierarchy* is also introduced, but with a meaning different from [53], representing the possible branches a dimension may have (for instance, a time dimension can have a month level that rolls up both to a quarter and a 4-month period; these are considered two different hierarchies of the dimension). A dimension is then defined as a composition of branches.

[17]Following the definition of composition in UML as a special kind of association where parts depend on the whole, graphically represented by a line connecting classes that ends with a filled diamond.

The relationship between a dimension and its levels (*DimensionLevel* class) is then indirectly established by linking the class *Hierarchy* to *RollUp* (*HierarchyToRollup* association) and *RollUp* to *DimensionLevel* (*isSourceOf*/*isTargetOf* associations). This reification design pattern is necessary to elicit some properties of the hierarchies, especially considering that the model does not include the concept of instances of a level, or level members (e.g., January 2024 could be a member of the level month of a time dimension). For instance, drilling down incomplete hierarchies or non-strict hierarchies [50] can be characterized on the basis of the values of multiplicity of the *isSourceOf* and *isTargetOf* associations. On the other hand, in [23, 53] a similar characterization requires the definition of proper assertions on levels and their members. The approach in [30] is even more extreme in terms of reification, introducing a class for a large part of relevant associations. It is the case, for instance, of the *HierarchyProperty*, *LevelProperty*, *LevelInHierarchy*, *HierarchyStep*, and *RollupProperty* classes. This makes the model much more verbose. On the other hand, in this way, it is possible to directly specify the correct structure of a dimension, in terms of constraints between levels and their dimension, or between levels and related members. Again, the same kind of capability requires the definition of assertions in [23, 53], e.g., an *inTime*-specific relation can be built specifying the association between the month and year levels. On the other hand, a reflexive rollup association, where the class *Level* is both the source and the target, would allow to link a specific instance (e.g., month) to a level belonging to a different dimension (e.g., to nation). To avoid these cross-dimensional rollups, a rule should be defined to constrain rollup operations, ensuring that they only occur between levels within the same dimension.

Aggregation. A representation of classical aggregation functions, e.g., sum, min, max, average, is provided in some articles, e.g., [60, 71, 88]. Some approaches provide a distinction between indicators and aggregated measures that can be calculated through aggregation functions; e.g., this is the case of [53] and COBRA [58]. As for the latter, the class of *Aggregation Metrics* is defined in terms of an aggregation construct (an operator) that is applied over the population obtained by applying a *Population Filter*. In [91] an aggregation function is specified for the calculation formula of a KPI, while in [34] the aggregation function is applied on a population.

Similarly, in the context of process analysis, a specific class for aggregate metrics is defined in [85], where a distinction is made between *instance metrics,* which measure the duration between two activities in the process, and *aggregate metrics*, which evaluate the performance of a process in a certain analysis period by aggregating values of instance metrics through classical aggregation operators. A similar approach is taken by PPINOT [19, 20, 28, 29, 67], in which the value of an aggregate measure is calculated by applying a certain aggregation function on a set of measures (belonging to different process instances) to obtain one single value. As an extension, in [28] an object aggregate measure is introduced to apply an aggregation function on a base measure but making reference to a different reference object (e.g., the sum of orders whose number of items is greater than a threshold).

In [63], the classical multidimensional model is extended with the notion of aggregation function, aggregation type, and summarizability of measures. As an extension of the RDF Data Cube [83], in QB4OLAP [30] aggregation functions are associated to a measure. Similarly, in KPIOnto ([23] and related work), an indicator is associated to an aggregation function. The same approach is taken by [3], which is inspired by KPIOnto. It is interesting to note that [63] provides a very detailed model of the summarizability concept introducing distinct notions of summarizability of measures along facts, dimensions, hierarchies, and rollups. Axioms allow to specify the meaning and relations among the different kinds of summarizability; for instance, summarizability of a measure along a fact implies summarizability along every dimension of that fact, summarizability along a dimension implies summarizability along every hierarchy of the dimension, and so on. The model is not strictly limited to the sum aggregation function, and other aggregation types (average, max, min) are considered as well. This level of detail enables the definition of a set of rules to assess the correctness of a multidimensional schema, like strict summarizability (which exploits properties of hierarchies discussed before) and level-by-level summarizability (which establishes that a measure can be correctly aggregated from the bottom to the top of a hierarchy). The latter property cannot be defined with other models.

On the other hand, in NPIOnto [72] the type of indicator, among which is the aggregate one, is represented only as a dimension. In [61], multiple instances of the same indicator can be defined for each aggregation level. A relation *aggregation_of* can then relate a couple of indicators to state that the former is an aggregation of the latter.

In [56] and related work, global indicators are associated with various aggregation models, such as numeric or linguistic, depending on the specific context and requirements.

Goal. The representation of goals within the context of indicator modeling assumes diverse forms. A first difference can be established according to the *subject-orientation* principle, that is, whether goals or indicators are the primary subject of the model. In fact, in the former case indicators are considered tools of goal management, while in the latter the goals are seen as mere properties of indicators. Considering the latter category, goals are represented as an attribute of a metric in [19, 29, 56] or implicitly embedded within the categorization of indicators, e.g., in [4, 10, 16, 60, 70]. Being designed for a specific application domain, categories are statically predefined. For instance, in the environmental sustainability domain, [16] incorporates explicit descriptors to indicate whether the indicator targets compliance, cost, quality of life, or quality of service, therefore providing a clear representation of the intended goal. [10] introduces a taxonomy of "themes" to specify the particular objective that an indicator serves, like "Material and energy consumption," while [70] inherits the taxonomy of themes from ISO 37120. In [39], the indicator is connected to the corresponding strategy, representing plans and programs that are employed toward the enterprise goals. In NPIOnto [72], goals are indirectly represented, with indicators linked to corresponding definitions in various indicator frameworks, which specify goals for national development. Similarly, in [38], an indicator is related to a target that is linked to a goal.

On the other hand, some articles provide a more explicit representation of goals. In EM-KPI [46], each indicator is associated to a performance goal, which in turn is linked to a relevant stakeholder. In KPIOnto [5, 23, 24] as well, an indicator is related to the corresponding business goal for which it is commonly used. In the extension [21] in the sustainability domain, the KPI is associated to the corresponding **Environmental, Social, or Governance (ESG)** goal. In these models, the capability to model and exploit the notion of goals is limited, providing mainly a standardization and categorization of terms and support for a common understanding, communication, knowledge sharing, and interoperability.

For what concerns articles with a specific focus on goals, a notable case is presented in [62], which extends [61] by formally formulating goals over goal patterns, i.e., Boolean properties for an organization, unit, or individual at a certain time point/period, that are associated to an indicator. Goals can be achieved through a set of tasks of a process and may be classified as hard goals if their satisfaction can be quantitatively established or, conversely, as soft goals. Rules for refining goals within a hierarchy of subgoals are defined, as well as relations among soft goals. Similarly, KPIOWL [18] follows the strategic modeling language introduced in [35], and a full decomposition of goals in subgoals is proposed. Goals in this framework are explicitly related to one another through relations of contribution and decomposition, allowing for a detailed representation of their hierarchical structure. A similar approach is proposed by the OWL-Q KPI Extension [44], where goals can be decomposed into subgoals through AND/OR relations. Richer analyses are enabled by these kinds of approaches, e.g., verifying goal satisfaction at strategic, tactical, and operational levels.

Process Perspective. There is no consensus in the literature regarding the relationship between KPIs and PPIs. Some authors do not establish any difference between them (e.g., [61]), while others consider PPIs as a particular case of KPIs, i.e., process-related KPIs (e.g., [20, 60, 85]). In some cases, they can be considered as focusing on different levels; i.e., KPIs are more closely associated to the tactical and strategic level, while PPIs are more aligned with the operational level (e.g., [69]).

In the analyzed articles, indicators are in some cases connected to the entity representing a process through properties. For instance, this happens in [58] and [56] and previous work thereof (where it is a generic entity) or in several other approaches where the indicator monitors/measures a process (e.g., [2, 39, 44, 61, 79, 85]). In most cases, indicators can be associated with control flow and timing aspects of processes, with limited traceability. Sometimes a higher expressiveness is enabled by the possibility to link performance indicators to process data and to a finer-grained process perspective, i.e., process elements. This is the case of PPINOT [19, 20, 29, 67], which explicitly aims at defining PPIs. The scope is used to select the process instances that must be considered to compute the PPI value (either considering every existing process instance or by means of a scope template).

5.2.2 Type of Model and Formalization. In the realm of data modeling, a common practice adopted by the majority of the selected articles is providing a graphical representation in the form of a UML schema or similar formalisms, often taking the shape of **Entity-Relationship (ER)** diagrams. This graphical approach serves as an intuitive means of communicating complex structures and relationships within a system. However, the landscape of system modeling is diverse, and there are exceptions where different formal languages and representation methods are employed to articulate the intricacies of a system's structure and behavior.

A notable alternative to graphical representations is using languages rooted in first-order logic. In particular, since the aim of the approaches is the conceptual modeling of a domain, it is natural to resort to languages specifically introduced for this purpose. In particular, **Description Logics (DL)** offer a formal, logic-based framework for specifying and reasoning about the properties and relationships within a system, providing a mathematically rigorous foundation for modeling with a good tradeoff between expressivity and complexity, and has became the de facto standard in knowledge representation. Most of the surveyed articles adopt DL to provide a formal definition of the model, e.g., [10, 18, 34, 88]. [19, 63] take a unique approach by initially defining a metamodel in UML and subsequently translating it into DL axioms. This method combines the visual clarity of UML with the formal precision of DL, providing a balance between intuitive representation and rigorous formalism. In contrast, in [53] the chosen formalism is Datalog with negation, a declarative logic programming language, originally introduced as a database query language.

Some approaches employ other formal languages to represent their system models when dynamic properties of the system must be captured. These include **Temporal Trace Language (TTL)**, a variant of an order-sorted predicate logic, as demonstrated in [61] and further discussed in [62]. This approach allows for the expression of temporal aspects and relationships within the model, which might be challenging to capture with standard graphical representations or DLs. A second example is the **Operational Conceptual Modelling Language (OCML)**[18] used in COBRA [58] in order to model how to actually compute metrics.

5.2.3 Categorization of KPIs. Various approaches have been proposed regarding categorization strategies for indicators, ranging from very simple categorizations to more advanced ones. Among

[18]http://technologies.kmi.open.ac.uk/ocml/

the former, while [78] groups indicators in categories based on custom classes, in [39] a simple categorization with three macro-categories for KPIs is provided, along with a more detailed one for intellectual capital indicators. A classical categorization is the one proposed in EM-KPI [46], grouping indicators according to their strategic, tactical, or operational objective. In the Smart Living Ontology [7], indicators are categorized in subclasses according to their scope (e.g., Pollution indicator is an Environmental indicator). A taxonomy of indicators is provided in [73] that are tailored to the specific case of **Service Level Agreement (SLA)** management for telecommunication, categorizing indicators in KPIs, **Key Quality Indicators (KQIs)**, and **Device Performance Indicators (DPIs)**. In PPINOT and related extensions [19, 20, 28, 29, 67], the base measures, which do not depend on others for their computation, are categorized in subclasses time, count, condition, and data measures.

In some cases, a categorization is derived from external KPI frameworks that are integrated in the proposed approach, as in NPIOnto [72], where the relation between a KPI and the *information source* inherently provides a classification, or in [70], where classes from the **Environmental Ontology (ENVO)** are reused. In [4], a taxonomy of KPIs is defined based on the UE's SDG framework,[19] while the Global City Indicator ontology and its extensions [33, 34, 84] rely on ISO 37120 for sustainable development of communities.

5.2.4 Representation of Dependencies. Relations of dependency are explicitly expressed in several models, to different aims, e.g., in [56] (*related_indicators* relation) or in [57] (*uses* relation). In [47], a tree of dependencies is represented to support exploration. In [3], a relation *takesDataFrom* between two indicators expresses that the former is mathematically dependent on the latter, while in [10] the measurement methods related to an indicator can accept other indicators as input (relation *has_input*). Both *atomic* and *composite* (i.e., aggregation of atomic ones) indicators are represented in [73], where a KPI has a Key Quality Indicator and interacts with other KPIs. In [2] the relation between qualitative and quantitative indicators is represented.

More expressive *dependency relations* are represented in [61], where different causality relationships between pairs of indicators are defined in terms of existence and strength of the dependency, through relations *causing*, *aggregation_of*, *correlated*. Positive and negative dependency is represented in KPIOWL [18], while a parent–child relation is defined in KPI-Q KPI Extensions [44]. A different case is [4], where dependency relations between indicators, goals, subgoals, and target are learned from data using ML/AI techniques.

Only a few articles provide an explicit representation for the calculation of an indicator or a metric. Among them, some approaches encode the mathematical expression as a string. Notable examples are [56], which does so for the function for a metric, and articles like [3, 47, 66, 85]. In the latter, metrics can be composed using arithmetic and logical operators. In PPINOT and related extensions [19, 20, 28, 29, 67], Process Performance Indicators can be defined through a *Derived measure* having a function as an attribute and explicit relations with other measures. The notion of *formula template* is introduced in [60] to categorize multiple formula expressions. The template is expressed as a string using placeholders. For each placeholder, a specific property associates the template to corresponding instances of indicators.

A simple form of *formula expression* is proposed in COBRA [58], where a special type of *Function Measure*, i.e., metrics that can be evaluated over a fixed number of inputs, is the "ratio" measure, which can specify dividend and divisor. Similarly, in the Global City Indicator Ontology [33] and related extensions [34, 84], an indicator is defined as a ratio and can be associated to its numerator and denominator, which decompose its definition.

[19]UN Sustainable Development Goals: https://sdgs.un.org/

In [91], the ontology explicitly represents KPI formulas in terms of input and output, parameters, and operators (unary, binary, and aggregation). Atomic KPIs are defined through queries on the extensional part of the ontology. In [22], a Mathematical Ontology includes the definition of the computation formulas for indicators, while other information is represented in a Business Ontology. Also in EIAW [88], complex indicators can be defined through a formula, which can contain basic mathematical operators and is expressed using a proprietary script language, while in [46] a formula can be represented through MathML, an XML language for definition of mathematical expressions. Similarly, in the earlier version of KPIOnto [23], a formula was represented as a MathML expression. In later versions (i.e., [5, 21, 24]), on the other hand, a KPI formula is represented through a set of classes and properties tailored to define a formula as the application of an operator to a set of arguments, which can be a constant, an indicator, or, recursively, another formula. This approach enables the encoding of the mathematical expression in a compositional way, by exploiting the same representation, namely OWL, for both the descriptive and the mathematical parts of the ontology.

In the OWL-Q KPI Extensions [44], a composite measure is provided with a formula that is explicitly linked to the component metrics, to support formula decomposition for KPI analysis.

5.3 Reasoning Capabilities

This subsection aims to discuss the use of reasoning by the analyzed approaches. In the following, we summarize and discuss results on tasks for which reasoning is used (D3.1), with a particular focus on the approaches that exploit reasoning for mathematical manipulation and dependency analysis (D3.2). See also Figure 5 for an overview.

5.3.1 Reasoning Tasks. Classic reasoning tasks, including subsumption and satisfiability, are often exploited to perform consistency check, as in [57]. Additionally, automatic classification may be provided, e.g., in [53], where a multidimensional ontology and its specification are expressed in terms of Datalog rules and constraints. A more advanced form of reasoning is proposed by [34], where Prolog rules are used to evaluate inconsistencies of various kinds, both intra- and inter-indicators: among them, "temporal inconsistencies" (between time intervals), "place inconsistencies" (when instances refer to different cities), and "measurement inconsistencies" (when they have inconsistent units of measurement). Also, "transversal inconsistencies" occur when two definitions of an indicator in different cities are not consistent with each other, while "longitudinal" inconsistencies may arise from changes in a city's boundaries over time. On a similar line, reasoning can be used to check the consistency of the multidimensional model, like in [63]. Here, an OWL-DL reasoner aims to check the completeness, consistency, correctness, and particularly the summarizability of measures along dimensions.

5.3.2 Reasoning on Dependencies. Among the articles providing an explicit representation of dependencies and/or mathematical formulas for Performance Indicators, reasoning is often used as a means to automatically support analysis [5, 18–24, 67, 73]. In [73], **Semantic Web Rule Language (SWRL)** rules are used together with the Jess rule engine to infer transitive relationships among PIs (e.g., the *influence* relation) or relationships between PIs and definition of SLAs. Mathematical reasoning is exploited through specific rules to calculate values of a composite PI from other PIs, based on the representation of the formula encoded in SWRL. The computed value can be used to derive violations on SLA thresholds. In [18], semantic rules in SWRL are defined on top of the KPIOWL ontology and are intended to support both the derivation of new relationships between indicators and the management of the ontology itself. One primary function of these rules is the detection of abnormal relationships, such as redundancy. An example is the identification of multiple indicators monitoring the same goal. Moreover, these semantic rules can be employed to identify inconsistencies based on PI values. For example, they can detect scenarios in which a goal, decomposed into subgoals, has children indicators that are satisfied, while the parent indicator is not. Stardog is used as a reasoner (a commercial version of Pellet OWL 2 [75]). In [20], inference rules in SWRL are used to detect positive or negative dependencies between indicators, and therefore to analyze potential conflicts and inconsistencies. Using a different approach, in [19, 67], automated analysis of PPIs is performed through DL reasoning in Hermit to infer new relationships between Business Processes and PPIs. By exploiting the mathematical representation of derived measures, reasoning is also used to find PPIs that are involved in the definition of a given one or to determine the resources potentially influencing a set of PPIs.

Prolog rules are used in [71] to verify if the knowledge graph contains the needed data to perform the calculation of a certain indicator. A set of Prolog rules are also defined for KPIOnto [5, 21, 23, 24] in order to support a variety of advanced tasks, including automated calculation of KPIs through mathematical manipulation via symbolic resolution of equations. Such services are capable of deriving all alternative formulas for a given indicator, evaluating the common set of dependencies among a set of indicators, and deriving what indicators can be computed starting from those already available. Other functionalities are defined in [24], with the purpose to support monitoring of indicators across multiple organizations, e.g., to assess the status of shared business processes. The set of KPIs produced by the organizations is compared in order to derive common dependencies and therefore identify mathematical expression to calculate common indicators. In [23], the reasoning framework is used for the management of a repository of indicators, supporting the creation of new indicators after checking that they are not equivalent to any other (to avoid duplication) and are consistent (to avoid contradictory definitions.

5.4 Objective of the Model

In this subsection, we report the objective of the selected models in order to identify the tasks they support (D4.1), their application fields (D4.2), and their level of specificity (D4.3).

5.4.1 Supported Tasks. Models of indicators serve various crucial functions within an organization. In the following, we discuss the most frequent tasks by grouping them into five categories, namely support to management, documentation, validation, support to analysis and monitoring, and advanced analysis. Results are summarized in Table 3.

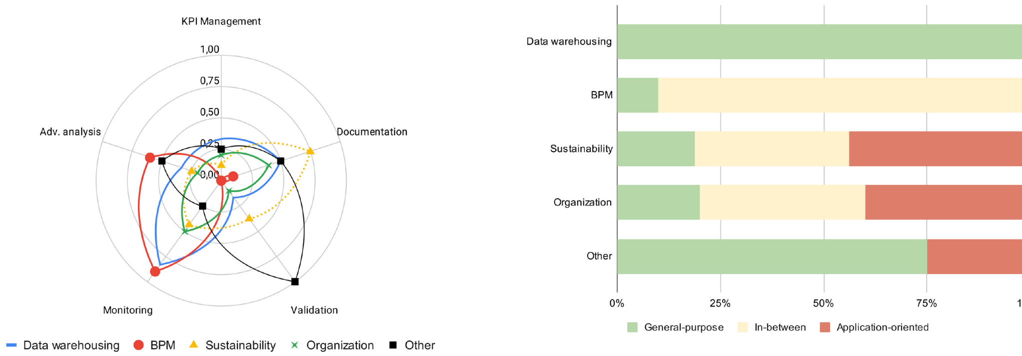
Some of the proposed approaches provide valuable support to *KPI management*. They assist in the selection and choice of appropriate KPIs ("Sel." in the table), aligning organizational objectives with measurable indicators. Moreover, they contribute to the process of building data marts, enabling the structured storage and retrieval of KPI-related data ("DM Building"). Several models of indicators also address *documentation* and repository needs. They help in creating comprehensive documentation for KPIs ("Doc."), ensuring clear understanding and knowledge sharing. They also facilitate disambiguation by linking documents and support exploration of KPI-related information ("Exp."), making it easier for stakeholders to navigate the KPI landscape. Some articles aim at supporting *validation*, by providing guidance in the definition of indicators ("Def.") through the exploitation of formal languages. This helps to ensure their correctness and reliability. Some validation approaches aim to check the properties of the metadata model underlying KPIs ("Met."). One of the most frequent tasks for a model of indicators is providing support for data *monitoring*, often for real-time evaluation of performance ("Mon."), or for facilitating analysis through Online Analytical Processing ("OLAP"), enabling the in-depth examination of KPI data. In some cases, when a computation formula for indicators is available or can be derived automatically, some frameworks can support their calculation ("Calc."), aiding in the precise quantification and evaluation of KPIs. Finally, more *advanced analysis* can be performed,

Table 3. Objective of the Model, in Terms of Tasks and Application Fields (D4.1 and D4.2)

Document	Field	KPI Management		Documentation		Validation		Monitoring			Adv. Analysis			
		Sel.	DM Building	Doc.	Exp.	Def.	Met.	OLAP	Mon.	Calc.	Root	Dep.	Goal	Comp.
[53]	Data Warehousing and OLAP							✓						
[63]							✓	✓						
[30]								✓						
[88]			✓	✓										
[22]			✓		✓					✓		✓		✓
[24]					✓			✓		✓		✓	✓	
[58]	Business Process Management									✓				
[85]									✓					
[44]									✓		✓			
[19, 20, 67]										✓		✓		
[29]									✓			✓	✓	
[28]									✓	✓				
[2]											✓	✓		
[79]				✓					✓					
[78]	Sustainability Management			✓										
[10]				✓		✓								
[38]				✓	✓									
[72]				✓										
[4]												✓		
[34]						✓								✓
[33]				✓										
[84]				✓		✓								
[70]				✓					✓					
[71]								✓						
[16]					✓				✓			✓		
[46]				✓					✓	✓				
[91]								✓	✓	✓				
[21]				✓		✓						✓		✓
[3, 7]		✓			✓				✓					
[52, 66]	Organization Management								✓	✓				
[61, 62]									✓				✓	
[18]				✓		✓								
[60]				✓										
[47]		✓												
[57]									✓					
[39]				✓										
[23]		✓		✓		✓								
[17, 56]	Other			✓					✓					
[73]										✓		✓		
[5]		✓				✓				✓		✓		

including analysis of dependencies among indicators ("Dep.") and evaluation of root-cause analysis ("Root"), helping organizations pinpoint the origins of issues and make informed decisions. Some approaches also support the evaluation of goal satisfaction ("Goal"), ensuring that KPIs are aligned with the organization's objectives. Additionally, they help to compare definitions of KPIs ("Comp."), fostering consistency and alignment across the organization's performance metrics.

5.4.2 *Application Field.* Papers have been grouped by the application field for which the model has been developed. The four most frequent groups, containing 42 papers, are reported in Table 3 and discussed below, along with details of their specific usage, while the remaining 4 papers are grouped under the category "Other." A comparative analysis is proposed in Figure 6(a), where the percentages of documents supporting different tasks are reported for each application field.



(a) (b)

Fig. 6. The radar chart (a) shows, for each application field, the percentage of documents related to the supported tasks; the stacked bar chart (b) represents the level of specificity of the approaches for different application fields.

Data Warehousing and OLAP. A number of approaches, which provide a model to represent dimensional hierarchies and measures, aim to support the management of multidimensional cubes and to perform OLAP operations on them, mostly supporting “monitoring” and “documentation” purposes, as shown in Figure 6(a). Here, the ontology serves as a conceptual layer between the business analysts and the multidimensional data, while multidimensional constraints (including the notion of aggregation) can then be enacted through rules. For instance, in [53] the MDO model supports business analysts in performing analytical processing in Datalog. As extensions of the RDF Data cube model, in [63] the OWL-DL ontology supports reasoning for analysis purposes, with also the capability to check properties of the multidimensional model (model completeness, consistency, correctness, and summarizability). Similarly, in [30] the ontology (which is an extension of QB4OLAP) is used to support OLAP operations, which are expressed through the SPARQL language, performed in a web scenario. In [22] an ontology of indicators is proposed to represent mathematical expressions for KPI calculation, with the purpose of supporting KPI elicitation, to analyze dependencies among them in terms of common components and check semantic correctness and redundancies of their definition. The approach aims to reduce the gap between a high-level managerial view of data and a technical view of DWH, hence simplifying and automating the main steps in data mart design. As an extension of such a work, the KPIOnto ontology is used to support self-service OLAP analysis, KPI browsing, and exploration in a Linked Data context with multidimensional sources [24]. More focused on the KPI management perspective, the EIAW framework [88] allows making business definitions explicit and supports the automatic deployment of data mart/warehouse-based BI applications.

Business Process Management. Applications of Semantic Business Process Management are aimed to minimize the support needed from IT staff throughout the business process lifecycle. To this goal, several articles focus on providing means for semi-automated monitoring and analysis of Business Processes, as also clear from Figure 6(a). Such approaches aim to reduce the risk of having multiple heterogeneous representations of information in the organization, which may lead to inconsistencies during information exchange, more complex maintenance, and erroneous analysis. In particular, in [58] **Business Process Analysis (BPA)** is done through a computation engine for metric calculation over domain-specific data.

Often, the model of indicators is complemented with an ontology for Business Processes, such that indicators are attached to specific activities in the process, like in [85], where the semantic model is then translated in an executable BPEL4SWS (BPEL for Semantic Web Services) process and the goal is detecting process events and performing the monitoring. Similarly, in [44], by formally specifying how KPIs can be measured over which **Business Process as a Service (BPaaS)** hierarchy components, a KPI analysis system is proposed in the context of Business Process monitoring, offering two main analysis capabilities: KPI assessment and drill-down. While the former is aimed at calculating KPI metric formulas on the fly, the latter can enable finding root causes of KPI violations.

In PPINOT [19, 20], the model is used to support monitoring of business processes and PPI value calculation along with the automation of many PPI management tasks, relying on the full traceability of the relations between PPIs and business process elements, e.g., for design-time estimation of those activities having an influence on indicators. In [67] the model is extended to support resource-aware PPI analysis, to identify people in the organization influencing a given PPI. As a further extension, in [29] the model is exploited for analysis of knowledge-based processes, supporting actors of a process during the process execution in complying with the established performance goals, through actionable guidelines. Finally, in [28], the model is extended to consider an object-centric context of process cases. In this context, a measure must specify the object type used as a reference for its calculation. On the other hand, in [2] the framework aims to perform queries and infer, through the a priori algorithm, relations between quantitative and qualitative indicators that are involved in a process.

In a different context, namely workflows executed on the Grid, in [79] performance metrics can be used to perform analysis through **RDF Data Query Language (RDQL)** queries and gain a better understanding of their results and facilitate the understanding of SLAs.

Sustainability Management. Several articles focused on representation of sustainability indicators in contexts related to public health, sustainable development, and smart cities, with the aim of enhancing data quality, comparability, and overall coherence among diverse databases with distinct objectives.

In [78] an ontology of public health indicators is proposed, serving as a tool to improve comparability and quality of data, ensuring coherence among different databases serving different purposes. The ISD-Economics ontology, introduced in [10], focuses on sustainable development indicators, with a specific emphasis on the economic dimension. It provides relevant elicited information to a software tool that is responsible for calculating sustainability levels for systems under analysis. One practical use of the ontology is in keeping the consistency of the unit of measurement. The extension to KPIOnto presented in [21] allows for the definition of ESG indicators, including their computation expressions, with the purpose to support sharing and comparison of indicators as well as the evaluation of a business's overall awareness and attention to social and environmental aspects. In the same domain, [38] addresses the representation of United Nations **Sustainable Development Goals (SDGs)** through an ontology. The SDG KOS platform supports users through their journey from documents to relevant statistics, aiding in contextualizing third-party information and facilitating data access. Furthermore, it supports multilingual semantic search and content linking by exploiting mappings between goals, targets, indicators, and data series to relevant terms in the **United Nations Bibliographic Information System (UNBIS)** and the EuroVoc vocabularies. The ontology proposed by [72] aims to represent indicators related to development agendas specific to Botswana, ensuring clear relationship mappings, data definitions, disaggregation, and integration with national, regional, and global development agendas. A taxonomy for SDG indicators, goals, subgoals, and targets is defined in [4] and is used to assist policy researchers in the design and evaluation of policies and in various types of analyses. The platform, tailored to a Policy Support System, manages data from diverse sources and features a policy "enunciator" for dynamic discovery and construction of semantic models.

In [33, 34] the **Global City Indicator (GCI)** is proposed to represent membership extent, temporal extent, spatial extent, and measurement of populations. It is used to represent city indicators as defined by ISO 37120 and supports their comparison to ascertain whether a city's interpretation is consistent with the standard. On the other hand, the GCI Finance Ontology [84], built on the GCI Ontology, focuses on ISO 37120 Finance Theme indicators. It aims to make the indicators, and the supporting data used to derive them, openly available on the Semantic Web.

The GCI Ontology has been extended by several articles in the domain of smart city. In [70], an IoT architecture enables data collection, storage, and processing from the city environment. It incorporates a semantic layer based on the **Environment Indicators Smart City Ontology (EISCO)**, granting interoperability to data representation in the platform. The ontology extends several other ontologies including **Semantic Sensor Network (SSN)** [15], **Environmental Ontology (ENVO)** [11], and GCI [33]. On the other hand, [71] proposes a city Knowledge Graph with an indicator ontology based on the GCI Ontology and ISO 37120 Indicator Definitions Ontology. The goal is to support indicator discovery and data visualization and the automated construction of dashboards according to discovered indicators.

In the same context of smart cities, the work by [16] introduces an integration layer to calculate a set of urban sustainability indicators from heterogeneous data sources. This allows users to explore and query the model and the associated data, aiding in understanding the effects of parameter value changes, uncovering and exploiting implicit dependencies, and selecting solutions optimizing performance metrics. In order to improve multi-level energy management, the EM-KPI ontology [46] intends to combine performance information and statistics for both the energy demand and supply sides in a district, to exchange KPIs among multiple stakeholders. In the work described in [91], an ontology-based approach is adopted for the automatic calculation of KPIs to support the evaluation of building energy efficiency. This method facilitates the comparison of building performance by analyzing extensive datasets at various levels of granularity. KPIs can be defined through inputs, formulas composed of a combination of parameters and operators, and outputs.

Finally, in [3, 7] a smart city ontology is defined containing the formal definition of indicators that aggregate and summarize urban data of interest, their formulas, and analysis dimensions. The ontology is exploited to support the generation of personalized exploration graphs for users of a smart city Data Lake. User profiles are defined by a set of constraints limiting the indicator instances that users are permitted to employ for data exploration.

Organization Management. In the domain of organization management, semantic models of indicators have been used to monitor performance, to assess strategic goals, or to support the definition of indicators within a reference repository.

The research conducted by [52, 66] focuses on the automated computation and visualization of KPIs from operational manufacturing systems, primarily for the purpose of operations management. This involves real-time performance assessment, with a specific focus on KPIs defined according to the ISO 22400 standard for automation systems and integration. The KPI implementation component is designed to receive event notifications from the manufacturing plant during runtime. These notifications are processed to extract valuable information and data pertaining to the production line. The extracted data is subsequently transmitted to a knowledge-based system for updating the knowledge base. KPI formulas are applied within this component whenever data is updated. The resulting KPI values are then conveyed to the user interface for visualization, typically presented in various graphical and descriptive formats. In a broader context, [61, 62] present an approach for modeling performance indicators within a comprehensive organizational modeling framework. It encompasses various perspectives, namely a process-oriented view, a performance-oriented view, an organization-oriented view, and an agent-oriented view. The framework incorporates KPIs along with goals, processes, and roles, with the ultimate objective of assessing organizational performance. It also defines mechanisms for assessing goal satisfaction by breaking down goals into subgoals. To ensure formal analysis and verification of PI specifications, a TTL checker in Prolog is employed.

KPIOWL, as described in [18], is designed in OWL2 to facilitate the identification and selection of KPIs. An ontology-driven approach is proposed to formally conceptualize essential elements of indicators, covering performance, results, measures, goals, and relationships of a given business strategy. By doing so, all the data involved in the process of selecting and analyzing KPIs are integrated and stored in shared repositories, enabling advanced querying and semantic validation through reasoning. Additionally, a set of SWRL rules is established to deduce relationships between indicators and goals. The article [23] introduces the SemPI framework that facilitates collaborative construction, management, and maintenance of a minimal and consistent repository of KPIs. By leveraging the KPIOnto schema and mathematical reasoning, the framework supports deducing identity and equivalence relationships among KPIs, thereby ensuring comprehensive consistency of the dictionary. With a similar aim, the ontology proposed by [60] offers the foundation for the development of a repository for Production Performance Indicators, where an indicator is described along with its properties. An indicator is associated to one of the predefined formula templates in order to make its calculation formula explicit.

In the context of [47], the objective is to support **Strategic Alliances (SAs)** among enterprises by monitoring and benchmarking them while facilitating the selection of appropriate indicators. This approach aids in gaining a deeper understanding of the factors driving SA performance, which in turn assists firms in making strategic and organizational decisions, such as determining whether to collaborate with other entities; structuring the alliance (e.g., number of nodes and control type); and defining monitoring strategies. In [39] the ontology serves to implement an information management tool, which guides the enterprise in its intellectual capital acquisition, measurement, monitoring, and exploitation. As such, it facilitates exploration of and navigation into enterprise knowledge, including KPIs for modeling intellectual capital indicators, which help in measuring the efficiency of the implementation of a strategy.

The ontology proposed by [57] addresses the need for organizations across various industrial sectors to maintain the high availability of their automation systems. It encompasses a condition monitoring ontology that incorporates ISO standards for condition monitoring and KPIs. The latter serve the purpose of aggregating a multitude of sensor values into easily interpretable figures, thereby offering more efficient insights into the condition of automation systems. This approach helps in understanding which sensors are used to calculate each KPI and how they are associated with state identification and fault diagnostics.

5.4.3 Specificity. The models proposed by the selected papers show a variable degree of specificity, ranging from models that include domain-specific classes and relations to general-purpose ones. In the following, we categorized them according to the extent to which they can be reused or adapted in different application scenarios. In particular, *general-purpose* models, which represent 32.6% of the models in the analyzed documents, typically only include very general concepts related to indicators and their properties, which are not tied to particular application domains. This makes such models reusable for various applications and domains but also asks for custom extensions in order to support specific functionalities.

On the other hand, *application-oriented* approaches model the indicator together with business terminology, making the ontology capable of supporting the application at hand, although less reusable. These include 26.1% of the documents. Among them, there are models defining specific classes or taxonomies dealing with public health [78], urban sustainability [7, 16, 33, 34], SDGs [4], finance [84], Service Level Agreements [73], strategic alliances [47], and specific industrial settings [52, 57, 66].

Finally, some models are *in between* the mentioned groups, as they show a medium degree of specificity, including generic classes and a few more specific ones. This group, which represents 41.3% of the total, includes models that refer to business processes, workflow management, events, or PPIs [19, 20, 28, 29, 60–62, 85], sometimes featuring a general-purpose module such as [79]. Other examples are models including a few classes from specific domains such as the health domain [2, 72], urban data [3], mappings to SDG targets [38], sustainability themes [10], or custom taxonomies in the context of IoT environmental data [70]. A custom categorization including intellectual capital indicators is proposed in [39], while in the business domain the model in [46] associates an indicator to an object building. Finally, the model proposed in [44] is an extension of OWL-Q, used for quality-of-service-based descriptions of web services.

Comparing the level of specificity with the application field, as shown in Figure 6(b), it is evident that approaches devoted to supporting data warehousing and OLAP favor general-purpose models, which can enable more flexible analyses across various domains. On the other hand, approaches in the context of "Sustainability" and "Organization Management" tend to focus on more application-oriented models, arguably because they are frequently designed to meet the requirements of projects with detailed research questions.

5.5 Technical Aspects

This subsection is devoted to discussing technical aspects of the models presented in the selected articles with respect to the format of the model (D5.1), reported implementation of the model (D5.2), availability of its specification (D5.3), and availability of a downloadable version of the model (D5.4).

We report a summary of the evaluation in Table 4, where we grouped documents that discuss the same work or extensions thereof and have equivalent technical aspects. In these cases, only the most advanced version was reported (e.g., we report that the model is downloadable if a link to an implementation is available in at least one of the documents in the group). According to the analysis, most of the models presented in the selected documents (73.9%) refer to OWL as the modeling language, with a few exceptions. Overall, 41 out of 46 approaches discuss in the paper a partial or a complete implementation (19.6% and 69.5% of the total, respectively). In 18 documents (39.1%) a link to a serialization of the model is available. In one case, however, the link appears to be broken. Finally, only nine documents include, for documentation purposes, the specification of the classes and properties defined in the ontology, typically as an external resource. However, for three documents (related to the same approach), the link to the specification is not operational.

6 Discussion

This section explores the impact of our studies on both research and practical applications and discusses general guidelines for selecting the most appropriate approach to modeling PIs based on a given scenario. Additionally, it summarizes limitations of the work and outlines open challenges and potential directions for future research.

6.1 Contributions

This survey provides a comprehensive overview of the research conducted in the past 20 years on semantic models of performance indicators. A set of 46 articles, selected through a systematic methodology (RQ1), were analyzed and compared in terms of representation approaches,

Table 4. Technical Aspects of the Models of the Selected Documents (D5)

Document	Format	Implementation	Specification	Downloadable
[17, 56]	OWL	✓	-	-
[78]	OWL	✓✓	-	-
[10]	OWL	✓✓	-	-
[39]	-	-	-	-
[85]	WSML	-	-	-
[73]	OWL	✓	-	-
[19, 20, 67]	OWL	✓✓	-	✓✓
[28]	-	-	-	-
[29]	OWL	✓✓	-	-
[16]	OWL	✓✓	-	-
[71]	OWL	✓✓	-	✓✓
[52, 66]	OWL	✓✓	-	-
[2]	OWL	✓	-	-
[72]	-	-	-	-
[61, 62]	TTL	✓✓	-	-
[3, 7]	OWL	✓✓	-	✓✓
[38]	OWL	✓✓	✓✓	✓✓
[70]	OWL	✓✓	-	-
[18]	OWL	✓✓	-	✓✓
[46]	OWL	✓✓	✓✓	✓✓
[60]	OWL	✓✓	-	-
[47]	-	-	-	-
[33, 34, 84]	OWL	✓✓	✓✓*	✓✓
[57]	OWL	✓✓	-	-
[91]	OWL	✓✓	-	-
[53]	Datalog	✓	-	✓✓
[63]	OWL	✓	-	-
[58]	OCML	✓✓	-	✓✓*
[88]	OWL	✓	-	-
[44]	OWL	✓✓	-	-
[30]	OWL	✓✓	✓	✓✓
[79]	OWL	✓	-	-
[22]	-	✓	-	-
[23]	OWL	✓✓	-	-
[5, 21, 24]	OWL	✓✓	✓✓	✓✓
[4]	OWL	✓✓	-	-

Documents are marked with a dash (-) if the feature is not provided or the information is not reported and a tick (✓) or a double tick (✓✓) if the feature is respectively partially or fully provided. The symbol (*) refer to broken links. Documents related to the same project with equivalent technical aspects are grouped.

expressivity, formalization level (RQ2), reasoning capabilities (RQ3), objectives and supported tasks for specific domain applications (RQ4), and technical aspects for sharing and reusability (RQ5). The study sheds light on the importance of clear and shared meanings of indicators and metrics in enabling effective decision-making, goal achievement, and data exchange across various application domains. Formal models, such as ontologies, have a strategic role in ensuring the consistency and trustworthiness of indicator data. By providing a structured framework for defining and representing indicators, semantic models facilitate data integration, sharing, and interpretation among stakeholders, ultimately leading to more informed decision-making processes.

6.2 General Guidelines

Besides being useful as a reference for scholars and practitioners in the field of performance measuring and management, the findings of this work can also provide guidance in selecting the most appropriate approach for modeling indicators. In particular, business analysts can exploit the results of the survey to navigate the complex landscape of semantic models, ultimately supporting them in selecting the most suitable model for effective and impactful representation of indicators in their respective domains.

Information useful to this aim is mostly condensed in Table 2, which provides insights on the expressivity of the models, and Table 3, summarizing tasks and application fields of the selected models. As the survey points out, each analyzed work focuses on a specific modeling approach and provides deeper details only for some of the analysis dimensions. As such, no single model outperforms the others for all the considered aspects. Hence, practitioners should prioritize dimensions based on their relative importance to specific scenarios. We summarize in the following some general guidelines, stemming from the results, that can help in identifying a set of relevant models starting from the requirements of a specific target scenario. We first consider the field and task of the model as a starting point, due to the fact that reuse is the easiest and most cost-effective solution if a suitable model for the scenario at hand is available or can be derived with little effort from an existing model:

— *Domain orientation.* In most practical cases, KPI modeling is driven by a specific domain. These types of models (see Section 5.4) typically include classes and properties that are closely mapped to domain concepts, e.g., the relation between a KPI and a Business Process in the BPM field. As such, they make the alignment between the model and the information system easier, reducing the burden of further customization of the model. In this sense, Table 3 can help in finding the approaches that are particularly oriented at specific application fields. Conversely, if the domain is significantly far from those considered in the present survey or if the scenario is not strictly domain based, general-purpose models should be preferred. Their peculiar features include the capability to generalize and thus an enhanced flexibility in modeling a wider set of indicators.
— *Task orientation.* In case the model, independently on the application domain, must support a specific task, both vertical models and general-purpose ones may be useful. The former may directly provide support for a specific set of tasks but may not be as general and flexible as the latter. Table 3 can support this search by considering a categorization based on columns and sub-columns in order to identify those models that are more aligned with the tasks at hand.
— *Concept representation.* Once a set of relevant models has possibly been identified, a fine-grained selection can be done based on their expressivity, namely what concepts are represented by the model and how. Even in case of novel models built from scratch, specific parts of existing models can be reused to represent certain concepts, e.g., dimensions, goal, process. Section 5.2 provides useful information to perform the selection, summarized in Table 2. Articles with three or two ticks in the table should be prioritized due to a more structured modeling of the respective concept.

A single model often cannot fully meet all scenario requirements. In these situations, starting from scratch should still be avoided whenever possible, as it demands more effort and can lead to incompatibility with existing models. Instead, reusing and integrating parts of existing models can enhance interoperability. Very often, referring to general-purpose ontologies in these cases is a preferable choice, since they can be more easily extended and customized than very detailed solutions. This would, however, require an accurate analysis and a moderate effort. To reduce the cost, preference should be given to already available and actionable models, such as those already implemented and downloadable (see Table 4). Further considerations can be taken into account to address specific aspects such as the complexity of the model and openness:

— *Model complexity.* In some cases, indicators are provided with several metadata, including dependencies or a computation formula. These situations ask for both complex representation capabilities to model such relations and tools for automated analysis and computation. Therefore, models that incorporate a formal representation and use a logic reasoner supporting advanced functionalities should be preferred. On the other hand, if the model is rather simple, reasoning techniques and complex representations may introduce unnecessary overhead. In such cases, keeping the representation minimal is advisable.
— *Openness and standards.* When openness is a priority in the design and application of KPI models, practitioners should prioritize approaches that align with the principles of Open and FAIR data (Findable, Accessible, Interoperable, Reusable) [86] and adhere to open standards such as RDF and OWL (see Table 4). These ensure that definitions of indicators are not only transparent but also universally interpretable, facilitating interoperability across systems and stakeholders. This aspect is particularly critical for public organizations and institutions, which are often required to openly document their processes to ensure accountability and transparency.

6.3 Limitations

While this manuscript aims to comprehensively survey semantic models for performance indicators, certain limitations must be acknowledged to contextualize its findings and scope. Despite employing comprehensive search strategies and a systematic approach, the extensive body of literature in the field means some relevant works may have been unintentionally omitted. Additionally, the subject of the study is inherently complex and multi-dimensional. While we have attempted to cover as many aspects as possible, as discussed in Section 4, some features of analysis possibly were left out of the study, such as the relationship between indicator models and real-world data, particularly regarding the ability to represent actual data points. These limitations highlight opportunities for further research and refinement. As the process is fully documented, future studies can build on this work and expand its findings.

6.4 Open Challenges

Despite the significant contributions of semantic models to PI management, several challenges remain open.

— *Standardization and interoperability.* Ensuring consistency in terminology and adopting shared semantic representations are crucial for interoperability and data exchange between systems and stakeholders, as well as to interlink different ontologies across different domains. To this aim, abstract, general-purpose KPI models should be preferred to domain-specific ones, as they provide greater flexibility and reduce the effort required for mapping between different ontologies. On the other hand, general-purpose models act as an overarching framework enabling specialized models to map to them and simplifying system alignment. This modular approach reduces integration complexity and reduces the risk of representation mismatches, streamlining the integration process, avoiding the creation of silos, and allowing for easier adaptation and evolution of domain-specific models. An alternative solution involves integrating multiple ontologies to expand concept coverage. This would, however, need to reconcile possibly alternative epistemological approaches in representing concepts, e.g., treating analysis dimensions as relations in an ontology and as classes in another, for which a direct integration is not possible.

— *Scalability and adaptability.* The scalability and adaptability of semantic models to evolving performance measurement needs pose significant challenges. As organizations and industries evolve, the ability of semantic models to accommodate new indicators, changing requirements, and diverse application scenarios becomes crucial for their continued relevance and effectiveness. General-purpose models are intrinsically more adaptable to future needs, since they do not overspecify classes and relations. As new systems, domains, and technologies emerge, such ontologies provide a scalable foundation, allowing the introduction of new concepts without requiring significant rework of the existing structure. On the other hand, domain-specific models tend to be rigid, making them harder to scale and adapt as requirements evolve, although they can immediately be used for specific tasks and therefore reduce the burden needed for their customization for the problem at hand.

Although not directly related to modeling, further aspects are very relevant in the practical usage of KPI definitions.

— *Alignment to FAIR principles.* The two above-mentioned challenges relate to compliance with the FAIR principles. Research should focus on definitions of KPIs that enhance usability and integration across systems. Well-documented, discoverable, and interoperable definitions improve data management and decision-making, while comprehensive, standardized metadata schemes describing their purpose, scope, and context facilitate their discovery. Advanced search tools can improve KPI findability by implementing efficient indexing and search algorithms based on attributes like category, measurement type, or business objectives. Reusability, on the other hand, can be enhanced by developing modular and extensible KPI definitions adaptable to different contexts and applications. Research should focus on flexible models that allow customization and extension without significant rework. Establishing best practices and standardized templates will further support reuse by providing a consistent format for defining and applying KPIs across projects and organizations.

— *Management tools.* There is a need for further research into tools that support the creation and management of KPI definitions. Developing comprehensive libraries or catalogs with specific functionalities can facilitate the efficient management of KPI definitions. Such tools should assist in the organization, standardization, and maintenance of KPI data, making it easier for users to manage and utilize KPIs effectively. The process of selecting appropriate KPIs can benefit from advanced support systems, including those leveraging reasoning-based techniques or Machine Learning methods. Research into automated and intelligent systems that aid in the selection and optimization of KPIs could enhance decision-making and ensure that KPIs align with organizational goals and performance objectives.

References

[1] Joud Al Dakheel, Claudio Del Pero, Niccolò Aste, and Fabrizio Leonforte. 2020. Smart buildings features and key performance indicators: A review. *Sustainable Cities and Society* 61 (2020), 102328.

[2] Emna Ammar El Hadj Amor and Sonia Ayachi Ghannouchi. 2017. Toward an ontology-based model of key performance indicators for business process improvement. In *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA'17)*. IEEE, 148–153.

[3] Ada Bagozi, Devis Bianchini, Valeria De Antonellis, Massimiliano Garda, and Michele Melchiori. 2019. Personalised exploration graphs on semantic data lakes. In *Proceedings of On the Move to Meaningful Internet Systems: OTM 2019 Conferences: Confederated International Conferences: CoopIS, ODBASE, C&TC 2019*. Springer, 22–39.

[4] Pooja Bassin, Niharika Sri Parasa, Srinath Srinivasa, and Sridhar Mandyam. 2022. Big data management for policy support in sustainable development. In *Big-Data-Analytics in Astronomy, Science, and Engineering*, Shelly Sachdeva, Yutaka Watanobe, and Subhash Bhalla (Eds.). Springer International Publishing, Cham, 3–15.
[5] Filippo Benvenuti, Claudia Diamantini, Domenico Potena, and Emanuele Storti. 2017. An ontology-based framework to support performance monitoring in public transport systems. *Transportation Research Part C: Emerging Technologies* 81 (2017), 188–208.
[6] Tim Berners-Lee, James Hendler, and Ora Lassila. 2001. The semantic web. *Scientific American* 284, 5 (2001), 34–43.
[7] Devis Bianchini, Valeria De Antonellis, Massimiliano Garda, and Michele Melchiori. 2018. Semantics-enabled personalised urban data exploration. In *International Conference on Web Information Systems Engineering*. Springer, 361–376.
[8] Pearl Brereton, Barbara A. Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. 2007. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software* 80, 4 (2007), 571–583.
[9] Dan Brickley and Ramanathan Guha. 2014. *RDF Schema 1.1*. W3C Recommendation. W3C. https://www.w3.org/TR/2014/REC-rdf-schema-20140225/
[10] V. Brilhante, A. Ferreira, J. Marinho, and J. S. Pereira. 2006. Information integration through ontology and metadata for sustainability analysis. In *Proceedings of the iEMSs 3rd Biennial Meeting, "Summit on Environmental Modelling and Software."*
[11] Pier Luigi Buttigieg, Norman Morrison, Barry Smith, Christopher J. Mungall, Suzanna E. Lewis, and Envo Consortium. 2013. The environment ontology: Contextualising biological and biomedical entities. *Journal of Biomedical Semantics* 4 (2013), 1–9.
[12] Luisa F. Cabeza, Esther Galindo, Cristina Prieto, Camila Barreneche, and A. Inés Fernández. 2015. Key performance indicators in thermal energy storage: Survey and assessment. *Renewable Energy* 83 (2015), 820–827.
[13] Simone Caruso, Manfredi Bruccoleri, Astrid Pietrosi, and Antonio Scaccianoce. 2023. Artificial intelligence to counteract "KPI overload" in business process monitoring: The case of anti-corruption in public organizations. *Business Process Management Journal* 29, 4 (2023), 1227–1248.
[14] Surajit Chaudhuri and Umeshwar Dayal. 1997. An overview of data warehousing and OLAP technology. *SIGMOD Record* 26, 1 (1997), 65–74.
[15] Michael Compton, Payam Barnaghi, Luis Bermudez, Raul Garcia-Castro, Oscar Corcho, Simon Cox, John Graybeal, Manfred Hauswirth, Cory Henson, Arthur Herzog, et al. 2012. The SSN ontology of the W3C semantic sensor network incubator group. *Journal of Web Semantics* 17 (2012), 25–32.
[16] Berta Cormenzana, Maria-Cristina Marinescu, Monica Marrero, Sergio Mendoza, Salvador Rueda, and Rosario Uceda-Sosa. 2015. Models for sustainability. In *2015 IEEE 1st International Smart Cities Conference (ISC2'15)*. IEEE, 1–6.
[17] Maria de los Angeles Martin and Luis Olsina. 2003. Towards an ontology for software metrics and indicators as the foundation for a cataloging web system. In *Proceedings of the IEEE/LEOS 3rd International Conference on Numerical Simulation of Semiconductor Optoelectronic Devices (IEEE Cat. No. 03EX726)*. IEEE, 103–113.
[18] María del Mar Roldán-García, José García-Nieto, Alejandro Maté, Juan Trujillo, and José F Aldana-Montes. 2021. Ontology-driven approach for KPI meta-modelling, selection and reasoning. *International Journal of Information Management* 58 (2021), 102018.
[19] Adela del Río-Ortega, Manuel Resinas, Cristina Cabanillas, and Antonio Ruiz-Cortés. 2013. On the definition and design-time analysis of process performance indicators. *Information Systems* 38, 4 (2013), 470–490.
[20] Adela del Río-Ortega, Manuel Resinas, and Antonio Ruiz-Cortés. 2010. Defining process performance indicators: An ontological approach. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems."* Springer, 555–572.
[21] Claudia Diamantini, Tarique Khan, Domenico Potena, and Emanuele Storti. 2022. Shared metrics of sustainability: A knowledge graph approach. In *Ceur Workshop Proceedings*, Vol. 3194. 244–255.
[22] Claudia Diamantini and Domenico Potena. 2011. Thinking structurally helps business intelligence design. In *Information Technology and Innovation Trends in Organizations: ItAIS: The Italian Association for Information Systems.* Springer, 109–116.
[23] Claudia Diamantini, Domenico Potena, and Emanuele Storti. 2016. SemPI: A semantic framework for the collaborative construction and maintenance of a shared dictionary of performance indicators. *Future Generation Computer Systems* 54 (2016), 352–365. https://doi.org/10.1016/j.future.2015.04.011
[24] Claudia Diamantini, Domenico Potena, and Emanuele Storti. 2021. Analytics for citizens: A linked open data model for statistical data exploration. *Concurrency and Computation: Practice and Experience* 33, 8 (2021), e4186.
[25] Eladio Domínguez, Beatriz Pérez, Ángel L. Rubio, and María A. Zapata. 2019. A taxonomy for key performance indicators management. *Computer Standards & Interfaces* 64 (2019), 24–40.
[26] Simme Douwe P. Flapper, Leonard Fortuin, and Paul P. M. Stoop. 1996. Towards consistent performance management systems. *International Journal of Operations & Production Management* 16, 7 (1996), 27–37.

[27] Alexander M. Elizarov, A. V. Kirillovich, Evgeny K. Lipachev, O. A. Nevzorova, Valery D. Solovyev, and N. G. Zhiltsov. 2014. Mathematical knowledge representation: Semantic models and formalisms. *Lobachevskii Journal of Mathematics* 35, 4 (2014), 348–354.
[28] Bedilia Estrada-Torres, Adela del Río-Ortega, and Manuel Resinas. 2022. Defining process performance measures in an object-centric context. In *International Conference on Business Process Management*. Springer, 210–222.
[29] Bedilia Estrada-Torres, Pedro Henrique Piccoli Richetti, Adela Del-Río-Ortega, Fernanda Araujo Baiao, Manuel Resinas, Flávia Maria Santoro, and Antonio Ruiz-Cortés. 2019. Measuring performance in knowledge-intensive processes. *ACM Transactions on Internet Technology (TOIT)* 19, 1 (2019), 1–26.
[30] Lorena Etcheverry, Alejandro Vaisman, and Esteban Zimányi. 2014. Modeling and querying data warehouses on the semantic web using QB4OLAP. In *Data Warehousing and Knowledge Discovery*, Ladjel Bellatreche and Mukesh K. Mohania (Eds.). Springer International Publishing, Cham, 45–56.
[31] Simme Douwe P. Flapper, Leonard Fortuin, and Paul P. M. Stoop. 1996. Towards consistent performance management systems. *International Journal of Operations & Production Management* 16, 7 (1996), 27–37.
[32] Farnaz Fotrousi, Samuel A. Fricker, Markus Fiedler, and Franck Le-Gall. 2014. KPIs for software ecosystems: A systematic mapping study. In *Proceedings of the 5th International Conference on Software Business: Towards Continuous Value Delivery (ICSOB'14)*. Springer, 194–211.
[33] Mark S. Fox. 2015. The role of ontologies in publishing and analyzing city indicators. *Computers, Environment and Urban Systems* 54 (2015), 266–279.
[34] Mark S. Fox. 2018. The semantics of populations: A city indicator perspective. *Journal of Web Semantics* 48 (2018), 48–65. https://doi.org/10.1016/j.websem.2018.01.001
[35] Jennifer Horkoff, Daniele Barone, Lei Jiang, Eric Yu, Daniel Amyot, Alex Borgida, and John Mylopoulos. 2014. Strategic business modeling: Representation and reasoning. *Software & Systems Modeling* 13 (2014), 1015–1041.
[36] Ivo Hristov and Antonio Chirico. 2019. The role of sustainability key performance indicators (KPIs) in implementing sustainable strategies. *Sustainability* 11, 20 (2019), 5742.
[37] Mirjana Ivanović and Zoran Budimac. 2014. An overview of ontologies and data resources in medical domains. *Expert Systems with Applications* 41, 11 (2014), 5158–5166.
[38] Amit Joshi, Luis Gonzalez Morales, Szymon Klarman, Armando Stellato, Aaron Helton, Sean Lovell, and Artur Haczek. 2021. A knowledge organization system for the United Nations Sustainable Development Goals. In *European Semantic Web Conference*. Springer, 548–564.
[39] Yelena Jussupova-Mariethoz and Andre-Rene Probst. 2007. Business concepts ontology for an enterprise performance and competences monitoring. *Computers in Industry* 58, 2 (2007), 118–129.
[40] Robert S. Kaplan and David P. Norton. 2005. The balanced scorecard: Measures that drive performance. *Harvard Business Review* 83, 7 (2005), 172.
[41] Megan Katsumi and Mark Fox. 2018. Ontologies for transportation research: A survey. *Transportation Research Part C: Emerging Technologies* 89 (2018), 53–82.
[42] Alexander Kipp, Tao Jiang, Maria Grazia Fugini, and Ioan Salomie. 2012. Layered green performance indicators. *Future Generation Computer Systems* 28 (2012), 478–489.
[43] Barbara Kitchenham and Stuart Charters. 2007. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. Technical Report. EBSE.
[44] Kyriakos Kritikos, Dimitris Plexousakis, and Robert Woitch. 2017. A flexible semantic KPI measurement system. In *International Conference on Cloud Computing and Services Science*. Springer, 237–261.
[45] Khadija Letrache, Omar El Beggar, and Mohamed Ramdani. 2016. Modeling and creating KPIs in MDA approach. In *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt'16)*. IEEE, 222–227.
[46] Yehong Li, Raúl García-Castro, Nandana Mihindukulasooriya, James O'Donnell, and Sergio Vega-Sánchez. 2019. Enhancing energy management at district and building levels via an EM-KPI ontology. *Automation in Construction* 99 (2019), 152–167.
[47] Barbara Livieri, Mario A. Bochicchio, and Antonella Longo. 2014. Ontologies and information visualization for strategic alliances monitoring and benchmarking. In *ICEIS*. 402–409.
[48] Barbara Livieri, Pierluca Di Cagno, and Mario Bochicchio. 2015. A bibliometric analysis and review on performance modeling literature. *Complex Systems Informatics and Modeling Quarterly* 2 (2015), 56–71.
[49] Florian Matthes, Ivan Monahov, Alexander W. Schneider, and Christopher Schulz. 2012. Towards a unified and configurable structure for EA management KPIs. In *Working Conference on Practice-driven Research on Enterprise Transformation*. Springer, 284–299.
[50] Jose-Norberto Mazón, Jens Lechtenbörger, and Juan Trujillo. 2009. A survey on summarizability issues in multidimensional modeling. *Data & Knowledge Engineering* 68, 12 (2009), 1452–1469.
[51] Dimitris Mourtzis, Sophia Fotia, and Michael Doukas. 2015. Performance indicators for the evaluation of product-service systems design: A review. In *Proceedings of the International Conference on Advances in Production*

Management Systems: Innovative Production Management Towards Sustainable Growth: IFIP WG 5.7 International Conference (APMS'15), Proceedings, Part II 0. Springer, 592–601.

[52] Usman Muhammad, Borja Ramis Ferrer, Wael M. Mohammed, and Jose L. Martinez Lastra. 2018. An approach for implementing key performance indicators of a discrete manufacturing simulator based on the ISO 22400 standard. In *2018 IEEE Industrial Cyber-physical Systems (ICPS'18)*. IEEE, 629–636.

[53] Bernd Neumayr, Stefan Anderlik, and Michael Schrefl. 2012. Towards ontology-based OLAP: Datalog-based reasoning over multidimensional ontologies. In *Proceedings of the 15th International Workshop on Data Warehousing and OLAP (DOLAP'12)*. Association for Computing Machinery, New York, NY, USA, 41–48. https://doi.org/10.1145/2390045.2390053

[54] Robert C. Nickerson, Upkar Varshney, and Jan Muntermann. 2013. A method for taxonomy development and its application in information systems. *European Journal of Information Systems* 22, 3 (2013), 336–359.

[55] Ian Niles and Adam Pease. 2001. Towards a standard upper ontology. In *Proceedings of the International Conference on Formal Ontology in Information Systems, Volume 2001*. 2–9.

[56] Luis Olsina, Fernanda Papa, and Hernán Molina. 2008. Ontological support for a measurement and evaluation framework. *International Journal of Intelligent Systems* 23, 12 Special Issue (2008), 1282–1300. https://doi.org/10.1002/int.20320

[57] Faruk Pasic, Benedict Wohlers, and Matthias Becker. 2019. Towards a KPI-based ontology for condition monitoring of automation systems. In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'19)*. IEEE, 1282–1285.

[58] Carlos Pedrinaci and John Domingue. 2009. Ontology-based metrics computation for business process analysis. In *Proceedings of the 4th International Workshop on Semantic Business Process Management (SBPM'09)*. Association for Computing Machinery, New York, NY, USA, 43–50. https://doi.org/10.1145/1944968.1944976

[59] Kai Petersen, Sairam Vakkalanka, and Ludwik Kuzniarz. 2015. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology* 64 (2015), 1–18.

[60] G. Pintzos, M. Matsas, and G. Chryssolouris. 2012. Defining manufacturing performance indicators using semantic ontology representation. *Procedia CIRP* 3 (2012), 8–13.

[61] Viara Popova and Alexei Sharpanskykh. 2010. Modeling organizational performance indicators. *Information Systems* 35, 4 (2010), 505–527. https://doi.org/10.1016/j.is.2009.12.001

[62] Viara Popova and Alexei Sharpanskykh. 2011. Formal modelling of organisational goals based on performance indicators. *Data & Knowledge Engineering* 70, 4 (2011), 335–364.

[63] Nicolas Prat, Imen Megdiche, and Jacky Akoka. 2012. Multidimensional models meet the semantic web: Defining and reasoning on OWL-DL ontologies for OLAP. In *Proceedings of the 15th International Workshop on Data Warehousing and OLAP (DOLAP'12)*. Association for Computing Machinery, New York, NY, USA, 17–24. https://doi.org/10.1145/2390045.2390049

[64] Hafizur Rahman and Md Iftekhar Hussain. 2020. A comprehensive survey on semantic interoperability for Internet of Things: State-of-the-art and research challenges. *Transactions on Emerging Telecommunications Technologies* 31, 12 (2020), e3902.

[65] Yves Raimond and Guus Schreiber. 2014. *RDF 1.1 Primer*. W3C Note. W3C. https://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/

[66] Borja Ramis Ferrer, Usman Muhammad, Wael M. Mohammed, and José L. Martínez Lastra. 2018. Implementing and visualizing ISO 22400 key performance indicators for monitoring discrete manufacturing systems. *Machines* 6, 3 (2018), 39.

[67] Adela del Río Ortega, Manuel Resinas Arias de Reyna, Cristina Cabanillas Macías, and Antonio Ruiz Cortés. 2013. Defining and analysing resource-aware process performance indicators. In *CAISE'13 Forum at the 25th International Conference on Advanced Information Systems Engineering (CAISE'13)*.

[68] Luis Fernando Castro Rojas and Carlos Mario Zapata Jaramillo. 2013. Executable pre-conceptual schemas for representing key performance indicators. In *2013 8th Computing Colombian Conference (8CCC'13)*. IEEE, 1–6.

[69] Ann Rosenberg, Greg Chase, Rukhshaan Omar, James Taylor, and Mark von Rosing. 2011. *Applying Real-world BPM in an SAP Environment*. Galileo Press, Bonn, Boston.

[70] Jesús Noel Suárez Rubí and Paulo Roberto de Lira Gondim. 2021. IoT-based platform for environment data sharing in smart cities. *International Journal of Communication Systems* 34, 2 (2021), e4515.

[71] Henrique Santos, Victor Dantas, Vasco Furtado, Paulo Pinheiro, and Deborah L. McGuinness. 2017. From data to city indicators: A knowledge graph for supporting automatic generation of dashboards. In *European Semantic Web Conference*. Springer, 94–108.

[72] Oarabile Sebubi, Irina Zlotnikova, and Hlomani Hlomani. 2019. A lightweight version of national performance indicator ontology (NPIonto). In *2019 Conference on Next Generation Computing Applications (NextComp'19)*. IEEE, 1–6.

[73] Sin-seok Seo, Arum Kwon, Joon-Myung Kang, and James Won-Ki Hong. 2011. OSLAM: Towards ontology-based SLA management for IPTV services. In *12th IFIP/IEEE International Symposium on Integrated Network Management (IM'11) and Workshops*. IEEE, 1228–1234.
[74] Craig Shepherd and Hannes Günter. 2006. Measuring supply chain performance: Current research and future directions. *International Journal of Productivity and Performance Management* 55, 3/4 (2006), 242–258.
[75] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. 2007. Pellet: A practical OWL-DL reasoner. *Journal of Web Semantics* 5, 2 (2007), 51–53. https://doi.org/10.1016/j.websem.2007.03.004
[76] Kristian Stancin, Patrizia Poscic, and Danijela Jaksic. 2020. Ontologies in education—State of the art. *Education and Information Technologies* 25, 6 (2020), 5301–5320.
[77] Mark Staples and Mahmood Niazi. 2007. Experiences using systematic review guidelines. *Journal of Systems and Software* 80, 9 (2007), 1425–1437.
[78] György Surján, Éva Szilágyi, and Tamás Kováts. 2006. A pilot ontological model of public health indicators. *Computers in Biology and Medicine* 36, 7–8 (2006), 802–816.
[79] Hong-Linh Truong, Schahram Dustdar, and Thomas Fahringer. 2007. Performance metrics and ontologies for Grid workflows. *Future Generation Computer Systems* 23, 6 (2007), 760–772. https://doi.org/10.1016/j.future.2007.01.003
[80] Montijn van de Ven, Paola Lara Machado, Alexia Athanasopoulou, Banu Aysolmaz, and Oktay Turetken. 2023. Key performance indicators for business models: A systematic review and catalog. *Information Systems and e-Business Management* 21, 3 (2023), 753–794.
[81] Amy Van Looy and Aygun Shafagatova. 2016. Business process performance measurement: A structured literature review of indicators, measures and metrics. *SpringerPlus* 5, 1 (2016), 1–24.
[82] W3C. 2009. *OWL 2 Web Ontology Language Document Overview*. W3C Recommendation. W3C. https://www.w3.org/TR/2009/REC-owl2-overview-20091027/
[83] W3C. 2014. *The RDF Data Cube Vocabulary*. W3C Recommendation. W3C. https://www.w3.org/TR/vocab-data-cube/
[84] Zhongxiu Wang and Mark S. Fox. 2016. Expenditure, revenue and taxes—Towards a standard for representing city finance open data. In *2016 IEEE International Smart Cities Conference (ISC2'16)*. 1–6. https://doi.org/10.1109/ISC2.2016.7580847
[85] Branimir Wetzstein, Zhilei Ma, and Frank Leymann. 2008. Towards measuring key performance indicators of semantic business processes. In *International Conference on Business Information Systems*. Springer, 227–238.
[86] Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, et al. 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3, 1 (2016), 1–9.
[87] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE'14)*. Association for Computing Machinery, New York, NY, USA, Article 38, 10 pages.
[88] Guotong Xie, Yang Yang, Shengping Liu, Zhaoming Qiu, Yue Pan, and Xiongzhi Zhou. 2007. EIAW: Towards a business-friendly data warehouse using semantic web technologies. In *The Semantic Web*, Karl Aberer, Key-Sun Choi, Natasha Noy, Dean Allemang, Kyung-Il Lee, Lyndon Nixon, Jennifer Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, et al. (Eds.). Springer, Berlin, 857–870.
[89] Neetu Yadav and Mahim Sagar. 2013. Performance measurement and management frameworks: Research trends of the last two decades. *Business Process Management Journal* 19, 6 (2013), 947–971.
[90] Carlos Mario Zapata Jaramillo and Luis Fernando Castro Rojas. 2016. A method based on patterns for deriving key performance indicators from organizational objectives. *Polibits* 53 (2016), 55–64.
[91] Yun-Yi Zhang, Zhen-Zhong Hu, Jia-Rui Lin, and Jian-Ping Zhang. 2021. Linking data model and formula to automate KPI calculation for building performance benchmarking. *Energy Reports* 7 (2021), 1326–1337. https://doi.org/10.1016/j.egyr.2021.02.044

Received 20 February 2024; revised 1 February 2025; accepted 14 February 2025