

## Architecture and Structural Organization of Terminological Databases

ABDUKHALIMOV ABDULAZIZ ABDUMUTAL UGLI  
*Andijan State Technical Institute, Uzbekistan*

### ABSTRACT

*This article examines the architecture, structural organization, and theoretical-methodological approaches employed in the development of modern terminological databases (TDB). The study presents a detailed analysis of the multi-level architectural model of TDBs, encompassing conceptual, logical, and physical tiers, and explores the interrelationship between macrostructural and microstructural components. Comparative analysis is conducted across thesaurus-based, ontological, and linguo-cognitive approaches to terminology, elucidating their respective roles in knowledge representation and transfer. The seven core elements of TDB microstructure headword, lexical-grammatical features, multi-component terminological phrases, definition, synonyms, related terms, and thematic lexis are systematically characterized. Furthermore, the applicability of principles emerging from digital technology advancement modularity, dynamism, and flexibility to TDB construction is discussed. The findings indicate that contemporary TDBs function not merely as tools for terminological standardization, but as integrated platforms for the management and transmission of scientific knowledge.*

**Keywords:** Terminological database, architecture, macrostructure, microstructure, ontology, thesaurus, terminology, linguo-cognitive approach, digital lexicography, knowledge management.

## INTRODUCTION

From a linguistic perspective, the analysis of scientific language represents an effective approach to semantic research and the systematization of terminology. Terminology constitutes a body of concepts formed as a result of humanity's scientific, research, and productive activities, reflecting the processes of cognizing and transforming the surrounding environment. Given that terminological work is grounded in cognitive, linguistic, and communicative principles, a broad research approach to terminology is recommended. Various disciplines continuously interact with one another; the synthesis of different scientific fields gives rise to new disciplines, which in turn influences the integration of scientific knowledge [1].

## THEORETICAL FOUNDATIONS OF TERMINOLOGICAL DATABASES

A review of the theoretical literature supports the view that there exists a methodological and theoretical foundation for developing a knowledge transfer model that incorporates thesaurus-based, ontological, and linguo-cognitive principles of organizing terminological material. The contemporary cognitive approach to terminology emphasizes wide-ranging research in this field. This domain referred to as cognitive-communicative terminology investigates the collective ideas of various linguistic disciplines, including the capacity to analyze terms and thermodynamics from innovative perspectives. Numerous studies by researchers such as S. V. Grinev (1993), L. M. Alekseyeva (1998), E. I. Golovanova (2008), S. L. Mishlanova (2011), and V. F. Novodranova (2013) confirm this trend [2]. There is also growing interest in cognitive research related to foreign terminology, which explores connections between various fields such as terminology and translation, lexicography, and computational linguistics. Therefore, a need exists to develop new methods of processing and analyzing terminology that lead to the integration of traditional terminological methods with the principles of ontology construction. The development and structural organization of terminological databases are typically

grounded in the analysis of large data arrays that emphasize the relationship between terminology, computational lexicography, and corpus linguistics.

When working with terminology, it is essential to recognize the dynamic and evolving nature of the terminological system and the significance of a synergetic approach [3]. This approach demonstrates that as a subject domain develops, points of inconsistency emerge within the terminological system of a given language, requiring a reassessment of established terminological frameworks to accurately reflect the corresponding conceptual system.

A semantic approach is also applied in terminology management systems. Each dictionary entry in a terminological database contains lexical units representing a single concept, along with classificatory and descriptive information [4]. Users have the ability to independently determine the number of descriptive fields and to integrate multimedia components and cross-references. From this perspective, a terminological database typically has a complex structure comprising several components: informational (data on the use of terms and documents), linguistic (search for a term in specific languages – source and target language – with search parameters such as terminoid → quasi-term → term), technical (search parameters related to linguistic and conceptual aspects), and organizational (external references to additional sources).

The thesaurus approach is employed in the development of terminological databases. This approach not only reflects the conceptual relations of a term within the terminological structure (logos) but also emphasizes its semantic, syntactic, and lexical characteristics (lexis). The thesaurus serves as a structured representation of material functioning as a “dictionary that includes semantic relationships between a conceptual entry and its elements” [5]. It operates as a model of the logical-semantic organization of terminology, combining thematic classifications with hierarchical and non-hierarchical relationships between lexical units.

In constructing a thesaurus, researchers move from individual terms to a conceptual system, then return to terms to

arrange them in accordance with that system [6]. The structural organization of terminology facilitates subsequent editing and preparation of machine translation outputs. The need to include specialized terms is associated with their widespread occurrence in specialized texts, which reflects their role in professional communication.

Practical terminological activity involves the creation of terminological dictionaries, also referred to as terminological lexicography or terminography. The development of terminological databases and automated knowledge bases is among the most important areas of activity in the field of terminology, aimed at standardizing terms through the facilitation, simplification, and harmonization of the translation of scientific and technical literature [7].

Creating a terminological system in a dictionary requires the selection of potential terms and a logical and conceptual analysis of the dictionary by grounding the dictionary project in a conceptual framework (whereby each terminological unit is associated with a specific element of the corresponding logical and conceptual system, serving to verbally express concepts within a particular field of scientific activity). Furthermore, terms must be organized either alphabetically or in accordance with systematically arranged elements, depending on the purpose of the dictionary and the data set defined in the dictionary's specifications [8].

With the advancement of digital technologies, a need arose to organize data in a format processable by computers. This necessitated the introduction of new approaches that harmonize ontological and terminological principles. Ontology reflects the complex relationships between concepts within a specific domain. It is distinguished from other knowledge organization systems by its higher degree of semantic detail. Ontologies may be constructed at various levels of detail: lists (the simplest form), taxonomies, thesauri (lower-level ontologies encompassing the domain's logical and conceptual relationships), and formal ontologies (the most detailed form). While ontology serves as a tool for formalizing knowledge, a thesaurus enables

the semantic description of concepts and terms within a given domain [9].

Ontology plays a decisive role in automated language processing, wherein sets of linguistic expressions (terminological phrases) correspond to specific conceptual systems and linguistic resources (such as dictionaries) define the conceptual foundations of a subject domain [2].

Thesauri organize terms together with their synonyms and related concepts based on linguistic associations, whereas ontologies emphasize conceptual relationships by defining logical connections between concepts. This distinction is particularly relevant for building terminological systems that represent specialized languages and are characterized by a dynamic approach to knowledge representation [10].

The classification of terminological dictionaries is determined by factors such as thematic scope, novelty of the dictionary, nature of entries, purpose and functions of the dictionary, and the range of languages covered. Specialized terminological dictionaries are intended for professional translators working with scientific and technical texts, and may be classified as “registering” (reflecting the current state of industrial terminology), “regulatory” (containing terms that must be used in accordance with official documentation), or “informational” (reflecting semantic relationships between terms to provide assistance) [11].

The principles applied in the development of automated systems and terminological databases include “modularity” (the relative independence of components within the structure and their capacity for integration into a unified system), “dynamism” (the ability to make operational updates to the dictionary), and “flexibility” [12]. These principles constitute the fundamental tenets of digital lexicography and serve as a methodological basis for the creation of modern linguistic databases. Adherence to such principles is equally necessary in Uzbek linguistics, as language is a continuously evolving process in which new words, terms, and meanings frequently emerge. Systems based on modular and dynamic approaches are capable of adapting to the natural changes in language and keeping the database up to date.

Flexibility, in turn, ensures a user-friendly interface for various user groups linguists, translators, students, and AI systems alike.

It is thus clear that terminological databases (TDBs, or term databases) represent collections of key terms that must be consistently interpreted and translated throughout a given document or project. Each term in a terminological database is provided with a comprehensive set of metadata, including the semantic definition of the term, its equivalents in other languages, abbreviated forms, synonymic variants (both recommended and non-recommended), scope of use, as well as references to normative documents such as national and international standards. Terms in the database are also categorized according to their intended use: 1) those intended to support translation services for scientific-technical literature and documentation; 2) those intended to provide information on standardized and recommended terminology.

#### TDB ARCHITECTURE AND STRUCTURE

Terminological databases (TDBs) represent an essential component of modern information technology systems, designed for the storage, management, and dissemination of standardized terminology across various domains. The architecture of a TDB is complex, encompassing a data model, organizational structure, functional components, and a technological platform [13].

It should be noted that a terminological database is a specialized information system that systematically stores terms, their definitions, contextual data, interlingual relationships, and metadata pertaining to a specific domain or set of domains. Unlike conventional dictionaries, a TDB does not merely store terms; it also records their semantic relationships, historical evolution, usage context, and versions in different languages. These multifunctional capabilities and the complex structure of a TDB are realized through a multi-level architectural model, which ensures the system's efficient operation and adaptability (Figure 1).

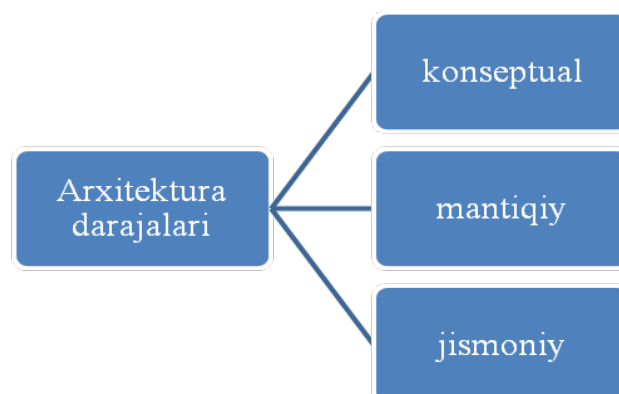


Figure 1. *Architectural levels of terminological databases*

The conceptual level defines the logical structure of the TDB and encompasses key elements such as the terminological entry, conceptual system, and multilingual model. The logical level determines the data model and the principles of its organization. At this level, the relational model, object-oriented model, and semantic network model are applied. The physical level determines how data is stored in memory and disk systems. At this level, storage structures and caching mechanisms are of particular importance.

The architecture of terminological databases is a multi-layered and complex structure that must satisfy modern technological requirements and user needs. An effective TDB architecture integrates a well-designed data model, flexible functional components, high processing speed, and reliable security mechanisms.

#### MACROSTRUCTURAL AND MICROSTRUCTURAL ELEMENTS

A terminological database is developed with due consideration of macro- and microstructures. The macrostructure consists of six entries that provide access to the database: 1) a concept or concepts manifested in a particular action-event; 2) the primary instrumental function; 3) the logical and epistemological classification of the concept; 4) the semantic segment of the text;

5) visual context; 6) a reference or encyclopedic article on the concept. Three of these the action-event, the primary instrumental function, and the logical-epistemological classification are innovative for this format.

The macrostructure of a terminological database serves as an interface between the user and the system. Each entry represents a separate informational layer, and together they provide a multifaceted description of the term.

The “action-event concept” reflects the dynamic context of the term. This approach, unlike traditional static dictionary systems, demonstrates the application of a term in real processes [14]. For example, the term “synthesis” manifests in the process of a chemical reaction, while “transmission” manifests in the event of mechanical motion transfer. This approach aligns with L. Vygotsky’s theory of concept development, as it emphasizes the study of terms in the context of actions [15].

The “primary instrumental function” reveals the pragmatic dimension of the term. This element clarifies the purposes and functions for which a term is employed. The concept of “instrumental function” is particularly important in technological and scientific terminology, as many terms denote names of tools, devices, or methods. This approach is associated with the principles of functional semantics emphasized in the linguo-cultural studies of Yu. S. Stepanov [16].

“Logical and epistemological classification” determines the position of a term within the system of knowledge. This classification combines ontological (pertaining to the nature of existence) and gnoseological (pertaining to the process of cognition) characteristics. For example, “electric current” as a physical phenomenon belongs to the ontological category, while the methods of measuring it belong to the gnoseological category. This approach develops the principles set forth in D.S. Lotte’s philosophy of terminology [17].

The semantic segment of the text defines the internal structure of each terminological entry and comprises a number of components. The lexicographic zone consolidates all morphological variants of a term, its grammatical information, and pronunciation features. This zone fully reflects the

morphological paradigm and indicates the term's position within the linguistic system. The semantic zone provides a detailed analysis of the meaning components of the term. Here, the primary meaning (denotative), secondary meaning (connotative), and contextual meanings are distinguished. The method of componential analysis is applied, which reveals the seme structure of the term.

#### CONCLUSION

In conclusion, terminological databases represent a strategic component of modern information technology systems, and their success is determined by a scientifically grounded multi-level architecture and a substantive structural approach. An architecture that integrates conceptual, logical, and physical levels, along with the coherent integration of macro- and microstructures, creates opportunities for users to obtain comprehensive information. The introduction of innovative elements such as the action-event concept, the instrumental function, and epistemological classification opens pathways for studying terms in their dynamic and practical contexts. As a result, TDBs are evolving not merely into tools for terminological standardization, but into modern platforms for the management and transmission of scientific knowledge.

#### REFERENCES

1. Tabanakova, V. D. 2023. Interdisciplinary synthesis of linguistic scientific knowledge. *Bulletin of PNRPU. Issues in Linguistics and Pedagogy*, 2, 8-24.
2. Fedyuchenko, L. G. 2021. *Terminological Database as a Transfer Model of Technical Knowledge* [Author's abstract of Doctoral dissertation in Philology]. Tyumen, 42.
3. Belyaeva, L. N. 2017. Translation dictionary in the synergetic aspect. *Scientific Notes of St. Petersburg University of Management Technologies and Economics*, 4/60, 29-34.
4. Kalinina, S. V. 1998. Terminological database in the system of automatic text processing: Vocabulary selection criteria. *Problems of Linguistics and Communication Theory*, 128-136.

5. Fedyuchenko, L. G. 2017. Structural and functional parameters of a terminological database. *Tyumen State University Herald. Humanities Research. Humanities*, 3/2, 45-48.
6. Nikitina, S. E. 2014. *Semantic Analysis of the Language of Science: Based on Linguistic Material*. Moscow: Librokom.
7. Miller, G. *WordNet: A Lexical Database for English*. Available online: <<https://aclanthology.org/H92-1116>>.
8. Leychik, V. M. 2022. *Terminology Science: Subject, Methods, Structure*. Moscow: LENAND, 248.
9. Belyaeva, L. N., Gerd, A. S. & Ubin, I. I. 1996. Automation in lexicography. In *Applied Linguistics* (pp. 68-90). St. Petersburg: St. Petersburg University Press.
10. Dobrov, B. V. & Lukashevich, N. V. 2007. Linguistic ontology on natural sciences and technologies for information retrieval applications. *Proceedings of Kazan State University. Physics and Mathematics Series*, 149/2, 49-72.
11. Zenko, M. I. 2023. Branding terminology in the ontological aspect. *Russian Linguistic Bulletin*, 7/43, 147-151.
12. Gerd, A. S. 1996. Scientific and technical lexicography. In *Applied Linguistics* (pp. 287-307). St. Petersburg: St. Petersburg University Press.
13. Belyaeva, L. N. 2010. The potential of automated lexicography and applied linguistics. *Izvestia: Herzen University Journal of Humanities & Sciences*, 134, 70-79.
14. ISO 12620:2019 Terminology and other language and content resources – Specification of data categories and management of a data category registry for language resources.
15. Danilenko, V. P. 1977. *Russian Terminology: An Experience of Linguistic Description*. Moscow: Nauka.
16. Vygotsky, L. S. 1999. *Thinking and Speech*. Moscow: Labirint.
17. Stepanov, Y. S. 2004. *Constants: Dictionary of Russian Culture*. Moscow: Akademicheskyy Proekt.
18. Lotte, D. S. 1982. *Issues of Borrowing and Ordering Foreign-Language Terms and Terminological Elements*. Moscow: Nauka.

**ABDUKHALIMOV ABDULAZIZ ABDUMUTAL UGLI**  
ASSISTANT, DEPT. OF LANGUAGES AND HUMANITIES,  
ANDIJAN STATE TECHNICAL INSTITUTE, UZBEKISTAN.  
E-MAIL: <ABDUXALIMOVABDULAZIZ1@GMAIL.COM>