Architecture of Ontology-Driven Multidimensional Analytical Systems Based on Formal Ontologies

Oleksandr Stryzhak[1], Andrii Yaremenko[2] and Viacheslav Gorborukov[1,3]
[1]*National Center "Junior Academy of Sciences of Ukraine", Dekhtyarivska Str. 38-44, 04119 Kyiv, Ukraine*
[2]*National University "Yuri Kondratyuk Poltava Polytechnic", Pershotravneva Avenue 24, 36011 Poltava, Ukraine*
[3]*Department of Multimedia Systems, National University of Kyiv-Mohyla Academy, Skovorody Str. 2, 04655 Kyiv, Ukraine*
sae953@gmail.com, y_aa@ukr.net, slavon07@gmail.com

Keywords: Formal Ontology, Multidimensional Analytics, Ontology Engineering, Ontology-Driven Analytical Systems, Semantic Normalization, Multicriteria Analysis, Inductive Construction of Analytical Structures, Intelligent Analytics.

Abstract: The article presents a formally grounded approach to the development of ontology-driven systems for multidimensional analytics. In this approach, a formal domain ontology is treated as the primary and invariant knowledge model that provides a sufficient basis for the inductive construction of analytical structures. Unlike traditional solutions, where ontologies are typically applied only as auxiliary semantic or integration layers, the proposed framework derives analytical dimensions, hierarchies, indicators, and permissible aggregation operations directly from the logical organization of the ontological model. The study formalizes several key processes required for building such systems. These include the structuring and semantic alignment of heterogeneous data sources, the construction of a formal domain ontology, and the subsequent use of this ontology to support multicriteria evaluation and the generation of multidimensional analytical representations. Within this framework, a mathematical definition of an ontology-analytic mapping operator is introduced. This operator ensures a rigorous transformation from the ontological model to a multidimensional analytical structure while preserving type information, semantic constraints, and interpretive properties. The results demonstrate that analytical facts can be induced from ontological elements, which enables the correct handling of context-dependent and partially defined dimensions. In addition, the article proposes a conceptual architecture for an ontology-driven analytical system organized into functional and conceptual layers. This layered architecture supports methodological consistency, reproducibility of analytical procedures, and scalability of the proposed approach. The presented formalization is general in nature and can be applied to the development of intelligent analytical systems in various domains. In particular, it is well suited for systems designed to analyze and evaluate educational and intellectual achievements, where consistency of semantics, criteria, and analytical procedures is critically important. Overall, the obtained results provide a formal foundation for constructing coherent, semantically consistent, and explainable analytical systems across diverse application areas.

1 INTRODUCTION

1.1 Motivation

The development of information and analytical systems is accompanied by a continuous increase in the complexity of domain environments, the diversity of data sources, and the requirements for semantic consistency, reliability, and interpretability of analytical results. In contemporary applications, data are generated from heterogeneous sources, including documents, registries, tabular datasets, and textual materials. These data often differ in their level of structure and semantic clarity and are used in contexts characterized by multiple layers of interpretation and strong contextual dependence. Under such conditions, analytical systems are increasingly viewed not merely as tools for data processing or aggregation but as instruments for constructing a formally consistent analytical space that enables comparison, generalization, and interpretation of results across different application contexts.

This situation creates a demand for approaches that combine rigorous formal modeling with the ability to support semantic interpretation, traceability, and structural extensibility. Ontology-based approaches play a particularly important role in addressing these challenges, as they allow the formal specification of domain structures, including key concepts, relationships, and interpretative constraints. The use of ontologies facilitates the transition from fragmented and heterogeneous datasets to coherent analytical models capable of supporting semantic consistency, reproducibility of model construction, and interpretability of analytical outcomes.

1.2 State of the Art

Modern information and analytical systems are traditionally built on data warehousing, OLAP, and multidimensional data model technologies, within which analytical facts are described through a combination of dimensions, hierarchies, measures, and aggregation operations [1], [2]. In classical OLAP systems, the main focus is on cube design, finding relevant analytical views, query optimization, and user interaction with multidimensional structures. Djiroun et al. address the problem of discovering and constructing OLAP cubes based on the analysis of user queries, demonstrating the importance of adaptive formation of analytical views [1].

At the same time, such approaches remain predominantly focused on existing collections of multidimensional cubes or on the construction of additional cubes in response to user queries, whereas the semantic nature of the subject domain typically does not serve as a formal basis for deriving the analytical structure itself.

A separate area of research is related to the transformation of traditional data warehouses under the influence of Big Data, NoSQL databases, and unstructured or semi-structured data sources. Martinez-Mosquera et al. systematize approaches to integrating OLAP with NoSQL databases and demonstrate that the transition to Big Data and NoSQL environments requires adapting OLAP cube construction models, warehouse types, modeling methods, processing modes, and performance criteria [3]. Ptiček et al. emphasize that for the integration of NoSQL data and traditional data warehouses, a mere technical combination of sources is insufficient, as the problem must be solved at a higher level of abstraction, specifically at the level of semantic modeling and design [4]. This points to the need for models capable not only of integrating heterogeneous data but also of ensuring its semantic interpretation in the analytical process.

The fundamental principles of this level of formalization are established in the works of Gruber, Guarino, Uschold, and Gruninger, as well as in subsequent generalizations by Staab and Studer [5]-[8]. Gruber defines an ontology as an explicit specification of conceptualization, oriented toward the sharing and reuse of formalized knowledge [5]. Guarino emphasizes the importance of formal ontology for separating domain knowledge from the specific procedures for its use [6]. Uschold and Gruninger view ontologies as a means of achieving shared understanding, interoperability, reusability, and reliability of software systems [7]. Within the scope of this study, an ontology is treated not as a reference dictionary or an external metadata system, but as a formal model of a subject domain capable of defining valid interpretations of objects, relations, attributes, values, and semantic constraints [8].

Semantic technologies and ontologies are actively used in semantic data warehouses, semantic ETL, and exploratory OLAP. Abelló et al. demonstrate that Semantic Web technologies can support the discovery, integration, semantic annotation, and analytical processing of external sources in exploratory OLAP scenarios [2]. Deb Nath et al. propose high-level RDF-based ETL constructs for defining, mapping, transforming, integrating, and loading semantic multidimensional data [9]. Antunes et al. demonstrate the application of Semantic Web techniques to link strategic management, the Balanced Scorecard ontology, and the BI environment [10]. In practice-oriented BI literature, the semantic layer is also viewed as a business-oriented abstraction between the data warehouse and user analytical applications, which improves data accessibility, terminology consistency, governance, and visualization [11]. However, in a significant portion of these approaches, ontologies or the semantic layer are used primarily for integration, annotation, contextualization, traceability, semantic ETL, or to support predefined or semi-generated analytical structures. At the same time, the rules for deriving a complete multidimensional analytical model from a unified ontological basis, including dimensions, hierarchies, measures, admissible aggregation operations, and analytical facts, remain only partially formalized.

The work closest to the subject matter of this study is that of Romero and Abelló, dedicated to the automation of multidimensional design based on ontologies [12]. It proposes an approach to identifying multidimensional concepts from a domain ontology and using these concepts for designing data warehouses. This work is particularly important for positioning the present study because it directly addresses the transition from an ontological description of a domain to multidimensional design.

At the same time, the present article is positioned within the line of ontology-based multidimensional design, but shifts the focus from the detection of multidimensional schema elements to the formal role of the ontology in analytical model construction. In Romero and Abelló's approach, the ontology serves as the input for identifying multidimensional design elements - facts, dimensions, measures, bases, and dimension hierarchies - and for constructing conceptual data warehouse schemas. In the present article, the ontology is treated as a formal domain model whose analytical function is defined after semantic normalization and interpretive specification of domain indicators. It is then used not only to identify multidimensional schema elements, but to derive a multidimensional analytical model under explicit semantic, functional, interpretive, and aggregation constraints.

Thus, the proposed ontology-analytical mapping operator formalizes a controlled transition from a formal domain ontology to a logical analytical cube, in which dimensions, hierarchies, measures, admissible aggregation operations, and analytical facts are systematically derived.

In the context of intelligent, educational, and explainable systems, ontologies are also viewed as a means of bridging the gap between formal knowledge, practical application, and the explainability of results. Mizoguchi and Bourdeau demonstrate that ontology engineering provides a shared vocabulary, supports consistent domain modeling, and facilitates the construction of theory-compliant systems in educational environments [13]. Futia and Vetrò, as well as Confalonieri and Guizzardi, link knowledge graphs and ontologies to the problem of explainable and neuro-symbolic AI, where explicit knowledge is used to enhance the comprehensibility, justifiability, and interpretability of results [14], [15]. For analytical systems, this means that an ontological model can perform not only an integrative but also an explanatory function, specifying the semantic origin of indicators, criteria, aggregations, and analytical conclusions.

Ontology-driven knowledge structuring and decision support systems constitute a separate line of application. Nadutenko et al. consider ontology-driven lexicographic systems as one example of applying the ontological approach to the organization, structuring, and interpretation of complex knowledge resources [16]. Stryzhak et al. propose a decision-making system based on the ontology of the choice problem, in which the task of rational choice is formalized through an ontological representation of alternatives, criteria, and ranking indicators [17]. Taken together, these works demonstrate that ontological models can support both the organization of complex information resources and decision-oriented analytical procedures. At the same time, this article shifts the focus from a specific choice problem or lexicographic structuring to the broader problem of inductively forming a multidimensional analytical model from a formal domain ontology.

Thus, an analysis of existing approaches allows us to identify three main limitations of the current state of research. First, in classical OLAP and DW approaches, analytical schemas are predominantly designed in advance or reconstructed according to the user's query, but their connection to the formal semantics of the subject domain is usually not defined as a separate, mathematically specified transition from ontology to analytical model [1], [2]. Second, in semantic OLAP, semantic DW, and semantic ETL, ontologies are mostly used for integration, annotation, contextualization, traceability, or ETL support, while their role as a generative basis for constructing multidimensional structures remains limited [2], [4], [9]-[11]. Third, even in approaches directly aimed at automating multidimensional design from ontologies [12], the main emphasis is placed on identifying multidimensional schema elements and supporting the construction of conceptual DW structures, whereas the present article focuses on defining a unified ontology-analytical mapping operator that specifies how dimensions, hierarchies, measures, admissible aggregation operators, and analytical facts are derived from the formal ontology.

Accordingly, the proposed approach builds upon existing work in ontology engineering, semantic OLAP, semantic data warehouses, semantic ETL, ontology-based multidimensional design, and ontology-based decision support, but shifts the focus from using ontology as an auxiliary semantic layer or a source of individual multidimensional design patterns to using it as a formal generative basis for multidimensional analytics. Within this approach, a multidimensional analytical structure is not simply projected from a predefined schema or user query, but is derived under explicitly defined mapping rules from concepts, relations, functions, types, and semantic constraints represented in the domain ontology.

1.3 Objectives and Tasks

Despite the widespread adoption of ontological models in modern information and analytical systems, the problem of formally aligning domain semantic models with analytical structures remains unresolved. This issue is particularly evident in relation to multidimensional representations, indicators, hierarchies, and aggregation operations.

In many existing implementations, analytical structures are still developed separately from the ontological model, which weakens the formal connection between domain semantics and analytical operations. This separation complicates the reproducibility of analytical models, limits their transferability across different domains, and weakens the ability to ensure the semantic correctness of analytical operations.

Consequently, there is a scientific need to develop a formal approach for ontology-driven construction of multidimensional analytical representations. Such an approach should support the inductive derivation of analytical dimensions, indicators, hierarchies, and permissible aggregation operations directly from the domain ontology while preserving its semantic constraints and interpretative properties. Addressing this problem is essential for the development of coherent, scalable, and methodologically consistent analytical systems capable of operating with heterogeneous data across a wide range of applied tasks.

The aim of this paper is to develop and demonstrate a formal ontology-driven framework in which a domain ontology serves as a generative formal basis for deriving multidimensional analytical structures, including dimensions, hierarchies, analytical measures, domain-specific indicators, admissible aggregation operations, and analytical facts.

To achieve the stated objective, the following tasks must be addressed:

- To formalize the ontological model of the domain as a structured system of entities, relationships, and semantic constraints.
- To establish principles for the inductive derivation of analytical dimensions, hierarchies, and analytical measures, including domain-specific indicators, from the components of the ontology.
- To develop a formal operator that maps the ontological structure to a multidimensional analytical model.
- To illustrate how the proposed formal constructs are instantiated in the domain of intellectual competition analytics.
- To substantiate the semantic consistency, traceability, and reproducibility conditions of the proposed approach.

The article is organized as follows. Section 2 presents the formal ontology-analytical model, including the formalization of input data, the construction of the domain ontology, the definition of the multidimensional analytical model, and the specification of the ontology-to-analytics mapping operator. Section 3 describes the conceptual multilayer architecture of the ontology-driven analytical system and its functional components. Section 4 demonstrates the application of the proposed approach to the problem of analyzing and ranking intellectual achievements within the context of competitive events. The final section summarizes the conclusions and outlines directions for future research.

2 MATERIALS AND METHODS

2.1 Formal Ontology-Analytical Model

The formal model introduced in this section describes the sequence of transformations from heterogeneous input sources to a normalized domain structure, then to a formal ontology, and subsequently to an enriched ontology that can be used for multidimensional analytical construction. The definitions below specify the main structural components of the proposed approach and provide the basis for the ontology-analytical mapping operator introduced in the next subsection.

The input data of the domain may originate from heterogeneous documentary sources, including various electronic forms, questionnaires, reports, registries, and textual descriptions. To ensure a unified approach to the further processing of such sources, a set of input documents is introduced:

$$S = \{s_1, s_2, \dots, s_n\}, \tag{1}$$

where S - denotes the set of all documents obtained directly from digital sources, s_i - represents an individual document within the domain.

Because input documents may differ significantly in their internal organization, level of structure, and degree of semantic completeness, a structural transformation operator is introduced. This operator maps the set of input documents to a unified normalized data structure N (which is suitable for further formal analysis and semantic interpretation):

$$\Phi_{\text{norm}}: S \to N. \quad (2)$$

Where:

- S - the set of input documents;
- Φ_{norm} - the operator of structural transformation and alignment;
- N- a unified normalized data structure.

The result of this transformation stage is described by the structure:

$$N = \left(X', A', V_A', F', C', \tau_X'\right). \quad (3)$$

Where:

- X'-the set of domain objects;
- A'- a set of normalized and semantically consistent attributes;
- V_A'- a set of normalized attribute values;
- $F' \subseteq X' \times A' \times V_A'$- a ternary relation of the «object-attribute-value type»;
- C'- a set of object types (classes);
- $\tau_X': X' \to C'$- an object typing function.

Based on the normalized structure, a formal ontology of the domain is constructed that provides a semantic interpretation of objects, attributes, their values, and the relationships between them. The construction of the ontology is described by an ontological mapping operator:

$$\Phi_{\text{ont}}: N \to O. \quad (4)$$

Where:

- N- is the normalized data structure;
- O- is the formal ontology of the domain;
- Φ_{ont}- is the ontological mapping operator that performs the construction of the ontological model.

After the ontology has been constructed, multicriteria evaluation becomes possible, followed by the extension of the ontological model with rating values.

$$\Phi_{MR}: O \to MR,\ \Omega_{MR}: O \times MR \to O^*. \quad (5)$$

Where:

- O- is the formal ontology of the domain;
- MR- is the formal model of the multicriteria evaluation and ranking problem;
- Φ_{MR}- is the operator for constructing the ranking model based on the ontology;
- Ω_{MR}- is the operator for applying the model MR to the ontology O with the formation and integration of evaluation results;
- O^*- is the extended ontology with integrated ranking results.

Thus, the first part of the model defines the transformation chain:

$$S \xrightarrow{\Phi_{olap}} N \xrightarrow{\Phi_{ont}} O \xrightarrow{\Omega_{MR}} O^*.$$

This chain is important for the architecture proposed in Section 3: the system does not transform raw data directly into an analytical cube. Instead, heterogeneous sources are first normalized, then interpreted ontologically, then enriched with evaluation results, and only after that used as the basis for ontology-driven multidimensional analytical construction.

The ontology-driven construction of a multidimensional analytical representation is based on the principle that a sufficiently specified formal ontological model provides an internally consistent semantic basis for analytical model construction. Such a description supports the derivation of analytical structures under explicitly defined mapping rules and reduces the need for ad hoc non-semantic assumptions. Within this framework, the ontology functions as a formal source of analytical interpretation. Formally, the ontology is defined as an ordered triple:

$$O = (X_o, R_o, F_o), \quad (6)$$

where: $X_o = X_C \cup X_I$- denotes the set of ontological elements, comprising both abstract concepts X_C, and concrete domain objects X_I, R_o- denotes the set of binary relations between ontology elements, and F_o - denotes the set of ontological functions that define the interpretation of attributes and values of ontology elements.

Each element $x \in X_o$ is typed according to the ontological typing function:

$$\tau_X: X_o \to C, \quad (7)$$

where τ_X- denotes the ontological typing function, and C- denotes the universal set of ontological types and categories.

Typing makes it possible to distinguish concepts of classificatory, contextual, and factual nature, as well as individual objects considered as instances of the corresponding concepts. Ontological relations R_o are defined as a subset of the Cartesian product of the set of ontological elements:

$$R_o \subseteq X_o \times X_o, \quad (8)$$

where $\left(x_i, x_j\right) \in R_o$ indicates the existence of a semantically interpreted relation between elements x_i

and x_j. Such relations may represent generalization, inclusion, composition, instance-class membership, or other structural dependencies that form the semantic framework of the ontology.

The set of ontological functions is defined as:

$$F_o = \{f \mid f: X_o \to V\}, \quad (9)$$

where f - an ontological function that interprets an attribute or property of an ontology element, V - the set of admissible values, which may include numerical, symbolic, or ordered domains.

The presence of functions with numerical or ordered value domains creates formal prerequisites for the subsequent analytical use of the ontology, in particular for the construction of indicators and aggregation operations within a multidimensional representation.

Ontological-analytical transformation consists in the formal mapping of the ontological model into a multidimensional analytical structure while preserving the semantic constraints, typing, and interpretations specified in the ontology. Unlike the procedural construction of analytical schemas, this transformation is considered a mathematically defined operator acting on formal structures.

For further formalization of ontology-driven analytical mapping, it is necessary to provide a rigorous formal description of the multidimensional analytical model. Such a representation is treated as an abstract mathematical structure independent of any specific implementation of data storage or computation. The multidimensional analytical model is defined as an ordered quintuple:

$$M_{OLAP} = (D, H, M, A, \mu). \quad (10)$$

Where:

- D - is a finite set of analytical dimensions;
- H - is the set of hierarchies defined over the dimensions, M - is the set of analytical measures;
- A - is the set of admissible aggregation operators;
- μ - is a multidimensional analytical mapping function.

Each dimension $D_i \in D$ is interpreted as an ordered set of discrete values or levels of generalization corresponding to a particular aspect of the subject domain. Formally, the set of admissible values for each analytical dimension is defined by the following mapping:

$$\mathrm{dom}: D \to 2^{V_D}, \quad (11)$$

for each $D_i \in D$ let $dom(D_i) \subseteq V_D$ - denote the set of admissible values of dimension D_i, while V_D - represents the universal set of all possible values of analytical dimensions, and 2^{V_D} - denotes the power set ofV_D. In this context $D_i$denotes the dimension itself (i.e., the axis), where as $d_i \in dom(D_i)$ - represents a specific value of that dimension (i.e., the coordinate of a fact).

The hierarchies H define admissible transitions between levels of granularity within a single dimension or across related dimensions. The set of measures M consists of numerical or ordered functions whose values can be aggregated along the dimensions. For each measure, a subset of admissible aggregation operators is defined to ensure the correctness of generalization and interpretation.

The multidimensional mapping functionμ formalizes the relationship between dimension values and measures, thereby determining the structure of facts in the multidimensional model:

$$\mu: dom(D_1) \times \cdots \times dom(D_k) \times M \to V', \quad (12)$$

Where:

- $D_i \in D$ - denotes an individual analytical dimension, $dom(D_i)$ - is the set of admissible values of the i -th analytical dimension,
- k - is the number of dimensions involved in the analytical mapping,
- M - is the set of measures, V' is the set of admissible measure values (numerical or ordered domains),
- μ - is a multidimensional function that maps a combination of dimension values and a selected measure to its corresponding value.

As a result, each element of the multidimensional model can be represented as an ordered tuple:

$$f_{OLAP} = (d_1, d_2, \dots, d_k, m, v), \quad (13)$$

where $d_i \in dom(D_i)$ -is a specific value of the corresponding dimension, $m \in M$ - is the measure associated with the given combination of dimensions, and $v \in V'$ - denotes the value of that measure.

Thus, the multidimensional analytical model provides the target structure for ontology-driven construction. The next subsection specifies how its components - dimensions, hierarchies, measures, admissible aggregation operations, and analytical facts - are derived from the formal ontology rather than externally imposed as a predefined OLAP schema.

2.2 Formal Operator of Ontology-Analytical Transformation

A key element of ontology-driven construction of multidimensional analytical representations is the ontology-analytical mapping operator, which provides a formal transition from an ontological model to the structure of multidimensional analytics. Unlike procedural or heuristic approaches to the construction of OLAP schemas, this operator is defined as a strict mapping between formal structures that preserves the semantic properties of the source ontology.

Formally, the ontology-analytical mapping operator is defined as:

$$\Phi_{olap}: O \to M_{OLAP}. \qquad (14)$$

Where:
- Φ_{olap} - denotes the ontology-analytical mapping operator;
- O - is the formal ontological model;
- M_{OLAP} - is the multidimensional analytical model.

The operator Φ_{olap} interprets ontological elements according to their semantic roles, which are determined by their typing and structural relations. In particular, concepts classified as categorical or contextual types may be mapped to the set of analytical dimensions; hierarchical relations between concepts may be transformed into dimension generalization structures; and ontological functions with numerical value domains may be mapped to the set of analytical measures.

For each component of the ontology, the operator Φ_{olap}defines a set of partial mappings:

$$\Phi_{olap} = (\Phi_D, \Phi_H, \Phi_M, \Phi_A, \Phi_\mu). \qquad (15)$$

Where:
- $\Phi_D: X_o \to D$- denotes the mapping of concepts to dimensions;
- $\Phi_H: R_o \to H$- represents the mapping of ontological relations to dimension hierarchies;
- $\Phi_M: F_o \to M$- denotes the mapping of ontological functions to analytical measures;
- $\Phi_A: M \to 2^A$ - represents the mapping of measures to the set of admissible aggregation operator;
- Φ_μ - denotes the mapping of ontology facts to the multidimensional function μ.

Thus, the ontology-analytical mapping operator formalizes the principle according to which the multidimensional analytical model is considered a derived structure of the ontology, rather than an independently designed data schema. The formation of analytical dimensions constitutes the first stage in the specification of the multidimensional analytical model. At this stage, ontological concepts are interpreted as axes of the multidimensional space. Within the ontology-driven approach, dimensions are not predefined; instead, they are inductively derived from the ontology based on the typing and semantic roles of concepts.

The set of analytical dimensions is defined as a subset of the set of ontological concepts:

$$D \subseteq X_C, \qquad (16)$$

where D - denotes the set of analytical dimensions and X_C - denotes the set of ontological concepts.

The criterion for including a concept $x \in X_C$ in the set of dimensions is its membership in types that are admissible for analytical decomposition. Formally, this condition is defined as follows:

$$x \in D \quad \Leftrightarrow \quad (x \in X_C \ \wedge \ \tau_X(x) \in C_{dim}), \qquad (17)$$

Where:
- $\tau_X(x)$ - denotes the type of the ontological concept x;
- $C_{dim} \subseteq C$- is the subset of ontological types permitted as analytical dimensions;
- C - is the universal set of ontological types.

The set C_{dim} typically includes concepts of a classification or contextual nature that determine the way facts are grouped, segmented, or compared. Concepts of a factual or interpretative nature are not included in the set of dimensions because they do not define independent axes of the analytical space.

In general, the set of analytical dimensions is formed as follows:

$$D = \{x \in X_C \mid \tau_X(x) \in C_{dim}\}. \qquad (18)$$

Where:
- D - is the set of analytical dimensions;
- X_C - is the set of ontological concepts;
- τ_X- is the typing function;
- C_{dim}- is the set of types admissible for analytical dimensions.

Thus, the formation of the set of analytical dimensions directly depends on the ontological typing and semantic structure of the domain, which ensures the correctness and semantic consistency of the analytical axes.

The hierarchical organization of dimensions is a necessary condition for supporting aggregation and drill-down operations in multidimensional analytics. Within the ontology-driven approach, dimension hierarchies are not constructed independently of the domain but are instead derived from ontological relations that already capture semantic dependencies between concepts. However, not every ontological relation can be interpreted as hierarchical in the analytical sense. Such a relation must define a direction of hierarchy from more detailed levels to more generalized ones and satisfy the properties of transitivity and acyclicity, which ensure the correct formation of hierarchical structures.

Formally, the set of dimension hierarchies can be represented as:

$$H = \{(x_i, x_j) \in R_o \mid x_i \prec x_j \wedge x_i, x_j \in D\}. \quad (19)$$

Where:

- H - is the set of dimension hierarchies,
- R_o - is the set of ontological relations,
- D - is the set of analytical dimensions,
- $\prec$ - denotes the binary generalization relation corresponding to the roll-up operation in multidimensional analytics.

Thus, the hierarchies of dimensions in the multidimensional analytical model constitute a direct mapping of the semantic hierarchies specified in the ontology. This ensures that analytical aggregation operations remain consistent with the conceptual structure of the domain and prevents the emergence of artificial or semantically inconsistent hierarchies.

Measures in the multidimensional analytical model represent quantitative or ordered characteristics of the domain whose values are subject to aggregation and comparative analysis. Within the ontology-driven approach, the set of measures is defined as the class of ontological functions that permit analytical operations:

$$M = \{f \in F_o \mid \forall x \in X_o : f(x) \in V'\}, \quad (20)$$

where M - is the set of analytical measures, F_o - is the set of ontological functions, X_o - is the set of ontological elements (concepts and/or objects) for which functions from F_o are defined, V' - is the set of admissible numerical or ordered value domains (i.e., domains that allow aggregation and/or comparison), and f - is an ontological function interpreted as a measure when the condition $f: X_o \to V'$.

The value of a measure $m \in M$ for an ontological element $x \in X_o$ is determined by applying the corresponding function:

$$m(x) = v,\ v \in V', \quad (21)$$

Where:

- m - is an analytical measure (a function from the set M);
- x- is the ontological element for which the measure value is computed;
- v - is the numerical or ordered value of the measure;
- V' - is the domain of admissible measure values.

Thus, measures are derived directly from ontological interpretations of values, ensuring semantic unambiguity of the measures and the correctness of their subsequent aggregation within analytical views.

Aggregation is a fundamental mechanism of multidimensional analysis; however, the admissibility of aggregation operations is determined not only by the data type but primarily by the semantics of the measures. In the ontology-driven approach, aggregation operations are aligned with measures at the level of their ontological interpretation. This prevents semantically incorrect generalizations and ensures the meaningful interpretability of analytical results.

The set of basic aggregation operators is defined as:

$$A = \{sum, avg, count, min, max\}, \quad (22)$$

Where:

- A - is the set of aggregation operators;
- sum- denotes the summation operator;
- avg - denotes the averaging operator;
- $count$ - represents the counting operator;
- min and max- denote the minimum and maximum operators, respectively.

The admissibility of aggregation is a property not of the operator itself but of a particular measure. Therefore, for each measure $m \in M$, the set of admissible aggregation operators is defined directly as a semantically determined subset of the set A:

$$A_m = \{a \in A \mid (m, a) \in Compat\}, \quad (23)$$

Where:

- A_m - is the set of aggregation operators admissible for the measure m;
- m - is an analytical measure;
- A - is the base set of aggregation operators;
- $Compat \subseteq M \times A$ - denotes the relation of semantic admissibility of the aggregation operator a for the measure m.

Thus, the correspondence between measures and admissible aggregation operations can be formalized by the mapping:

$$\Phi_A: M \to 2^A, \quad (24)$$

where Φ_A - is the function that semantically aligns measures with aggregation operators, and 2^A - denotes the power set of the set of aggregation operators.

The definition of the set A_m is based on the ontological properties of the measure. Additive measures allow summation and averaging operations; intensive measures allow averaging and extreme-value operations; whereas ordinal or nominal measures are typically restricted to minimum and maximum operations. Such semantic alignment ensures the correctness of aggregation and prevents the application of operations that contradict the semantic meaning of the measures.

The ontology-driven construction of a multidimensional analytical representation is based on identifying ontological elements that serve as sources of analytical facts. This role is not an inherent property of a concept and is not determined solely by its type; rather, it is induced by its ontological interpretation, particularly by the presence of ontological functions whose values permit analytical aggregation or comparison.

Fact-generating ontological elements are defined by the following condition:

$$X_F = \{x \in X_o \mid \exists f \in F_o \colon f(x) \in V'\}, \tag{25}$$

Where:

- X_F- is the subset of ontological elements that serve as sources of analytical facts;
- X_o- is the set of ontological elements of the domain;
- F_o - is the set of ontological functions;
- $f\colon X_o \to V$- is an ontological interpretation function;
- $V' \subseteq V$ - is the domain of values admissible for analytical interpretation (numerical or ordered).

Thus, fact generation is a role induced by ontological interpretation, rather than a consequence of membership in a predefined class. An ontological fact is represented as an ordered triple:

$$f_O = \big(x, f, f(x)\big), \tag{26}$$

Where:

- $x \in X_F$- is a fact-generating ontological element;
- $f \in F_o$ - is an ontological function;
- $f(x) \in V'$ - is an analytically significant value.

An ontological fact is semantically defined; however, it does not yet possess coordinates within the multidimensional analytical space. The coordinates of an analytical fact are formed through the induction of relevant analytical dimensions from the ontological context of the fact-generating element.

Let D- denote the set of analytical dimensions previously defined as a subset of the set of ontological concepts. For each $x \in X_F$ the set of dimensions that are semantically relevant for constructing the corresponding analytical fact is defined as:

$$D_{axes}(x) = \left\{ y \in D \;\middle|\; \begin{matrix} (x,y) \in R_o \vee \\ (y,x) \in R_o \end{matrix} \right\}. \tag{27}$$

Where:

- $D_{axes}(x) \subseteq D$ - is the set of analytical dimensions (ontological axes induced for the fact-generating element$x$);
- $x \in X_F$ - is the fact-generating ontological element;
- X_o- is the set of ontological elements of the domain;
- $R_o \subseteq X_o \times X_o$- is the set of ontological relations;
- $y \in D$ - is an ontological element considered as a candidate dimension in the context of the fact induced by the element x.

The set D defines the complete space of admissible analytical dimensions of the domain, whereas $D_{axes}(x)$ specifies only those dimensions that are semantically relevant in the context of a particular fact. This makes it possible to formally describe situations in which not all dimensions participate in the formation of each analytical fact, without introducing artificial or semantically empty values.

Because a multidimensional analytical model operates not with dimensions themselves but with their values, a mapping is introduced for each fact-forming element that induces the values of the corresponding dimensions within the context of a given fact:

$$\delta\colon \{(x,y) \mid x \in X_F,\;\; y \in D_{axes}(x)\} \to V_D. \tag{28}$$

Where:

- X_F - denotes the set of fact-forming ontological elements;
- V_D - is the set of admissible dimension values;
- $\delta(x,y)$ - is the value of dimension $y \in D_{axes}(x)$, induced in the context of the fact associated with element x.

The function δ therefore provides a formal transition from ontological concepts to the coordinates of the analytical space. Based on this, the ontologically induced multidimensional mapping function is defined as a specialization of μ, restricted to the set of ontological facts:

$$\begin{gathered} \mu_F\colon (x,f) \mapsto (d_1, d_2, \dots, d_k, m, v), \\ d_i = \delta(x, y_i) \in V_D, \quad i = 1, \dots, k. \end{gathered} \tag{29}$$

Where:

- μ_F - denotes the ontologically induced multidimensional mapping function;
- $x \in X_F$ - is a fact-forming ontological element;
- $f \in F_O$ - represents the ontological function;
- $D_{axes}(x) \subseteq D$ - the set of dimensions relevant to the fact x, $\langle y_1, \dots, y_k \rangle = ord(D_{axes}(x))$;
- $k = | D_{axes}(x) |$ - an ordered list of relevant dimensions and their cardinality;
- δ - a mapping defined by formula (28), that induces the values of the dimensions for $y \in D_{axes}(x)$;
- $m = \Phi_M(f) \in M$ - an indicator corresponding to the ontological function,;
- $v = f(x) \in V'$ - the value of the indicator.

The resulting tuple $(d_1, \dots, d_k, m, v)$ constitutes an analytical fact of the multidimensional model and corresponds to a single row of the fact table. The coordinates of the fact are determined by the values of semantically relevant dimensions, whereas the measure is given by the value of the ontologically interpreted indicator.

The set of analytical facts is formed as the image of the set of ontological facts under the mapping μ_F:

$$F_{OLAP} = \left\{ \mu_F(x, f) \left| \begin{matrix} x \in X_F, f \in F_O, \\ f(x) \in V' \end{matrix} \right. \right\}, \tag{30}$$

where F_{OLAP} denotes the set of analytical facts of the multidimensional model, μ_F - is the ontologically induced fact mapping, X_F - denotes the set of fact-forming ontological elements, F_O is the set of ontological functions, and V' - is the set of admissible indicator values.

Thus, the multidimensional analytical structure emerges as a formally induced derivative of the domain ontology, in which the fact table is constructed directly from ontological elements, relations, and functions. The logical model of analytical facts is based on the semantic relevance of dimensions and does not require filling missing coordinates with zero or null values. This property ensures correct aggregation, consistency of interpretation, and independence from the specific characteristics of the physical data storage implementation.

Consequently, the induced analytical facts are semantically traceable to their ontological origin. Each fact is determined by a fact-generating element, the relevant dimensional context, and an ontologically interpreted measure. The practical instantiation of these constructs is presented in Section 4.

3 ONTOLOGY-DRIVEN ANALYTICAL SYSTEM ARCHITECTURE

3.1 Conceptual Layered Structure

The proposed architecture of the ontology-driven analytical system is designed as a layered model that supports the transformation of heterogeneous domain data into semantically grounded multidimensional analytical representations. Its purpose is to organize the interaction between data structuring, ontology construction, multicriteria enrichment, analytical transformation, logical cube construction, and interpretation of results within a unified formal framework. The general layered architecture of the proposed system is shown in Figure 1.

The architecture reflects a sequence of ontology-analytical transformations, within which primary information sources are reduced to a normalized structure, interpreted as a formal ontology, enriched with evaluation results, and transformed into a multidimensional analytical model. In this framework, the ontology serves as an invariant semantic core from which dimensions, hierarchies, indicators, admissible aggregation operations, analytical facts, and a logical multidimensional cube are inductively derived.

Conceptually, the system consists of four functional-conceptual layers:

1) Data Structuring and Ontological Preparation Layer.
2) Ontological Formation, Management, and Cognitive Interpretation Layer.
3) Multicriteria Enrichment and Analytical Transformation Layer.
4) Interpretation, Navigation, and Visualization Layer.

These layers do not represent independent subsystems. Rather, they specify successive functional roles in the implementation of the ontology-analytical transformation. The first layer prepares semantically aligned data, the second constructs and maintains the ontology, the third enriches the ontology and constructs the multidimensional analytical model, and the fourth provides interpretation of the induced analytical structures through projections, rankings, filters, and visual representations.

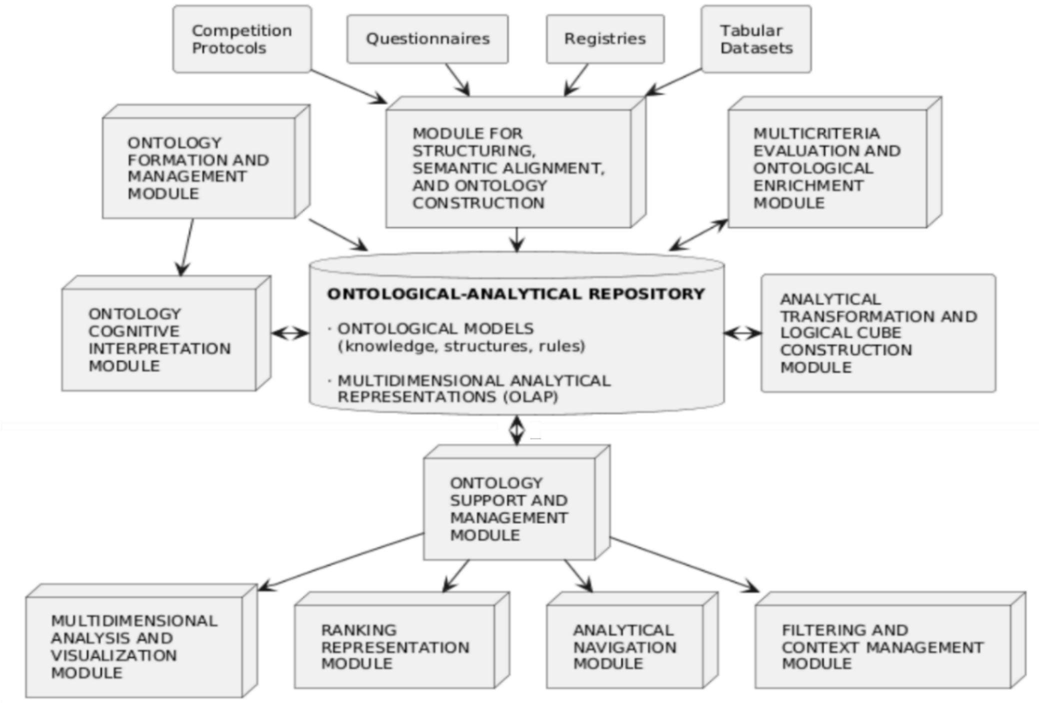

Figure 1: Conceptual architecture of the ontology-driven analytical system.

Table 1: Correspondence between architectural layers and ontology-analytical transformations.

Architectural layer	Formal component	Input	Output	Functional role
Data Structuring and Ontological Preparation	Structural transformation and semantic alignment operator - (1)-(3)	Documents, protocols, tables, registries, textual descriptions	Normalized object-attribute-value structure	Formation of a semantically aligned representation of source data
Ontological Formation, Management and Cognitive Interpretation	Ontological mapping operator and formal ontology model - (4), (6)-(9)	Normalized domain structure	Formal domain ontology with typed elements, relations and interpretation functions	Construction and maintenance of the invariant semantic model
Multicriteria Enrichment and Analytical Transformation - stage 1	Ranking model construction and ontology enrichment operators - (5)	Formal ontology	Enriched ontology with integrated evaluative or ranking values	Integration of multicriteria evaluation results into the ontology as semantically interpreted properties
Multicriteria Enrichment and Analytical Transformation - stage 2	Analytical transformation model, ontology-analytical mapping operator and multidimensional model - (10)-(30)	Enriched ontology	Multidimensional analytical model / logical analytical cube with dimensions, hierarchies, measures, aggregation rules and analytical facts	Construction of the ontology-induced logical cube from semantic structures and fact-generating elements
Interpretation, Navigation and Visualization	Analytical projections and multidimensional interpretation mechanisms - (12), (13), (29), (30)	Analytical facts, indicators, rankings and multidimensional structures	Rankings, analytical slices, filters, visual representations	Practical interpretation and exploration of ontology-induced analytical results

Within this architecture, the third layer is particularly important, as it implements the transition from formal ontology to a multidimensional analytical structure. This transition corresponds to the formal operators introduced in Section 2 and links the extended ontology O^* to the multidimensional analytical model M_{OLAP}.

The first stage of the analytical transformation forms an extended ontology with integrated evaluation values; the second stage transforms this ontology into a multidimensional model containing dimensions, hierarchies, indicators, aggregation rules, and analytical facts. The transformation implemented by the architecture can be represented as the following sequence:

$$S \xrightarrow{\Phi_{olap}} N \xrightarrow{\Phi_{ont}} O \xrightarrow{\Omega_{MR}} O^* \xrightarrow{\Phi_{olap}} M_{OLAP}.$$

This sequence shows that the multidimensional analytical model is not introduced independently of the ontology. It is constructed as a derived structure based on the enriched ontology and the ontology-analytical mapping operator. This correspondence is summarized in Table 1.

3.2 Data Structuring and Ontological Preparation Layer

The first layer transforms heterogeneous sources into a normalized representation suitable for further ontological interpretation. Its input may include electronic forms, competition protocols, registries, tabular datasets, questionnaires, reports and textual descriptions. These sources may differ in structure, terminology, completeness and level of formalization; therefore, the primary task of the layer is to eliminate structural, lexical and semantic heterogeneity.

At this level, the transition from source documents to a normalized object-attribute structure takes place. Such a structure captures the entities of the domain, their attributes, values, types, and basic relationships between them. It is not yet an ontology and does not contain an analytical model. Its function is to prepare a formally consistent representation for subsequent ontological mapping.

This layer includes the Module for Data Structuring and Semantic Alignment. It performs ingestion and consolidation of source data, normalizes terms and attributes, aligns equivalent or related values, and produces a coherent representation that can be used as input for ontology construction.

No comparison, ranking, aggregation or cube construction is performed at this stage. The output of the layer is a normalized structure that preserves the relevant domain content while eliminating inconsistencies that would prevent further semantic interpretation.

3.3 Ontological Formation, Management and Cognitive Interpretation Layer

The second layer transforms the normalized domain structure into a formal ontology and maintains this ontology as the central semantic mechanism of the system. The ontology defines concepts, individual objects, relations, typing rules and interpretation functions that determine how domain data may be consistently understood and further processed.

Within this layer, normalized records are given an ontological interpretation. Values that may have been simple fields at the input structure level can, after ontological mapping, become distinct entities, relationships, or functions. Contextual, classificatory, spatial, temporal, organizational, and personal characteristics are interpreted as ontological elements associated with relevant events or objects in the domain. Numerical or ordered values can be interpreted as ontological functions if they have domain-specific, evaluative, or analytical semantics.

The Ontology Formation and Management Module supports the structure and consistency of the ontology as the domain evolves. It controls changes in concepts, relations, types and interpretation functions without requiring manual redesign of analytical tables.

The Ontology Cognitive Interpretation Module uses the ontology as a source of semantic constraints for subsequent evaluation and analytical transformation. This layer does not calculate rankings, does not perform aggregation and does not construct the multidimensional cube. It defines the semantic conditions under which such operations become valid.

Thus, the ontology formed at this layer serves as the invariant semantic model of the domain. It specifies which elements can later function as contextual dimensions, which relations may define hierarchies, which functions may become measures, and which elements may generate analytical facts.

3.4 Multicriteria Enrichment and Analytical Transformation Layer

The third layer implements the analytical core of the system. At this level, the formal ontology is used not only as a semantic description of the domain but also as the basis for multicriteria enrichment, analytical transformation and construction of the multidimensional analytical model.

This layer includes two functionally distinct modules: the Multicriteria Evaluation and Ontological Enrichment Module and the Analytical Transformation and Logical Cube Construction Module.

The Multicriteria Evaluation and Ontological Enrichment Module defines evaluation criteria, weights and rules of interpretation according to the ontology of the domain. Primary evaluative properties are interpreted within the domain model and transformed into semantically defined evaluative values. The resulting values are integrated into the ontology as properties of the corresponding ontological elements rather than as external technical fields. At this stage, an enriched ontology is formed in which the primary evaluative characteristics of the subject domain may be supplemented with integral rating values, normalized scores, or other interpreted indicators.

The Analytical Transformation and Logical Cube Construction Module implements the transition from the enriched ontology to the multidimensional analytical model. This is the stage at which the logical analytical cube is constructed. Contextual and classificatory ontology elements are mapped to dimensions, ontological relations define hierarchies, analytically meaningful functions are mapped to measures, admissible aggregation rules are assigned to measures, and fact-generating ontology elements produce analytical facts.

This module is responsible for actually building the multidimensional cube. Here, the cube is not necessarily understood as a physical OLAP structure in a specific storage system, but rather as a logical multidimensional analytical model. It includes a set of dimensions, hierarchies, measures, admissible aggregation operations, and analytical facts. It is this model that is subsequently used to build rankings, slices, projections, filters, and visualizations.

As a result, the multidimensional model is not manually imposed on the data. It is derived from the logical organization of the ontology and from the enriched semantic representation of the domain. The analytical cube therefore preserves the connection between domain semantics and analytical operations.

For the domain of intellectual competitions, such a transformation may derive dimensions such as year, region, institution, competition, section, participant and mentor. Measures may include ranking score, achievement indicators and aggregated counts. The admissibility of aggregation depends on the semantic role of the measure: counts may be summed, ranking scores may be averaged or compared, and ordinal characteristics require restricted interpretation.

Thus, the third layer is responsible not only for ranking or evaluation, but also for the formal construction of the ontology-induced multidimensional analytical cube.

3.5 Interpretation, Navigation and Visualization Layer

The fourth layer provides practical access to the results of ontology-driven analysis. It operates on analytical facts, indicators, rankings and multidimensional projections generated by the previous layer. It does not modify the ontology and does not construct a new analytical cube. Instead, it supports interpretation, filtering, visualization and navigation across levels of detail.

At this level, the user gains access to various projections of a single ontologically induced analytical structure. These may include ranking views, aggregated snapshots, comparative projections, filters by dimensions, as well as transitions between aggregated measures and detailed analytical facts.

The layer includes Semantic Consistency Support, Multidimensional Visualization, Ranking Representation, Analytical Navigation and Contextual Filtering. These functions allow users to explore analytical slices, restrict results by selected dimensions, compare aggregated indicators and move between levels of granularity without changing the underlying ontology or the induced multidimensional model.

The fourth layer therefore works with the already constructed logical cube. It produces user-oriented representations of the analytical model but does not change the model itself. This separation prevents visualization, filtering or navigation procedures from altering the semantic basis of the analysis.

Thus, the proposed architecture establishes a coherent interaction between semantic modeling, multicriteria enrichment, analytical transformation, logical cube construction and interpretation. The ontology remains the invariant foundation of the system, while the multidimensional analytical model and user-oriented representations are derived from it in a controlled and semantically consistent manner.

4 RESULTS AND DEMONSTRATION OF THE PROPOSED MECHANISM

4.1 Demonstration Scenario and Input Records

To demonstrate the formal instantiation of the proposed mechanism, we consider the problem of multidimensional analysis of the results of intellectual competitions. This subject area is representative of the proposed approach, as it combines diverse data sources, the organizational structure of competitions, personal relationships between participants and mentors, the spatial distribution of educational institutions, as well as evaluation characteristics that require a coordinated analytical representation.

The system receives a set of documents defined in the formal model (1). Each document may correspond to a competition protocol, a participant registry, a results table, a list of sections, a list of educational institutions, or a description of a competition event. The demonstration fragment examines four records of student participation in competitions. A sample of the input records is presented in Table 2.

Each entry in Table 2 describes not an isolated numerical result, but an instance of a specific student's participation in a particular competitive context.

Therefore, the entry s_i is interpreted as a source for constructing an individual ontological element Participation:

$$s_i \to p_i,\ \tau_X(p_i) = Participation.$$

In the example snippet, the set of elements looks like this:

$$\mathcal{P} = \{p_1, p_2, p_3, p_4\}.$$

The Participation element is interpreted as an ontological representation of a participation event, rather than as the technical equivalent of a table row. It integrates personal, organizational, temporal, spatial, classificatory, and evaluative contexts, on the basis of which analytical facts will subsequently be derived.

4.2 From Input Fields to Ontological Interpretation

At the first stage, the structural normalization operator Φ_{norm} defined in (2) is applied. Its result is a normalized object-attribute structure N formally specified in (3).

For the corresponding record s_i, the normalized object p_1 can be represented as a structured description of Student 1's participation in the National research competition in 2025:

$$p_i = \begin{pmatrix} 2025, NRC, CS, AI, KyivRegion, \\ Lyceum_A, Student_1, Mentor_1, 92, I \end{pmatrix}.$$

At the level of the normalized structure, the fields competition, student, region, institution, section, and mentor may be represented as attributes of the object p_i. However, after ontological mapping, they do not remain simple attribute values. The corresponding values are interpreted as separate ontological entities connected with Participation through semantic relations. Table 3 summarizes how the main input fields are interpreted in the ontological model.

Table 2: Fragment of input competition records.

Record	Year	Competition	Field	Section / discipline	Region	Institution	Student	Mentor	Protocol score	Place
s_1	2025	National research competition	Computer Science	Artificial intelligence systems	Kyiv region	Lyceum A	Student 1	Mentor 1	92	I
s_2	2025	National research competition	Computer Science	Artificial intelligence systems	Lviv region	Lyceum B	Student 2	Mentor 2	84	II
s_3	2025	Informatics olympiad	Informatics	Programming	Kyiv region	Lyceum A	Student 3	Mentor 1	78	III
s_4	2024	National research competition	Applied Mathematics	Mathematical modeling	Kharkiv region	Gymnasium C	Student 4	Mentor 3	88	II

Table 3: Interpretation of input fields in the ontological model

Input field	Normalized representation	Ontological interpretation
competition	value of competition attribute in p_i	individual of class Competition, connected with p_i by *hasCompetition*
student	value of student attribute in p_i	individual of class Student, connected with p_i by *hasStudent*
mentor	value of mentor attribute in p_i	individual of class Mentor, connected with p_i by *hasMentor*
institution	value of institution attribute in p_i	individual of class Institution, connected with p_i by *hasInstitution*
region	value of region attribute in p_i	individual of class Region, connected with p_i by *hasRegion*
field	value of field attribute in p_i	individual of class Field, connected with p_i by hasField
section / discipline	value of section or discipline attribute in p_i	individual of class Section or Discipline, connected with p_i by *hasSection* or *hasDiscipline*
year	temporal value in p_i	contextual value connected with p_i by $hasYear$
protocol score	numerical value in the protocol	source evaluative function $ProtocolScore(p_i)$
place	ordered competition result	ordinal evaluative function $Place(p_i)$

This distinction is essential for the correct transition from a tabular representation to an ontological model. Not every field of a normalized record is an attribute in the narrow ontological sense, and not every numerical value automatically becomes an analytical measure. For example, the year value is numerical, but it functions as a temporal coordinate. By contrast, the protocol score and award place have evaluative semantics because they characterize the result of a specific participation.

4.3 Ontological Fragment Centered on Participation

At the second stage, the normalized structure is interpreted as a formal ontology using the operator Φ_{ont} defined in (4). The structure of the formal ontology $O$is defined in (6), while its typing, relations, and functions are specified according to (7)-(9).

Within the demonstration domain, the ontology fragment is constructed around the Participation element, which serves as the central fact-forming node (Fig. 2).

In this fragment, the relations between *Participation* and the entities *Student*, *Mentor*, *Competition*, *Year*, *Section*, *Institution*, and *Region* belong to the set of ontological relations defined in (8). They specify the semantic context of participation. The evaluative characteristics *ProtocolScore*, *Place*, and *RankingScore* belong to the set of ontological functions F_O defined in (9), since they are defined as functions on *Participation*.

For the corresponding Participation instance p_1, the relations are instantiated as follows:

- $hasStudent(p_1, Student\ 1)$;
- $hasMentor(p_1, Mentor\ 1)$;
- $hasCompetition(p_1, NationalResearchComp)$;
- $hasYear(p_1, 2025)$;
- $hasField(p_1, ComputerScience)$;
- $hasSection(p_1, ArtificialIntelligenceSystems)$;
- $hasInstitution(p_1, Lyceum\ A)$;
- $hasRegion(p_1, KyivRegion)$;

The primary evaluative characteristics are defined as functions on the set of participations:

$$ProtocolScore: \mathcal{P} \to V_{score},\ Place: \mathcal{P} \to V_{place}.$$

For the first participation:

$$ProtocolScore(p_1) = 92,\quad Place(p_1) = I.$$

4.4 Multicriteria Enrichment and Logical Cube Construction

After constructing the ontology, the multicriteria evaluation model is applied $MR,$ and the evaluation results are integrated into the ontology by the operator Ω_{MR}defined in (5). In the demonstration example, the model converts the primary evaluative characteristics of a participation event - the protocol score and award place - into an integral rating value.

The following function is used to illustrate this:
$R(p_i) = 0.7 \cdot ProtocolScore(p_i) + 30 \cdot W_{place}(p_i),$
where: $W_I = 1.00,\quad W_{II} = 0.85,\quad W_{III} = 0.70.$
This function is a demonstration instance of the model. In a complete system, the evaluation model may take into account the level of the competition, the weight of the section, the type of competition, scale normalization rules, criterion weights, and other parameters defined by the domain ontology. The results of enriching the Participation elements are presented in Table 4.

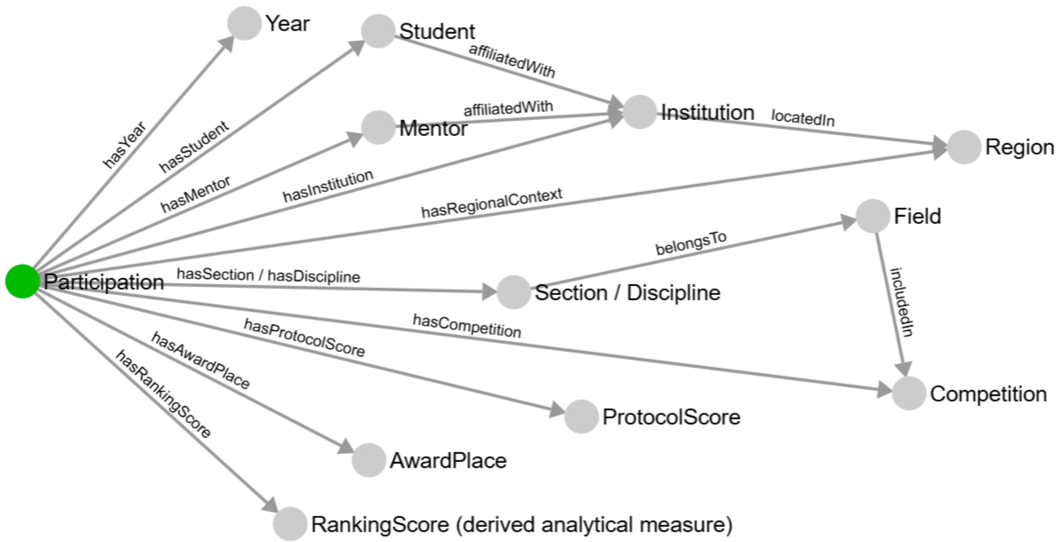

Figure 2: Fragment of the domain ontology centered on participation.

Table 4: Multicriteria enrichment of participation elements.

Participation	*Protocol score*	*Place*	W_{place}	*RankingScore*
p_1	92	I	1.00	94.4
p_2	84	II	0.85	84.3
p_3	78	III	0.70	75.6
p_4	88	II	0.85	87.1

After applying the evaluation model Ω_{MR}, the ontology transitions to an enriched state O^* in which the rating values are interpreted as ontological functions of participation:

$$RankingScore(p_1) = 94.4;$$
$$\mathrm{RankingScore}(p_2) = 84.3;$$
$$\mathrm{RankingScore}(p_3) = 75.6;$$
$$RankingScore(p_4) = 87.1.$$

Next, the ontology-analytical mapping operator Φ_{olap} defined in (14) is applied to map the extended ontology O^* to the multidimensional analytical model M_{OLAP} defined in (10). In this section, M_{OLAP} is viewed as a logical analytical cube, that is, as a structure containing dimensions, hierarchies, measures, admissible aggregation operations, and analytical facts.

In the example provided, the dimensions are inductively derived from ontological elements related to *Participation*:

$$D = \begin{Bmatrix} Year, Region, Institution, Competition, \\ Field, Section, Discipline, Student, Mentor \end{Bmatrix}.$$

Measures are not derived from all numerical attributes, but only from those ontological functions or derived values that have analytical semantics with respect to the fact-forming element. For example:

$$M = \{RankingScore, AchievementIndicator\}.$$

The achievement indicator is defined for individual participationas:

$$AchievementInd(p_i) = \begin{cases} 1, & Place(p_i) \in \{I, II, III\}, \\ 0, & otherwise. \end{cases}$$

Measures for the number of participations or achievements are generated at the level of an aggregated analytical group G, for example, a region, educational institution, section, or supervisor:

$$ParticipationCount(G) = |\mathcal{P}_G|,$$
$$AchievementCount(G) = \sum_{p_i \in \mathcal{P}_G} AchievementInd(p_i),$$

where $\mathcal{P}_G$ - is the set of participations belonging to the group G. The set of basic aggregation operations is defined in (22), and for each indicator, a semantically admissible subset of operations is determined according to (23). The semantic roles of indicators and admissible aggregation operations are summarized in Table 5.

Whether aggregation is permissible depends not only on the type of value but also on its subject-specific semantics. Therefore, participation counts can be summed, achievement indicators can be aggregated, and ranking scores can be summed, averaged or compared, but ordinal characteristics such as rank should not be arbitrarily summed.

The hierarchies of the logical cube are derived from ontological relations R_o (8). The following links are relevant in the demo domain:

$$Student \rightarrow Institution \rightarrow Region;$$
$$Mentor \rightarrow Institution \rightarrow Region;$$
$$Section \rightarrow Field \rightarrow Competition;$$
$$Institution \rightarrow Region \rightarrow Country.$$

Thus, the logical analytical cube is derived from an extended ontology rather than being a pre-designed data schema.

4.5 Ontologically Induced Analytical Facts

The fact-forming ontological elements are defined in the formal model (25). In the demonstration domain, the set of instances is a subset of these elements: $\mathcal{P} \subseteq X_F$, since for each entry $p_i \in \mathcal{P}$ there are functions that have an evaluative or analytical interpretation:

$$ProtocolScore(p_i) \in V_{score}, \quad Place(p_i) \in V_{place}, \quad RankingScore(p_i) \in V_{rank}$$

For each instance p_i a set of relevant dimensions is determined D_{p_i}, which corresponds to the formal description of context-dependent axes of analysis in (27). For the participation instance p_1:

$$D_{p_1} = \begin{Bmatrix} Year, Region, Institution, Competition, \\ Field, Section, Student, Mentor \end{Bmatrix}.$$

The values of these dimensions in the context of a specific instance are determined by the function δ (28). For the participation instance p_1:

$$\delta_{p_1}(Year) = 2025, \quad \delta_{p_1}(Region) = KyivRegion,$$
$$\delta_{p_1}(Institution) = Lyceum\ A,$$
$$\delta_{p_1}(Competition) = NationalResearchCompetition,$$
$$\delta_{p_1}(Field) = ComputerScience,$$
$$\delta_{p_1}(Section) = ArtificialIntelligenceSystems,$$
$$\delta_{p_1}(Student) = Student\ 1, \delta_{p_1}(Mentor) = Mentor\ 1$$

The formulation of an analytical fact follows the general structure f_{OLAP} (13), and a specialized ontologically induced mapping μ_F, (29). For the participation instance p_1, based on the RankingScore measure, we obtain the following analytical fact:

$$f_{OLAP}(p_1, RScore) = \begin{pmatrix} 2025, KyivRegion, Lyceum\ A, \\ NRC, CS, AI, Student\ 1, \\ Mentor\ 1, RankingScore, 94.4 \end{pmatrix}$$

Table 5: Measures and admissible aggregation operations.

Ontological function and Semantic role	Analytical use	Admissible aggregation
$ProtocolScore(p_i)$, source evaluative function	input for ranking model; optional analytical variable	avg, min, max, if analytically allowed
$Place(p_i)$, ordinal evaluative function	input for ranking model	comparison, ordering
$RankingScore(p_i)$, derived analytical measure	ranking and aggregation	sum, avg, min, max
$AchievementInd(p_i)$, *derived analytical measure*	basis for counting achievements	sum, count
$Year(p_i)$, contextual value	dimension value	not a measure

Table 6: Analytical facts induced from Participation elements.

Fact	Year	Region	Institution	Competition	Field	Section / discipline	Student	Mentor	Measure	Value
f_1	2025	Kyiv region	Lyceum A	National research competition	Computer Science	Artificial intelligence systems	Student 1	Mentor 1	RankingScore	94.4
f_2	2025	Lviv region	Lyceum B	National research competition	Computer Science	Artificial intelligence systems	Student 2	Mentor 2	RankingScore	84.3
f_3	2025	Kyiv region	Lyceum A	Informatics olympiad	Informatics	Programming	Student 3	Mentor 1	RankingScore	75.6
f_4	2024	Kharkiv region	Gymnasium C	National research competition	Applied Mathematics	Mathematical modeling	Student 4	Mentor 3	RankingScore	87.1

The analytical facts induced from the demonstration recordsare presented in Table 6.

For the same participation instance, other analytical facts may be generated when the indicator changes. For example, for the achievement indicator:

$$f_{OLAP}(p_1, AchInd) = \begin{pmatrix} 2025, KyivRegion, Lyceum\ A, \\ NRC, CS, AI, Student\ 1, \\ Mentor\ 1, AchInd, 1 \end{pmatrix}$$

Thus, a single Participation element can generate several analytical facts if it is associated with multiple analytically significant functions or derived indicators. The set of such facts f_{OLAP} is formally defined in (30).

4.6 Multidimensional Aggregation and Analytical Projections

Once the analytical facts have been formed, the multidimensional function defined in (12) provides indicator values for various combinations of dimensions. For example, a regional projection is formed by aggregating facts by dimension $Region.$

For each regional group $G,$ the number of participations is determined as:

$$ParticipationCount(G) = |\mathcal{P}_G|.$$

The sum of the rating values:

$$TotalRankingScore(G) = \sum_{p_i \in \mathcal{P}_G} RankingScore(p_i),$$

and the average rating:

$$AverageRankingScore(G) = \frac{TotalRankingScore(G)}{ParticipationCount(G)}.$$

The regional analytical projection is presented in Table 7.

Table 7: Regional analytical projection.

Region	Number of participations	Sum of ranking scores	Average ranking score
Kyiv region	2	170.0	85.0
Lviv region	1	84.3	84.3
Kharkiv region	1	87.1	87.1

Similar projections can be constructed based on other dimensions, such as educational institution, supervisor, section, field, or competition. These results are not distinct independent models, but rather different projections of a single multidimensional structure M_{OLAP} induced from the ontology.

Changing the level of analysis does not require redesigning the data schema. It involves selecting a different subset of dimensions and applying semantically valid aggregation operations to the corresponding measures. The key result is that the analytical structure is formed not through prior manual design of an OLAP schema, but through an ontological interpretation of participation as an event to which personal, organizational, spatial, temporal, classificatory, and evaluative characteristics are attached. In such a model, the participation result is not an isolated numerical field, but a function defined on the ontological element Participation.

The resulting multidimensional model supports analytical views at the level of the student, the administrator, the educational institution, the region, the section, the field, or the competition. All of these views are different projections of a single ontologically induced structure. This allows for maintaining the semantic integrity of the data, controlling the validity of aggregation, and correctly working with context-dependent dimensions.

5 CONCLUSIONS

The article develops a formally grounded approach to the construction of ontology-driven information and analytical systems, in which the formal domain ontology serves as the primary invariant knowledge model, used for the inductive formation of analytical structures, indicators, hierarchies, and admissible aggregation operations. The proposed approach is oriented toward the consistent integration of the semantic model of the domain and multidimensional analytics within a unified formal framework, while minimizing ad hoc assumptions through explicitly defined ontology-analytical mapping rules.

A key result of the study is the formalization of the ontology-analytical mapping operator, which defines a formal transition from the ontological model to a multidimensional analytical representation, taking into account the typing of ontological elements, semantic relations, and interpretative functions. Such an operator supports the derivation of analytical dimensions, hierarchies, and indicators directly from the logical organization of the domain, while preserving semantic constraints at the stages considered in the proposed transformation model.

Within the proposed formalization, it is shown that analytical facts can be induced from ontological elements and their interpretations without the prior specification of fixed analytical schemas. This makes it possible to formally describe analytical situations with context-dependent or partially defined dimensions, support the semantic consistency of aggregation, and reduce the dependence of the logical analytical model on specific implementations of physical data storage or computation.

The proposed architecture of the ontology-driven analytical system formalizes the interaction of semantic, analytical, and interpretative components in the form of functional-conceptual layers with clearly defined roles. Such an architecture separates data structuring, ontology construction, multicriteria enrichment, analytical transformation, logical cube construction, and user-oriented interpretation, while preserving the connection between the semantic model of the domain and the induced analytical structures.

The developed approach is general in nature and can be applied to various classes of information and analytical tasks, including the analysis and evaluation of intellectual and educational achievements as a specific applied case. The demonstration presented in the article illustrates how input records can be transformed into Participation-centered ontological elements, enriched with evaluation values, and further used to induce dimensions, measures, aggregation rules, and analytical facts.

The obtained results provide a formal foundation for further research on ontology-driven analytical systems, including the automated derivation of analytical models, the integration of multicriteria evaluation with semantic structures, and the development of semantically traceable analytical and ranking mechanisms in complex information and analytical environments.

ACKNOWLEDGMENTS

The authors express their gratitude to colleagues and experts who contributed to the discussion of the conceptual provisions of the study.

REFERENCES

[1] R. Djiroun, K. Boukhalfa and Z. Alimazighi, “Data cubes retrieval and design in OLAP systems: From query analysis to visualisation tool,” International Journal of Business Intelligence and Data Mining, vol. 14, nos. 1/2, pp. 267-298, 2019, doi: 10.1504/IJBIDM.2019.096813.

[2] A. Abelló et al., “Using semantic web technologies for exploratory OLAP: A survey,” IEEE Transactions on Knowledge and Data Engineering, vol. 27, no. 2, pp. 571-588, 2015, doi: 10.1109/TKDE.2014.2330822.

[3] D. Martinez-Mosquera, R. Navarrete, S. Luján-Mora, L. Recalde and A. Andrade-Cabrera, “Integrating OLAP with NoSQL databases in big data environments: Systematic mapping,” Big Data and Cognitive Computing, vol. 8, no. 6, art. 64, 2024, doi: 10.3390/bdcc8060064.

[4] M. Ptiček, B. Vrdoljak and M. Gulić, “The potential of semantic paradigm in warehousing of big data,” Automatika, vol. 60, no. 4, pp. 393-403, 2019, doi: 10.1080/00051144.2019.1630582.

[5] T. R. Gruber, “A translation approach to portable ontology specifications,” Knowledge Acquisition, vol. 5, no. 2, pp. 199-220, 1993, doi: 10.1006/knac.1993.1008.

[6] N. Guarino, “Understanding, building and using ontologies,” International Journal of Human-Computer Studies, vol. 46, nos. 2-3, pp. 293-310, 1997, doi: 10.1006/ijhc.1996.0091.

[7] M. Uschold and M. Gruninger, “Ontologies: Principles, methods and applications,” Knowledge Engineering Review, vol. 11, no. 2, pp. 93-136, 1996, doi: 10.1017/S0269888900007797.

[8] S. Staab and R. Studer, Eds., Handbook on Ontologies. Berlin, Heidelberg: Springer, 2009, doi: 10.1007/978-3-540-92673-3.

[9] R. P. Deb Nath, O. Romero, T. B. Pedersen and K. Hose, “High-level ETL for semantic data warehouses,” Semantic Web, vol. 13, no. 1, pp. 85-132, 2022, doi: 10.3233/SW-210429.

[10] A. L. Antunes, J. Barateiro and E. Cardoso, “Application of semantic web techniques in DW/BI systems for strategic management,” IEEE Access, vol. 13, pp. 163103-163119, 2025, doi: 10.1109/ACCESS.2025.3610577.

[11] H. P. Bomma, “Enhancing business analytics with a semantic layer,” International Journal of Business Quantitative Economics and Applied Management Research, vol. 7, no. 9, pp. 93-99, 2023.

[12] O. Romero and A. Abelló, “Automating multidimensional design from ontologies,” in Proceedings of the ACM 10th International Workshop on Data Warehousing and OLAP (DOLAP ’07), Lisbon, Portugal, 2007, pp. 1-8, doi: 10.1145/1317331.1317333.

[13] R. Mizoguchi and J. Bourdeau, “Using ontological engineering to overcome AI-ED problems: Contribution, impact and perspectives,” International Journal of Artificial Intelligence in Education, vol. 26, no. 1, pp. 91-106, 2016, doi: 10.1007/s40593-015-0077-5.

[14] G. Futia and A. Vetrò, “On the integration of knowledge graphs into deep learning models for a more comprehensible AI: Three challenges for future research,” Information, vol. 11, no. 2, art. 122, 2020, doi: 10.3390/info11020122.

[15] R. Confalonieri and G. Guizzardi, “On the multiple roles of ontologies in explanations for neuro-symbolic AI,” Neurosymbolic Artificial Intelligence, vol. 1, pp. 1-14, 2025, doi: 10.3233/NAI-240754.

[16] M. Nadutenko, V. Prykhodniuk, V. Shyrokov and O. Stryzhak, “Ontology-driven lexicographic systems,” in Advances in Information and Communication. FICC 2022. Lecture Notes in Networks and Systems, vol. 438. Cham, Switzerland: Springer, 2022, pp. 204-215, doi: 10.1007/978-3-030-98012-2_16.

[17] O. Stryzhak, V. Horborukov, V. Prychodniuk, O. Franchuk and R. Chepkov, “Decision-making system based on the ontology of the choice problem,” Journal of Physics: Conference Series, vol. 1828, no. 1, art. 012007, 2021, doi: 10.1088/1742-6596/1828/1/012007.