**LICENSE AGREEMENT**

**concluded according to § 65 of Act no. 185/2015 Coll.**

**(Copyright Act) as amended**

**about other uses of works**

**I. Contracting Parties**

**Provider:**

Title:

Seat:

ID:

Represented by:

Person authorized to negotiate:

(hereinafter "Provider")

**Acquirer:**

Title: Institute of Linguistics Ľ. Štúra Slovak Academy of Sciences

Registered seat: Panská 26, 811 01 Bratislava

ID: 16 70 88

Represented by:

Person authorized to negotiate:

phone:

e-mail:

(hereinafter "Acquirer")

**II. Subject of the contract**

1. The Provider provides the Acquirer with the following works in electronic form:

*here is a list of works, or. the text "works as listed in the Annex"*

1. The Purchaser shall use the said works to the extent and under the conditions agreed in this contract.

**III. Method and scope of use of works**

1. The Provider grants the Licensee a license to an unlimited territorial extent.
2. The Purchaser is entitled to include the provided works in an automated unit called the Slovak National Corpus, in its respective subcorpora and parallel corpora . The corpus is a set of plain text data in a single form equipped with programs that allow the end user to search for and display individual words, word combinations and other language information in a predetermined context ( max. 50 units to the left and 50 units to the right of the search term ). The end user has no direct possibility to dispose in any way with the full texts of the works provided.

More detailed information on the Slovak National Corpus can be found in the Appendix, which is an integral part of this contract.

1. The Purchaser is entitled to enable third parties to study language phenomena in the Slovak National Corpus via the computer network Internet, to make available statistical data on the functioning of individual means in Slovak and frequency lists in the range of individual words, pairs, triplets and quadruples. Relevant contexts can also be used in the Slovak terminology database and for training language models .

**IV. Time of use of parts**

1. The works included in the corpus will be used indefinitely due to the specifics of the research linguistic work.

**V. Remuneration**

1. Due to the non-commercial use of the works, t. j. for scientific, research and teaching purposes , the Provider and the Purchaser have agreed to provide a license free of charge .

**VI. Obligations of the contracting parties**

1. The Purchaser undertakes to carry out the works provided in the Slovak National Corpus at its own expense for sorting the provided works.
2. The Purchaser undertakes that the works in question, as part of the Slovak National Corpus, will be used exclusively for scientific, research and teaching purposes and will not be used commercially.
3. The Purchaser undertakes to secure the Slovak National Corpus by all available means so that end users will not be able to dispose of the entire texts of individual works, resp. use the full texts of individual works as one coherent whole.
4. The Purchaser undertakes to ensure that the right of access to the provided works processed within the Slovak National Corpus has only those persons who have agreed to the non-commercial use of the Corpus. If the end user does not comply with the specified conditions, his access to the corpus will be immediately and irrevocably revoked.
5. Transferee agrees to allow Provider continuous access to sources of national Slovak him the body in the terms binding for each user Slovak National Corpus.
6. The Provider undertakes to use the resources of the Slovak National Corpus in accordance with this Agreement exclusively for non-commercial purposes .
7. In the case of published scientific di ela, in which were included in the reasonable scope of the text of the relevant part (generally one sentence) by scientific works indicate the source of the quote, the name of the author, Mr. Rip. copyright holder and other bibliographic data. There is no obligation to provide royalties for such use in accordance with § 44 of Act no. 185/2015 Coll. (Copyright Act) as amended.
8. The Provider declares that the use of the works agreed in this contract is not an infringement of the rights of the author, another holder of copyright or distribution rights or third parties.

**VII. Final provisions**

1. This Agreement shall enter into force and effect on the date of signature by both parties.
2. This contract is made in two copies, of which each of the contracting parties will receive one copy.
3. The contract may be terminated or amended only by mutual agreement and a written amendment signed by both parties.
4. In the case of a serious breach of any of the contractual obligations or in the event of loss of ability to fulfill the obligations arising from this contract s one of the parties has the opposite party is entitled to terminate the contract with two months' notice with effect from the first day of the calendar month follow increased anti- the month in which the notice was delivered to the other party. Any termination of the contract will only affect the provision of other texts, it will not apply in any way to the already provided and processed data included in the existing sources of the Slovak National Corpus.
5. In the event of a change in the status and objectives of the Purchaser in relation to the provided text, a new contract must be concluded.
6. This contract is governed by the law of the Slovak Republic. Any dispute in connection with this contract shall be settled by agreement between the parties as a matter of priority.
7. The contracting parties declare that they have concluded the contract freely, seriously and clearly, that the contract was not concluded in distress and under noticeably unfavorable conditions, that they became acquainted with its content, understood it and agreed to it, which they confirm by signing i.

In Bratislava

Provider Acquirer

                                                                                                 Director of JÚĽŠ SAS

**Appendix**

**Slovak National Corpus**was established in r. 2002 as a specialized department of the Institute of Linguistics Ľ. Štúra SAS in Bratislava aims to build and manage a corpus (s) of Slovak language texts.

The text corpus is an ***extensive internally structured*and *comprehensive set of language data, which are electronically stored and processed***, which allows to comprehensively and repeatedly examine them, on the basis of it to describe the system of the language and the real functioning of language means , which also serves the results frequency word lists and their combinations and other statistical analyzes. The use of corpus databases presupposes searching for language information on the Internet .

Corpora have different content and formal composition and focus.

***The primary database of the***Slovak National Corpus is being built as a ***general monolingual corpus***, in which *written*texts of contemporary Slovak from r. 1 955. The whole corpus is *lematized*using specialized software (each word is assigned a basic form) and *morphologically annotated*(each word is assigned information about its morphological properties in a given context). *Bibliographic and stylistic data are*manually assigned to each text . The selected sample of texts is manually annotated syntactically (each word is assigned information about its function in the sentence). In the corpus, it is possible to search online on the Internet for linguistic information for scientific research, teaching and other non-commercial uses, while the user has access only to short contexts containing the searched word, connection or specific forms. The specific components of the Slovak National Corpus are: manually morphologically annotated corpus, morphological databases and (dictionary of all word forms), Slovak dependency corpus (syntactically annotated corpus), parallel corpora, Slovak terminological database, Slovak spoken corpus, dialect and historical corpus, linguistic resources digitized lexicographic and grammatical works, monographs, collections, magazines, etc. from the current production of JÚĽŠ SAS, selectly also older works).     

The Slovak National Corpus is built primarily for the needs of lexicographers (creation of various dictionaries, currently especially the Interpretive *Dictionary of the Contemporary Slovak Language*), for the needs of grammar and stylistic research (grammar manuals, spelling rules; comparative research). It is also used in schools (creation of exercises; teaching Slovak to foreigners). A special area is computer processing of Slovak as a natural language.

The systematic building of the Slovak National Corpus primarily represents the systematic collection of data of various styles, genres, authorial, generational or publishing names from all regions of Slovakia. Vocabulary processing of contemporary Slovak since 1955 requires a lot of manual work in scanning and recognizing texts that do not exist in electronic form. Although the latest texts are usually directly in electronic form, they need to be converted into a uniform format and technically cleaned (graphics, images, tables, etc. are removed ) into a clean continuous text. The text is then segmented into words, text units to which linguistic information is added.

Procedure for processing texts in the Slovak National Corpus JÚĽŠ SAS :

● in the ARCHIVE, to which only directly involved members of the SNK department have access, the obtained texts are kept in the original format with basic information about their origin / source, form and content;

● after removing the characters and symbols of the editors and programs in which the texts were created, and after removing the graphic components, the texts are converted into a uniform format in which the structural properties of the text are recorded; an external annotation (bibliographic and stylistic-genre data) is added to each text - this form of the corpus forms a BANK, which is accessible exclusively to employees of the SNK JÚĽŠ SAS department for possible full-text search;

● in the next phase, the text divided into basic units (words, punctuation, *marks*, symbols) is linguistically *marked*: information on the text structure is added to the text, language information at the word level (word type, morphological categories of word form, basic word form - *lemma*), language information at the sentence level (sentence function, semantics ), etc. - create a CORPUSOID, from which they can obtain linguistic information for scientific research use workers JÚĽŠ SAS and external collaborators in the areas JÚĽŠ SAS;

● thus processed texts that are providers of the license for public areas JNE use are as data made available Nené the online search language means interested in Slovak language and researchers in the field of corpus linguistics who agree to the terms of a non-commercial resource use Slovak National Corpus pr via the Internet; the searched language information is displayed only in a precisely defined short context with information about the source and authorship of the text.

**You can find more information at https://korpus.juls.savba.sk.**

1 / 4