

L I C E N Ć N Á Z M L U V A

uzavretá podľa § 65 zákona č. 185/2015 Z. z.
(autorský zákon) v znení neskorších predpisov

I. Zmluvné strany

Poskytovateľ:

Názov:

Sídlo:

IČO:

Zastúpený:

Osoba oprávnená rokovať:

(ďalej „Poskytovateľ“)

Nadobúdateľ:

Názov: Jazykovedný ústav Ľ. Štúra Slovenskej akadémie vied, v. v. i.

Sídlo: Panská 26, 811 01 Bratislava

IČO: 16 70 88

Zastúpený: doc. Mgr. Gabriela Múcsková, PhD.

Osoba oprávnená rokovať: Mgr. Jana Levická, PhD.

tel.: 02/5441 0307

e-mail: jana.levicka@korpus.juls.savba.sk

(ďalej „Nadobúdateľ“)

II. Predmet zmluvy

1. Poskytovateľ poskytuje Nadobúdateľovi tieto diela v elektronickej podobe:
tu sa uvedie zoznam diel, príp. text „diela podľa zoznamu uvedeného v prílohe“.
2. Nadobúdateľ použije uvedené diela v rozsahu a za podmienok dohodnutých v tejto zmluve.

III. Spôsob a rozsah použitia diel

1. Poskytovateľ udeľuje Nadobúdateľovi licenciu v územne neobmedzenom rozsahu.
2. Nadobúdateľ je oprávnený zaradiť poskytnuté diela do automatizovaného celku nazývaného Slovenský národný korpus, do jeho príslušných podkorpusev a paralelných korpusev. Korpus predstavuje súbor čistých textových dát v jednotnom formáte vybavený programami, ktoré umožňujú konečnému používateľovi vyhľadať a zobrazíť jednotlivé slová, spojenia slov a iné jazykové informácie vo vopred stanovenom kontexte (max. 50 jednotiek vľavo a 50 jednotiek vpravo od hľadaného výrazu). Konečný používateľ nemá žiadnu priamu možnosť disponovať akýmkoľvek spôsobom s celými textami poskytnutých diel.

Podrobnejšia informácia o Slovenskom národnom korpuse sa nachádza v Prílohe, ktorá je neoddeliteľnou súčasťou tejto zmluvy.

3. Nadobúdateľ je oprávnený umožniť tretím osobám skúmanie jazykových javov v Slovenskom národnom korpuse prostredníctvom počítačovej siete internet, sprístupňovať aj štatistické údaje o fungovaní jednotlivých prostriedkov v slovenčine a frekvenčné zoznamy v rozsahu jednotlivých slov, dvojíc, trojíc a štvoríc slov. Relevantné kontexty môžu byť využité aj v Slovenskej terminologickej databáze a na tréning jazykových modelov.

IV. Čas použitia diel

1. Diela zaradené do korpusu sa budú používať vzhľadom na špecifiká výskumnej jazykovednej práce neurčitý čas.

V. Odmena

1. Vzhľadom na nekomerčné použitie diel, t. j. na vedecké, výskumné a vyučovacie ciele, sa Poskytovateľ a Nadobúdateľ dohodli na bezplatnom poskytnutí licencie.

VI. Závazky zmluvných strán

1. Nadobúdateľ sa zaväzuje, že zaradenie poskytnutých diel do Slovenského národného korpusu vykoná na svoje náklady.
2. Nadobúdateľ sa zaväzuje, že predmetné diela ako súčasť Slovenského národného korpusu sa budú používať výlučne na vedecké, výskumné a vyučovacie ciele a nebudú sa používať komerčne.
3. Nadobúdateľ sa zaväzuje všetkými dostupnými prostriedkami zabezpečiť Slovenský národný korpus tak, že koneční používatelia nebudú môcť disponovať celými textami jednotlivých diel, resp. používať úplné texty jednotlivých diel ako jeden súvislý celok.
4. Nadobúdateľ sa zaväzuje zabezpečiť, aby právo prístupu k poskytnutým dielam spracovaným v rámci Slovenského národného korpusu mali len tie osoby, ktoré súhlasili s nekomerčným používaním korpusu. V prípade, že konečný používateľ nedodrží stanovené podmienky, bude mu prístup ku korpusu okamžite a nevratne zrušený.
5. Nadobúdateľ sa zaväzuje umožniť Poskytovateľovi plynulý prístup k zdrojom Slovenského národného korpusu v rámci podmienok záväzných pre každého používateľa Slovenského národného korpusu.
6. Poskytovateľ sa zaväzuje používať zdroje Slovenského národného korpusu v súlade s touto zmluvou výlučne iba na nekomerčné ciele.
7. V prípade publikovania vedeckého diela, do ktorého boli zaradené v odôvodnenom rozsahu časti textu predmetných diel (spravidla jedna veta), autor vedeckého diela uvedie prameň citátu, meno jeho autora, príp. nositeľa autorských práv a ďalšie bibliografické údaje. Za takéto použitie nevzniká povinnosť poskytnúť autorskú odmenu v súlade s § 44 zákona č. 185/2015 Z. z. (autorský zákon) v znení neskorších predpisov.

8. Poskytovateľ vyhlasuje, že použitia diel dohodnuté touto zmluvou nie sú porušením práv autora, iného nositeľa autorských alebo distribučných práv ani tretích osôb.

VII. Záverečné ustanovenia

1. Táto zmluva nadobúda platnosť a účinnosť dňom podpísania oboma zmluvnými stranami.
2. Táto zmluva sa vyhotovuje v dvoch rovnopisoch, z ktorých každá zo zmluvných strán dostane jedno vyhotovenie.
3. Zmluvu je možné ukončiť alebo zmeniť iba na základe vzájomnej dohody a obidvoma stranami podpísaného písomného dodatku.
4. V prípade závažného porušenia niektorej zo zmluvných povinností alebo v prípade straty schopnosti plniť povinnosti vyplývajúce z tejto zmluvy jednou zo zmluvných strán má opačná strana nárok na vypovedanie zmluvy s dvojmesačnou výpovednou lehotou, ktorá začína plynúť prvým dňom kalendárneho mesiaca nasledujúceho po mesiaci, v ktorom bola výpoveď doručená druhej zmluvnej strane. Prípadné vypovedanie zmluvy bude mať dosah iba na poskytovanie ďalších textov, nebude sa nijako vzťahovať na už poskytnuté a spracované dáta zaradené do existujúcich zdrojov Slovenského národného korpusu.
5. V prípade zmeny štatútu a cieľov Nadobúdateľa vo vzťahu k poskytnutým textom musí byť uzavretá nová zmluva.
6. Táto zmluva sa riadi právnym poriadkom Slovenskej republiky. Akýkoľvek spor v súvislosti s touto zmluvou budú zmluvné strany riešiť prednostne dohodou.
7. Zmluvné strany vyhlasujú, že zmluvu uzatvorili slobodne, vážne a zrozumiteľne, že zmluva nebola uzatvorená v tiesni a za nápadne nevýhodných podmienok, že sa s jej obsahom oboznámili, porozumeli mu a súhlasia s ním, čo potvrdzujú svojimi podpismi.

V Bratislave

Poskytovateľ

Nadobúdateľ

doc. Mgr. Gabriela Múcsková, PhD.
riaditeľka JÚLŠ SAV, v. v. i.

Príloha k Licenčnej zmluve

Slovenský národný korpus vznikol v r. 2002 ako špecializované oddelenie Jazykovedného ústavu Ľ. Štúra SAV v Bratislave s cieľom budovať a spravovať korpus(y) textov slovenského jazyka.

Textový korpus je **rozsiahly vnútorne štruktúrovaný a ucelený súbor jazykových dát, ktoré sú elektronicky uložené a spracovávané**, čo umožňuje všestranne a opakovane ich skúmať, na základe výskumu opisovať systém príslušného jazyka a reálne fungovanie jazykových prostriedkov, na čo slúžia aj výsledky spracovania textov v podobe frekvenčných zoznamov slov a ich spojení a iné štatistické analýzy. Používanie korpusových databáz predpokladá vyhľadávanie jazykových informácií na internete.

Korpusy majú rôzne obsahovo-formálne zloženie a zameranie.

Primárna databáza Slovenského národného korpusu sa buduje ako **všeobecný jednojazyčný korpus**, v ktorom sa spracúvajú *písané* texty súčasnej slovenčiny od r. 1955. Celý korpus je pomocou špecializovaných softvérov *lematizovaný* (každému slovu sa priradí základný tvar) a *morfológicky anotovaný* (každému slovu sa priradí informácia o jeho morfológických vlastnostiach v danom kontexte). Ručne sa ku každému textu priradujú *bibliografické a štýlovo-žánrové údaje*. Vybraná vzorka textov je ručne anotovaná syntakticky (každému slovu sa priradí informácia o jeho funkcii vo vete). V korpuse je možné vyhľadávať on-line na internete jazykové informácie na vedecko-výskumné, učebné a iné nekomerčné použitia, pričom používateľ má prístup iba ku krátkym kontextom obsahujúcim hľadané slovo, spojenie či konkrétne tvary. Špecifickými súčasťami Slovenského národného korpusu sú: ručne morfológicky anotovaný korpus, morfológická databáza (slovník všetkých tvarov slov), Slovenský závislostný korpus (syntakticky anotovaný korpus), paralelné korpusy, Slovenská terminologická databáza, Slovenský hovorený korpus, nárečový a historický korpus, lingvistické zdroje (digitalizované lexikografické a gramatické diela, monografie, zborníky, časopisy a pod. zo súčasnej produkcie JÚĽŠ SAV, výberovo aj staršie diela).

Slovenský národný korpus sa buduje predovšetkým pre potreby lexikografov (tvorba rôznych slovníkov, aktuálne najmä výkladového *Slovníka súčasného slovenského jazyka*), pre potreby gramatického a štýlistického výskumu (gramatické príručky, pravopisné pravidlá; porovnávacie výskumy). Svoje využitie nachádza aj v školách (tvorba cvičení; výučba slovenčiny pre cudzincov). Osobitnou oblasťou je počítačové spracovanie slovenčiny ako prirodzeného jazyka.

Systematické budovanie Slovenského národného korpusu predstavuje predovšetkým systematický zber dát najrôznejších štýlov, žánrov, autorských, generačných či vydateľských úzov zo všetkých regiónov Slovenska. Spracúvanie slovnej zásoby súčasnej slovenčiny od r. 1955 vyžaduje veľa ručnej práce pri skenovaní a rozpoznávaní textov, ktoré neexistujú v elektronickej podobe. Najnovšie texty síce bývajú zväčša priamo v elektronickej podobe, no treba ich konvertovať do jednotného formátu a technicky čistiť (odstraňuje sa grafika, obrázky, tabuľky a pod.) do podoby čistého súvislého textu. Text sa potom rozsegmentuje na slová, textové jednotky, ku ktorým sa pridávajú lingvistické informácie.

Postup pri spracúvaní textov v Slovenskom národnom korpuse JÚĽŠ SAV:

- v ARCHÍVE, ku ktorému majú prístup iba priamo zainteresovaní členovia oddelenia SNK, sa získané texty uchovávali v pôvodnom formáte so základnou informáciou o ich pôvode/zdroji, forme a obsahu;
- po odstránení znakov a symbolov editorov a programov, v ktorých texty vznikli, a po odstránení grafických súčastí sa texty prevedú do jednotného formátu, v ktorom sa zaznamenávajú štruktúrne vlastnosti textu; ku každému textu sa doplní vonkajšia anotácia (bibliografické a štýlovo-žánrové údaje) – táto podoba korpusu tvorí BANKU, ktorá je prístupná výlučne pracovníkom oddelenia SNK JÚĽŠ SAV na prípadné fulltextové vyhľadávanie;
- v ďalšej fáze sa text rozdelí na základné jednotky (slová, interpunkcia, značky, symboly) lingvisticky *značkuje*: textu sa pridávajú informácie o štruktúre textu, jazykové informácie na úrovni slov (slovný druh, morfológické kategórie tvaru slova, základný tvar slova – *lema*), jazykové informácie na úrovni viet (funkcia vo vete, sémantika) a pod. – vytvorí sa CORPUSOID, z ktorého môžu získať lingvistické informácie na vedecko-výskumné využitie pracovníci JÚĽŠ SAV a externí spolupracovníci v priestoroch JÚĽŠ SAV;
- takto spracované texty, ktoré majú od poskytovateľov textov licenciu na verejné využívanie, sú ako DÁTA sprístupnené na on-line vyhľadávanie jazykových prostriedkov záujemcom o slovenský jazyk a bádateľom v oblasti korpusovej lingvistiky, ktorí súhlasia s podmienkami nekomerčného používania zdrojov Slovenského národného korpusu prostredníctvom siete internet; vyhľadané jazykové informácie sa zobrazujú len v presne vymedzenom krátkom kontexte s informáciou o zdroji a autorstve textu.

Ďalšie informácie nájdete na <https://korpus.juls.savba.sk>.