

Slovak Terminology Database

Mária Šimková^a, Jana Levická^b, Velislava Stoykova^c

**Ľudovít Štúr Institute of Linguistics, Slovak Academy of Sciences^{a,b},
Institute for Bulgarian Language Prof. Lyubomir Andreychin – Bulgarian Academy of Sciences^c
marias@korpus.sk, janal@korpus.sk, vstoykova@yahoo.com**

Abstract

The authors of this paper aim to present the current situation and problems of the terminology work and management in Slovakia. The core of the work focuses on the history and development of the Slovak Terminology Database Project, based partly on corpus linguistics. Besides, they draw attention to the importance of gathering and centralising terminology information and resources by means of a terminology portal. The last section of the paper is devoted to a comparison of the representation of terminology in the Slovak Terminology Database and the Bulgarian Terminological Electronic Archive.

Keywords: *Slovak terminology database, terminology work, terminology portal, multilingual terminology*

1. Background and contemporary work on terminology in Slovakia

In every scientific and professional domain, it is essential to use exact, comprehensible and unambiguous terminology, accepted and consistently used by domain experts. However, the reality is significantly different – both in theories of scientific disciplines and in scientific practice, variability of terms is an everyday phenomenon, not to mention synonymy and polysemy on a trans- and intradisciplinary level. Creation and standardisation of terms and terminologies is a part of the terminological and terminographical work, which on the one hand is specialised, and on the other hand has an interdisciplinary or transdisciplinary character – it requires knowledge of language and domain-specific knowledge, as well as knowledge of related domains. Systematic terminology work consists of multiple phases, beginning with a description of terms and concepts in use (sometimes requiring harmonisation of conceptual and, consequently, terminological systems), through coining new terms all the way to the unification, normalisation or standardisation of terminological units (Levická 2005).

In Slovakia, extensive terminology work had been in the 1950s and 1960s by efficient terminological committees, which resulted in publishing some tens of terminological dictionaries from various domains. Moreover, the leading figure in this process, prof. Ján Horecký (Horecký 1956), presented the theoretical guidelines of this terminology work in his thesis *Základy slovenskej terminológie* (Foundations of Slovak Terminology, 1956).

After the political, social and economic turn of 1989, Slovakia, together with other countries in Central and Eastern Europe, witnessed significant changes in vocabulary, especially in terminology (see also Šimková 1999). Another period of intense development resulted in ambiguities in the terminologies of law, economics and administration was during the accession process and integration

of the Slovak Republic into the European Union, which created an urgent need for translation of many documents in a short period of time. These processes, as well as the development of new scientific disciplines, called for a new period of terminology work and terminology unification.

Some institutions, aware of the urgent need, started with limited but key terminological activities, i.e. building terminology databases – the National Bank of Slovakia and the Slovak Office of Standards, Metrology and Testing, where a Department of Technical Standardisation was established in 2014. However, those are only domain-limited databases with specific and very narrowly defined aims – the former is an in-house tool for employees of the bank while the latter is used by the creators and translators of technical norms and has been offered to public recently on an annual fee basis. One should also mention the English-Slovak and Slovak-English specialised dictionary published by the Research Institute of Posts and Telecommunication in Banská Bystrica. The dictionary contains terminologies of post and electronic communication, updated and enriched annually (<http://www.vus.sk/iecd/new/Vyklad.asp>).

The need for coordination of terminological activity is regulated by the **State Language Policy of the Slovak Republic**, approved by the Slovak government in 2001 (<http://www.culture.gov.sk/posobnost-ministerstva/statny-jazyk/dokumenty-c3.html>). In this document, a terminological database built in order to enhance the protection and development of the Slovak language is recognized as one of the main aims: “**to build a database of terms** and to make it electronically accessible is in the interest of the whole scientific and professional community and thus of high priority. It is a difficult task that requires significant funding, human resources and cooperation between ministries. Building the database is a task that can be best and most professionally handled by The Linguistic Institute.”

Since the need for **centralisation** of existing terminological resources as well as the need for **coordination** of terminology work became more and more pressing, the Slovak National Corpus, a department of the Ľudovít Štúr Institute of Linguistics, Slovak Academy of Sciences (<http://korpus.juls.savba.sk>), started the preparation phase for the Slovak Terminology Database project (hereinafter referred to as STD) in 2005.

Two years later (2007) the EC’s Directorate-General initiative was started, aiming at unification of national terminological systems via terminological networking, discussion and projects. The initiative resulted in establishing the Slovak Terminology Network (*Slovenská terminologická sieť*; <http://www.slovenskaterminologickasiet.eu/>).

Based on the 2014 resolution of the Slovak government, terminological committees were re-established at all central institutions of the Slovak public administration.

2. Project of the Slovak Terminology Database (2005-2017)

2.1. Goals and development of the Slovak Terminology Database

The unification of terminology presupposes centralisation and coordination of terminology work, which means “to collect terminological documentation which already exists in various forms and is often difficult to obtain, to promote a methodical approach to terminological research in sectors as yet unexplored, and to make all information available to clients at reasonable cost through fast and efficient processing” (Dubuc 1972: 203). Currently, all these functions can be ensured by terminology databases. With this aim in mind, the project of the Slovak Terminology Database was started.

The preparation of the STD project methodology (2005-2007) necessitated a thorough analysis of the terminological activities and needs in Slovakia and worldwide in the last 20 years (i. e. approximately since the mid-1980s). The analysis also focused on the structure and functioning of leading foreign terminology databases (especially Eurodicautom, its successor IATE – InterAgency Terminology for Europe, as well as the Canadian databases Termium and Grand dictionnaire terminologique). Consequently, the methodology of the STD was elaborated, based partly on the

traditions of Slovak terminology theory, as formulated mostly in the works of J. Horecký and I. Masár (see e. g. Masár 1991), and partly on knowledge of current trends in terminological management and, more specifically, in text-based terminology and socioterminology.

Even the title of the project required actual terminology work, since there is some variety in the naming of the concept *terminologická databáza* (terminology database). This is, at the same time, a textbook example of a serious problem that the terminologist comes across when a term is eventually established in several fields and is not only used by various professionals, but by lay public as well. In Slovak, for instance, there are three terms used for the given concept (*terminologická banka*, *terminologická databanka*, *terminologická databáza*), which reflects the French usage (*banque de terminologie*, *banque de termes*, *base de données terminologiques*), whereas English has four terms (*terminological data bank*, *terminological database*, *terminology database*, *term bank*). This topic is described in greater detail in (Levická 2008).

Apart from other arguments, given that two major terminological resources – IATE and Termium – are called “databases”, we opted for the Slovak adapted term *databáza* (database), and the definition by J. C. Sager who claims that a database is “[a] collection, stored in a computer, of specialized language vocabularies, including nomenclatures, standardized terms and phrases, together with the information required for their identification, which can be used as a mono- or multilingual dictionary for direct consultation, as a basis for dictionary production, as a control instrument for consistency of usage and term creation and as an ancillary tool in information and documentation” (Sager 1990: 169). This definition emphasizes the multitude of the functions of the database, corresponding to the expected interests of various target groups of database users.

Main characteristics of the terminology database:

1. macrostructure, i.e. **logical order**, usually reflecting the conceptual system of individual fields;
2. microstructure, i. e. **content** of individual terminological records;
3. **method of** term record creation and **choice of resources** used for obtaining terminological information.

From the view point of general linguistics, the terminology database can be

- | | |
|---------------|---|
| prescriptive: | it enables unification of terminologies and, being accessible to the general public, it ensures and facilitates the spreading of already standardised terminologies; |
| descriptive: | <ol style="list-style-type: none">1. it offers the possibility of mapping the current state and development of terminologies, as well as updating, modification and adding of terminological information;2. it can provide outputs of different kinds, especially thematic glossaries and dictionaries. |

From the terminological point of view, the database

1. facilitates the **coordination** of terminological activities;
2. serves as a **referential source** of professional and linguistic information that can be used partly by professionals, partly by linguists, translators and, last but not least, by the general lay public.

Hence, the terminology database is neither a library nor an electronic archive of printed dictionaries, but a specific collection of terminological records with its own structure and content. In the event of enriching and enlarging the terminology database with terms from existing resources, it is necessary to unify, correct and complete these terms according to the principles of the terminology database.

The STD project aims to set up a terminology database provided with both conceptual and linguistic information, inspired by foreign examples but, of course, adapted to Slovak needs and the possibilities of the Ľ. Štúr Institute of Linguistics. The team is open to cooperation and exchange of data with the leading European database IATE, which is why the EUROVOC 4.3 Thesaurus was chosen as the classification system (see section 2.2).

The abovementioned project methodology drew inspiration from the textual terminology approach to the terminology extraction of lexical units – potential terms from specialised running texts and identification of concepts they refer to. Of no lesser importance is the development of software tools for automatic extraction of terms and possibly definitions from specialised subcorpora.

The centralisation of various terminologies, continuous modification and updating of term records and narrow collaboration of terminological boards, translators and specialists is nowadays considered to be the only way of terminological unification, and consequently standardisation. Contemporary terminological tendencies stress the model of the text and corpus approach as a *sine qua non* prerequisite of every terminological project. The process of systematic gathering of terms is based exclusively on representative corpora, supervised and validated by specialists and terminologists. As Sager (Sager 1990: 131) puts it, information extracted from a text represents a reliable indicator of changes and ensures the only plausible data for building and revising terminological records.

It is therefore only natural that a terminology database project be started by the Corpus Department of the Ľ. Štúr Institute of Linguistics for it had all the resources and tools at its disposal – the textual base of the Slovak National Corpus itself and software ones for automatic annotation of Slovak texts.

The team chose and adapted the appropriate software for the database along the way. The SNC policy of text acquisition also had to be modified and the focus was shifted towards economic and legal texts for the purpose of creating specialised subcorpora (http://korpus.sk/legal_en.html, http://korpus.sk/ecn_en.html).

The beta version of the STD, launched online in 2007, was focused on reusing and adapting existing quality terminology resources published in particular in the Slovak linguistic revue *Kultúra slova*. The testing period allowed verifying and evaluating the proposed methodology, which resulted in several subsequent modifications, and training team editors and proofreaders. Since the collection of terminological resources already existing in various forms is a part of the terminology unification and coordination of terminology work (Dubuc 1972: 203), it was necessary to formulate principles for harmonisation of the work with respect to the original concept. At the same time, these analyses helped to envisage problems and specificities regarding individual terminological projects, in the framework of which new terminological records have been and will be created.

Whether the STD is to include new records or already existing terminographical products, the goal is to provide the highest quality possible, continuously add data to the obligatory categories missing in the original resources, and to minimize the duplicity of records, which began to appear already in the beta version of the project and has since then, as the STD grows, been more and more significant.

Nowadays, the STD is integrated into the terminological portal (https://terminologickyportal.sk/wiki/Hlavná_stránka) and offers more than 6,500 term records that can be classified into domains corresponding to various soft and hard sciences: Astronomy, Bilingualism, Construction, Corpus Linguistics, Criminal Law, Criminology, Fire Protection, History, Labor Law, Migration Policy etc.

2.2. Structure and content of term records

The study of current trends in terminology management and the development of terminology databases has shown that for the needs of the Slovak professional and lay public, it is best to create a corpus-based database with the use of the EUROVOC Thesaurus (as of today, it is 4.4. version) as the classification system. As far as the term record design is concerned, the team drew inspiration especially from the ISO 10241 (ISO 10241 1992) International terminology standards – Preparation and layout. In compliance with the best practices of terminology work, the relevance of individual fields for the end user (terminologists, scientists, professionals, translators, teachers, laymen) had to be considered and an evaluation of project possibilities had to be made. Further consideration was given to the results of research conducted at professional forums and targeted at the needs of translators. The first version of a terminological record contained 12 categories, 7 – 9 being obligatory (depending on the category), whilst the edited version contained 11 categories, with 7 regarded as obligatory. The new updated version of the term record includes 13 data categories, 5 of which are obligatory: term, field, definition,

source of the definition and term usage. The optional categories are: synonyms, context, context source, related terms, authority, equivalents in foreign languages, note, URL. However, most of the existing records have only the obligatory part completed; less often they include synonyms, sometimes related terms, and a comment. Individual categories of the term record are described in the section Description of terminological record at https://terminologickyportal.sk/wiki/O_projekte.

absolútne

Odporúčaný termín

OBLASŤ:	FILOZOFIA
DEFINÍCIA:	to, čo je nezávislé, vo svojej existencii ničím iným nepodmienené
ZDROJ:	Slovník spoločenských vied. Bratislava : SPN 1997.
KONTEXT:	Absolútne je pojem idealistickej filozofie na označenie niečoho večného, nekonečného, nepodmieneného, dokonalého a nemeniteľného, ktoré nezávisí od ničoho iného, samo osebe obsahuje všetko jestvujúce a tvorí ho.
ZDROJ KONTEXTU:	PODĽA Filozofický slovník. Bratislava : Pravda 1982.
PRIBUZNÉ TERMÍNY:	idealizmus, relativizmus
CUDZAJAZYČNÝ EKIVALENT:	en: absolute, la: absolutus
POZNÁMKA:	Opakom absolútne je relatívne, ktoré je podmienené vzťahmi a súvislosťami s inými vecami.
URL:	http://ii.fmph.uniba.sk/~kravcik/filii/pojem.php?obl=Absolutno

Kategória: Filozofia

Figure 1: Example of a term record in the philosophy domain: *absolútne* (absolute)

Adding existing terminological resources to the STD requires, first of all, a qualitative evaluation of these resources and, consequently, classification of terminological information contained therein. The attention is given to the harmonisation of the form and content of individual term records. Items and phenomena to be removed or corrected are identified. A significant amount of time is dedicated to completion of obligatory categories that were not included in the original resources. The issue of terminological homonymy/polysemy and synonyms (Levická 2015) is also a topic of research.

Rules for editing terminological records (https://terminologickyportal.sk/wiki/O_projekte/Ako_editovat) are another instrument for harmonising the form and content of individual records in the STD. They comprise the description of every category, explain individual subcategories, and, in separate sections, one will find a detailed description of what the given record should not contain.

POUŽITIE TERMÍNU

Charakterizuje termín a naznačuje jeho pragmatický charakter, t. j. **informuje o schválení, prijateľnosti či použiteľnosti termínu v danom kontexte**. Pri výbere kvalifikátorov, ktoré možno použiť pri tvorbe záznamov STD, sa vychádzalo z noriem ISO 12620:1995 a ISO 740:2000.

Legislatívny	termín definuje zákon alebo vyhláška.
Odporúčaný	termín odporučila terminologická komisia príslušného odboru, prípadne iná relevantná inštitúcia.
Normalizovaný	termín definuje norma/norma.
Nesprávny	termín sa odporúča nepoužívať vzhľadom na jeho nekorektnú formu alebo obsah.
Neologizmus	novovytvorený, resp. novoprevzatý, ktorého forma a/alebo obsah sa ešte môže zmeniť.
Zastaraný	termín sa odporúča nepoužívať vzhľadom na jeho zastaranosť.
Profesionalizmus	termín používaný v rámci odborných slangov.
Eurolegislatívny	termín definuje európska legislatíva.

Figure 2: Example of the Term usage category description, concerning the pragmatic aspect of a term, its status on the scale legal – recommended – standardised – – incorrect.

3. Terminology Portal

In 2015, the Department of the Slovak National Corpus at the Ľudovít Štúr Institute of Linguistics launched the preparation of a terminology portal, whose pilot version was made public in November 2016 (<https://terminologickyportal.sk>). The building of the portal is first of all related to the project *Analysis of terminology work of Ján Horecký as an inspiration for the terminology management of the 21st century in Slovakia*. The aim of the portal is to gather and present in one place as much information about terminology and terminological issues in Slovakia and abroad as possible, as well as to make some terminological resources on-line accessible to all interested in terminology.

The portal currently provides 7 main sections with the following content:

A. Terminological resources

A.1. Slovak Terminology Database (see section 2)

A.2. Overview of mono- and multilingual Slovak terminological dictionaries – annotated, in some cases even commented bibliographic information about older or more recent dictionaries and encyclopedias. Every record contains information on the domain of the dictionary and languages involved.

The screenshot shows the Terminology Portal interface. On the left is a sidebar titled 'Prehľad výkladových a prekladových terminologických slovníkov' (Overview of explanatory and translation terminological dictionaries) containing a list of various dictionaries. The main content area on the right displays a detailed record for 'Slovenské mená lišajníkov (lichenizovaných húb)' (Slovak names of lichens). This record includes the authors 'Pišút, I. – Lackovičová, A. – Lisická, E.', a citation from the journal 'Kultúra slova', the domain 'mykológia' (mycology), and the languages 'slovenský' (Slovak) and 'latinský' (Latin). It also provides a comment and a link to the full text PDF.

Figure 3: Example of a record on the dictionary *Slovenské názvoslovie lišajníkov* (Slovak Nomenclature of Lichens)

A.3. Monolingual terminological dictionaries SAV (1953 – 1965) – e-versions of terminological dictionaries created and/or published by the Slovak Academy of Sciences

B. Theoretical resources

B.1. *Slovenské odborné názvoslovie* (Slovak Professional Nomenclature) – all issues of the journal published in 1953 – 1963, pdf files

The screenshot shows the 'Slovenské odborné názvoslovie' (Slovak Professional Nomenclature) journal interface. On the left is a sidebar with a list of journal issues from 1953 to 1954. The main content area on the right displays the table of contents for issue 10, 1953. The table lists various articles and their authors, along with their page numbers. The articles include 'Ferdinand Buffa: Sto rokov od vydania Reusovej Květny Slovenska', 'Z terminológie gymnastiky', 'Terminológia hŕti', 'Plachtárske názvoslovie', 'Rusko-slovenská terminológia stolného riadu a príborov', 'DISKUSIE', 'Peter Popesko: K novej veterinárskej anatomickej terminológii', 'ZPRÁVY', 'V. Dužčíková: MVDr. Pavol Laurenčík, inseminácia hovädzieho dobyčka', 'J. Horecký: Gabriel Čeněk—Václav Medek, Kurz deskriptivnej geometrie pre technikov I.', 'V. Dužčíková: Pekáreň—Cacek—Krátký—Kvičera, Rusko-český technický slovník', 'F. Buffa: Stavebný materiál', and 'J. Horecký: Sportszótár. Sportívny slovar. Dictionnaire sportif. Sportwörterbuch. Sports Dictionary'.

Figure 4: A sample from the journal *Slovenské odborné názvoslovie*, including contents

B.2. Československý *terminologický* časopis (Czechoslovak Journal for Terminology) – all issues published in 1962-1966, pdf files

B.3. Database of terminological works by J. Horecký – annotated bibliographical information on all of J. Horecký's published terminological works; e-versions of the texts are being added

O našej terminologickej práci

[Prehľad PDF](#)

Jazykovedný zborník Slovenskej akadémie vied, 4, 1950, s. 60 – 62.

Oblasť: **Ingenierstvo**

Kľúčové slová: [Terminológia](#) [Terminologický jazyk](#) [Terminologické práce](#) [Terminologická práca](#) [Terminologické práce](#) [Terminologické práce](#)

Notícia:
 Autor v príspevku nastojuje potrebu kritiky vo vzťahu k teórii pravej fonologickéj, resp. štrukturalistickej školy. Upozorňuje, že v prvom období po uzložení slovenčiny sa pri tvorbe terminov tradične siahlo k ľudovému jazyku, v dôsledku čoho sa terminologická práca vyvíjala napríklad chémií, strojárstvom, stavebníctvom. Podľa Horeckého by súčasná terminológia nemala slovenskú terminológiu vraciat do obrodeneckého obdobia a jeho súhrnného tvorenia neologizmov. Výhodnejšie je podľa Horeckého prevziať ako termín medzinárodný, ale v danej dobe žilé slovo jeho príklad sa uvidia ekonomizér namiesto užitočnosť/hospodárnik, dľa namiesto hubice/trysky/fúkačky. V závere vyzdvihuje potrebu neustálnej konfrontácie terminologickej teórie s praxou tvorenia a ušatovania konkrétnych terminov.

Figure 5: Example of an annotation of J. Horecký's paper on terminology work

C. Useful links – annotated links to Slovak and foreign terminology databases, dictionaries, portals, networks and centers

Both the Terminology Portal and the Slovak Terminology Database contained therein serve as a dynamic source of information about terminology work and its results, which are added and updated on regular basis.

4. Multilingual Terminology Representation

Multilingual electronic terminological databases are reliable sources of correct translations of terms since they present structured conceptual knowledge of a certain domain and allow electronic multilingual search with various criteria. In this part, we will compare representations and definitions of terms in the domain of *linguistics* given in the STD and Bulgarian Terminological Electronic Archive (BTEA). The BTEA is the electronic archive of the Department of Bulgarian Terminology and Terminography of the Institute for Bulgarian Language containing terms in various subject domains including several electronic terminological dictionaries in Bulgarian language. For the purpose of our work, we shall compare related terms in the domain of *corpus linguistics* from the STD repository and from an electronic version of the Dictionary of Terms in the Humanities (Попова / Popova 2007) of the BTEA repository, and we shall analyse the term *linguistics*. The definition from BTEA is as follows:

ЕЗИКОЗНАНИЕ, само ед.; **лингвистика**. Ез. Наука за езика, която установява, класифицира и систематизира конкретните езикови факти, създава и преценява теоретико-методологическите основи на знанието в тази научна област, разкрива неговите закономерности, изучава възникването и развитието на езика.

СЧ: вж. *общо езикознание* и *приложно езикознание*; *сравнително-историческо езикознание*; *математическо езикознание*; *фонетика* и *фонология*, *граматика* (*морфология* и *синтаксис*), *морфемология*, *словообразуване*, *лексикология*, *етимология*, *стилистика*, *текстология*, *етнолингвистика*, *социolingвистика*, *психolingвистика*, *диалектология*, *топонимия*, *лингвостатистика* и др.

В: вж. диахронно езикознание и синхронно езикознание. (БП)

ЛИНГВИСТИКА, лингвистики. Ез. Вж. езикознание. (БП)

The presented examples show that the Bulgarian language uses two terms for *linguistics* – *езикознание* which is of domestic origin and *лингвистика* which is of foreign origin. The domestic term is presented with its definition and also with related hyperonyms classified by their semantic relations, and the foreign term is given as a synonym. However, the entry for the foreign term is referred to the definition of the domestic term.

The screenshot shows the STD website interface. On the left is a sidebar with navigation links: Hlavná stránka, Zoznam terminologických zbierok, O projekte STD, Často kladené otázky, Kontakt, and a section for nástroje (tools) including odkazy na túto stránku, súvisiace úpravy, špeciálne stránky, verzia na tlač, and trvalý odkaz. The main content area displays the term 'lingvistika'. It includes a search bar at the top right with the text 'Hľadať STD' and a search icon. Below the search bar, there are tabs for 'Termín', 'Diskusia', and 'Zobrazíť formulár'. The term 'lingvistika' is highlighted. Under the heading 'Odporúčany termín', there is a table with the following information: OBLAST: LINGVISTIKA; DEFINÍCIA: veda o jazyku, jeho systéme a používaní, vývine, vzťahoch s mimojazykovými skutočnosťami a pod.; ZDROJ: Šimková, M.: Výberový slovník termínov z korpusovej lingvistiky. 2006.; SYNONYMUM: jazykoveda; CUDZOJAZYČNÝ EKVALENT: en: linguistics, bg: лингвистика, bg: езикознание; URL: http://korpus.juls.savba.sk/; and a category 'Kategória: Korpusová lingvistika'.

Figure 6: Example of a term record for the term *lingvistika* (linguistics)

The STD includes the term *linguistics* in its domain of *corpus linguistics* and uses the word *lingvistika* for it. The structure of the database allows adding equivalents in foreign languages, and the entry of the Slovak term *lingvistika* includes also two equivalents in English – *linguistics* and in Bulgarian – *езикознание* and *лингвистика* given as synonyms. The entry is as follows:

Also, the related compound terms are presented in both electronic databases. The Slovak and Bulgarian terms for *mathematical linguistics* are defined as hyponyms of term *linguistics* – *matematická lingvistika* and *математическо езикознание*, respectively. However, the Bulgarian database also defines the terms *математическа лингвистика*¹ and *математическа лингвистика*² as follows:

МАТЕМАТИЧЕСКА ЛИНГВИСТИКА¹. Математическа дисциплина с твърде широко съдържание, което включва: създаване на математически модели на езика чрез апарата на алгебрата, теорията на алгоритмите и математическата логика; статистика и статистически методи; автоматичен превод от един естествен език на друг естествен език; приложни изследвания. (МП)

МАТЕМАТИЧЕСКА ЛИНГВИСТИКА². Ез. Вж. математическо езикознание. (БП)

МАТЕМАТИЧЕСКО ЕЗИКОЗНАНИЕ; математическа лингвистика². Ез. 1. Дял от езикознанието, който изучава езика чрез математически методи.

2. Използването само на такива математически методи, които описват фактите в естествените езици и водят до резултати, съществени или полезни за лингвистиката.

From the software point of view, the STD is open to modifications and add-on. For example, a unilingual term record can be extended to a bilingual one comprising not only a foreign equivalent but also a definition or other categories in a given foreign language.

However, the key feature of any terminological database or electronic terminographical product is the constant updating of their content - both the STD and the BTEA are to prove their flexibility in the near future.



Slovenská terminologická databáza

[Vytvoríť účet](#) [Prihlásiť sa](#)

Termin
Diskusia
Zobrazíť formulár
Čítať

Hľadať STD 🔍

dezambiguácia ¹

Odporúčaná termin

OBLASŤ:	LINGVISTIKA
DEFINÍCIA:	výsledok automatizovaného odstraňovania viacznačnosti, homonymie tvaru a priradovanie jednoznačných kategórií pri počítačovom spracovaní textu
ZDROJ:	Šimková, M.: Výberový slovník termínov z korpusovej lingvistiky. 2006.

CUDZOJAZYČNÝ	en: disambiguation, bg: снемане на многозначност
EKVIVALENT:	
URL:	http://korpus.juls.savba.sk/

[Skryť]

Cudzojazyčné údaje

CUDZOJAZYČNÁ	bg: резултат от автоматичен процес, с който се елиминират омонимични значения
DEFINÍCIA	на дадена дума чрез анализ на таговете, присъединени към тази дума

Kategória: Korpusová lingvistika

Figure 7: Example of a term record for the term *dezambiguácia* (disambiguation)

References

- Dubuc 1972: Dubuc, R. Termium: System description. – *Meta*, 17, 4, 203 – 219.
- Horecký 1956: Horecký, J. *Základy slovenskej terminológie*. Bratislava: Vydavateľstvo SAV.
- ISO 10241 (1992): *Normes terminologiques internationales*. International Organization for Standardization.
- Levická 2005: Levická, J. Proces ustáľovania terminologických sústav a prínos terminologických databáz. – In: *Vedecká rozprava k problémom terminológie obchodu, obchodných vied a marketingu obchodu*. Ed. K. Viestová. Bratislava: Ekonomická univerzita, 22 – 29.
- Levická 2008: Levická, J. Slovenská terminologická databáza. – *Kultúra slova*, 42, 3, 139 – 157.
- Levická 2015: Levická, J. Synonymá v terminológii marketingu – sú anglicizmy preferovanejšie? – In: *Terminologické fórum V. Úklady terminológie*. Ed. Ingrid Cíliková. Trnava: Univerzita sv. Cyrila a Metoda v Trnave, Filozofická fakulta, 92 – 107.
- Masár 1991: Masár, I. *Príručka slovenskej terminológie*. Bratislava: VEDA.
- Sager 1990: Sager, J. C. *A Practical Course in Terminology Processing*. Amsterdam – Philadelphia: John Benjamins.
- Šimková 1999: Šimková, M. Pohyby v slovenskej lexike v 90. rokoch. – In: *Internacionalizácia v súčasných slovanských jazykoch: za a proti*. Red. J. Bosák. Bratislava: Veda, 116 – 136.
- Попова 2007: Попова, М. и др. *Терминологичен речник по хуманитарни науки*. София: Наука и изкуство. [Popova 2007: Popova, M. et al. *Terminologichen rechnik po humanitarni nauki*. Sofia: Nauka i iskustvo.]