

Nataša Logar, Iztok Kosem, Tomaž Erjavec

ALEKS:

**Leksikalno-skladenjska podatkovna zbirka slovenskega strokovno-
znanstvenega jezika
(Zasnova in zgradba)**

Ljubljana, 2019

ALEKS: Leksikalno-skladenjska podatkovna zbirka slovenskega strokovno-znanstvenega jezika

Verzija: v1.0 (25. maj 2019)

Avtorji: Nataša Logar, Iztok Kosem, Tomaž Erjavec

Sodelavca: Petra Zaranšek, Cyprian Adam Laskowski

Izdajatelj: Fakulteta za družbene vede; Center za jezikovne vire in tehnologije; Institut "Jožef Stefan", 2019

Založnik: Amebis, d. o. o., Kamnik

Javni financer: Javna agencija za raziskovalno dejavnost Republike Slovenije

Način dostopa (URL): www.termania.net.

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

811.163.6'373.46(0.034.2)

LOGAR, Nataša, 1976-

Aleks [Elektronski vir] : leksikalno-skladenjska podatkovna zbirka slovenskega strokovno-znanstvenega jezika / [avtorji Nataša Logar, Iztok Kosem, Tomaž Erjavec ; sodelavca Petra Zaranšek, Cyprian Adam Laskowski ; izdajatelj Fakulteta za družbene vede, Center za jezikovne vire in tehnologijo [in] Institut "Jožef Stefan"]. - V. 1.0 (25. maj 2019). - El. knjiga. - Kamnik : Amebis, 2019. - (Zbirka Termania)

Način dostopa (URL): <https://www.termania.net/slovarji/223/kas-slovar-splosnostrokovne-leksike>

ISBN 978-961-6474-39-9

1. Gl. stv. nasl. 2. Kosem, Iztok 3. Erjavec, Tomaž, 1960-

COBISS.SI-ID 30093798



Vsebina dela »ALEKS: Leksikalno-skladenjska podatkovna zbirka slovenskega strokovno-znanstvenega jezika« je dostopna pod pogoji licence Creative Commons [Attribution 4.0 International](https://creativecommons.org/licenses/by-nc/4.0/).

Delo »ALEKS: Leksikalno-skladenjska podatkovna zbirka slovenskega strokovno-znanstvenega jezika« je nastalo pri projektih *Slovenska znanstvena besedila: viri in opis* (J6-7094; 2016–2018) in *Kolokacije kot temelj jezikovnega opisa: semantični in časovni vidiki* (J6-8255; 2018–2020). Oba projekta je sofinancirala Javna agencija za raziskovalno dejavnost Republike Slovenije iz državnega proračuna.

Več o podatkovni zbirki ALEKS je mogoče prebrati na spletni strani projekta *Slovenska znanstvena besedila: viri in opis*.

KAZALO

I. Predgovor	4
II. Priprava zbirke.....	5
III. Mikrostruktura	7
IV. Dostop in vizualizacija	8
V. Krajšave in simboli.....	9
Literatura.....	10
Priloga.....	12

I. Predgovor

Večina podatkovnih zbirk (v ožjem smislu tudi slovarjev), povezanih s katerimkoli strokovno-znanstvenim – oz. v izrazoslovju žanrske analize *akademske* – jezikom, je po vsebini terminoloških. V projektu z internim naslov KAS (uradnim pa [Slovenska znanstvena besedila: viri in opis](#)) smo se zato odločili pozornost nameniti tudi tistemu drugemu, precej redkeje proučevanemu segmentu te besedilne zvrsti: leksiki, ki je za natančno, jasno in jedrnato ubesedovanje strokovne vsebine zelo pomembna ter je značilna za strokovno-znanstveno pisanje nasploh, ni pa terminološka. Lahko bi rekli, da je skupna akademskemu pisanju ne glede na vedo, stroko ali področje. Poimenovali smo jo splošno strokovna.

Zbirka ALEKS v1.0 (*'akademska leksika slovenščine'*) obsega 463 iztočnic. Prikazali smo jih v njihovi tipični okolici: s kolokatorji in zgledi rabe, pri čemer smo pri naboru kolokatorjev sledili načinu, ki smo ga za celotno slovenščino preizkusili v projektu KOLOS ([Kolokacije kot temelj jezikovnega opisa: semantični in časovni vidiki](#)). Zglede rabe smo – kot tudi že omenjene iztočnice in kolokatorje – pridobili iz [Korpusa akademske slovenščine KAS](#), in sicer z aplikacijo GDEX, ki je bila na slovenščino prilagojena pred nekaj leti. Zbirka ALEKS je enako leksikalna, kot je slovnična, saj ob iztočnicah, kot so npr. *analiza*, *pomemben*, *pogosto* in *opredeliti*, prikazuje njihovo ponovljivo (ne naključno ali redko) sopojavljanje z drugimi besedami, ki so z njimi v neposrednem skladenskem razmerju, npr. *analiza [podatkov, rezultatov, stroškov, odgovorov ...]*; v neabstrahirani obliki pa si lahko vsako od iztočnic skupaj z vsakokratnim kolokatorjem ogledamo še v treh avtentičnih primerih iz besedil. ALEKS tako vsekakor presega zgolj geslovniško naravo, ki jo je (zavajajoče) mogoče slutiti iz naslova vsebinsko sorodnega vira za ameriško angleščino [Academic Vocabulary Lists](#), obenem pa z navedenim sega tudi že proti pomenu in besedilnosti.

Zbirko smo pripravili in jo objavili na prosto dostopnem portalu Termania z namenom, da bi bila v pomoč vsem, ki šele začenjajo pridobivati znanja o strokovno-znanstvenem pisanju v slovenščini, pa tudi tistim, ki si želijo svoje tovrstne veščine le še izpiliti. Zavedamo se, da zbirka v različici 1.0 še ni popolna in brez napak, ocenjujemo pa, da bodo govorci slovenščine, ki različna nova znanja pridobivajo na univerzitetni ravni (ali pa so jih na 1. oz. 2. stopnji visokošolskega študija celo že usvojili), te nepopolnosti lahko presegli brez večjih težav.

24. maj 2019

Avtorji

II. Priprava zbirke

A. Geslovník

Iztočnice za podatkovno zbirko smo pridobili tako, da smo z metodo frekvenčnega profila primerjali dva korpusa, in sicer Korpus akademske slovenščine KAS (v0.2, 2000–2015) in korpus splošne slovenščine [Kres](#) (2012). Besede (leme) na seznamih smo najprej ločili v tri skupine: (a) terminološko skupino; (b) nepredmetnopomensko skupino (zaimki, vezniki, predlogi) in (c) splošnostrokovno skupino. Razvrščanje le v te skupine, pa sicer tudi celotna preostala analiza je potekala v kombinaciji z ogledom konkordančnih vrstic v korpusu KAS in značilne besedne okolice, ki smo jo pridobili z orodjem Besedne skice (Sketch Engine v lokalni inštalaciji [Centra za jezikovne vire in tehnologije Univerze v Ljubljani](#)). Pri prepoznavanju besed kot splošnostrokovnih smo izhajali iz dejstva, da gre za besede, ki so se uvrstile med rezultate frekvenčnega profila, kar jih je že samo po sebi potrdilo kot značilne za strokovno-znanstveno pisanje. Če je vlogo neterminološkosti, področne nespecializiranosti ali neterminološke kolokacijskosti vsaj v enem od pomenov nato potrdil še ogled njihovih konkordanc in besednih skic, smo oboje šteli za zadosten pogoj za uvrstitev v nabor leksike, ki nas je zanimala. Analiza je na koncu dala skoraj 500 splošnostrokovnih besed, izmed katerih smo v geslovník uvrstili 226 samostalnikov, 119 pridevnikov, 86 glagolov in 32 prislovov. Členkov in okrajšav v prvo različico zbirke nismo vključili, ker zanje avtomatska pridobitev besedilnega okolja še ni mogoča, tako da ALEKS v1.0 skupaj obsega, kot že omenjeno, 463 iztočnic (Priloga 1).

B. Izvoz kolokatorjev iz korpusa KAS: številske nastavitve

Iztočnice vseh štirih besednih vrstah smo razdelili v tri skupine, pri čemer smo upoštevali njihovo pogostost v korpusu KAS in minimalno pogostost kolokacije:

Samostalnik

- I. skupina: do 200.000 pojavitev, minimalna pogostost kolokacije: 50
- II. skupina: od 200.001 do 800.000 pojavitev, minimalna pogostost kolokacije: 100
- III. skupina: 800.001 in več pojavitev, minimalna pogostost kolokacije: 500

Glagol

- I. skupina: do 60.000 pojavitev, minimalna pogostost kolokacije: 30
- II. skupina: od 60.001 do 300.000 pojavitev, minimalna pogostost kolokacije: 40
- III. skupina: 300.001 in več pojavitev, minimalna pogostost kolokacije: 100

Pridevnik

- I. skupina: do 60.000 pojavitev, minimalna pogostost kolokacije: 30
- II. skupina: od 60.001 do 300.000 pojavitev, minimalna pogostost kolokacije: 40
- III. skupina: 300.001 in več pojavitev, minimalna pogostost kolokacije: 100

Prislov

- I. skupina: do 80.000 pojavitev, minimalna pogostost kolokacije: 30
- II. skupina: od 80.001 do 200.000 pojavitev, minimalna pogostost kolokacije: 40
- III. skupina: 200.001 in več pojavitev, minimalna pogostost kolokacije: 100

Za razmeroma visoko mejo minimalne pogostosti kolokacije smo se odločili zaradi velikosti korpusa KAS (v različici v0.2 je vseboval več kot 1,1 milijarde pojavnic) ter zaradi želje po zmanjšanju (pre)velike količine podatkov, ki bi lahko v končni, na portalu objavljeni obliki za uporabnika postala nepregledna in neobvladljiva.

C. Izvoz kolokatorjev iz korpusa KAS: slovnične relacije

Slovnični relaciji, za katere smo pridobili podatke iz korpusa KAS, sta bili skupaj 302. Med njimi je bilo 161 samostalniških, 59 glagolskih, 48 pridevniških in 34 prislovnih. Za ponazoritev navajamo 10 samostalniških relacij:

S_bлізу-d_x_g2
S_brez-d_s2
S_brez-d_x_g2
S_brez-d_x_s2
S_do-d_s2
S_do-d_x_g2
S_do-d_x_p2
S_do-d_x_s2
S_glede-d_s2
S_glede-d_x_g2

Izvoz kolokatorjev po slovničnih relacijah je zajemal že tudi prvo, grobo slovnično-semantično skupinjenje. To je pri pridevniških iztočnicah najprej pomenilo členitev kolokatorjev po spolu, nato pa še pomensko razvrščanje glede na statistično podobnost besednih skic kolokatorjev (pridobljeno v funkciji [Tezaver](#) v orodju Sketch Engine). Pri samostalniku, glagolu in prislovu je bil upoštevan samo zadnji kriterij. Znotraj iste skupine so bili kolokatorji avtomatsko razvrščeni po pogostosti zveze kolokatorja z iztočnico (kolokacije) od najbolj pogoste do najbolj redke. Tako predpripravljene podatke smo nato uvozili v slovarski urejevalnik iLEX in jih tam ročno pregledali.

Č. Izvoz zgledov rabe iz korpusa KAS: GDEX

Za vsako kolokacijo smo iz korpusa izvozili po tri zgleda. Pri izvozu smo uporabili orodje GDEX (angl. *Good Dictionary EXamples*), ki razvršča zgleda glede na njihovo kakovost po različnih značilnostih (npr. dolžina povedi, skladenjska zgradba, pogostost posameznih besed v povedi). Za izhodišče smo vzeli slovenske različice konfiguracij za vse štiri besedne vrste, ki so podrobneje opisane v Kosem (2015). Ključni deli konfiguracij oz. klasifikatorji so:

- Prednost imajo cele povedi.
- "Kaznuje se" vsaka pojavnica v povedi, ki ima manj kot tri pojavitve v korpusu.
- Poved mora biti daljša od 7 pojavnic.
- Poved ne sme vsebovati ponovitve iztočnice.
- Poved ne sme vsebovati elektronskega ali spletnega naslova.

- "Kaznujejo se" pojavnice, daljše od 12 znakov.
- "Kaznujejo se" povedi z več kot tremi vejicami.
- "Kaznujejo se" lastna imena.
- Upošteva se drugi kolokator, torej kolokator iskanih kolokacij. Na ta način iščemo tipične strukture, v katerih se kolokacije pojavljajo.
- Jaccardov indeks meri podobnost med nizi, s čimer se izognemo pridobivanju podobnih ali enakih zgledov.

Večina razlik med nastavitvami za posamezne besedne vrste se nahaja zgolj v nastavitvah posameznih klasifikatorjev. S to metodo pridobljene zglede smo uporabili pri potrjevanju pravilnosti kolokatorjev, kot tudi za ponazoritev rabe kolokatorjev in iztočnic v sobesedilu.

D. Prva redakcija

Zbirka ALEKS v1.0 je bila enkrat ročno pregledana. Prva redakcija je med drugim zajemala grobo razdruževanje in drugačno druženje pomensko ter strukturno sorodnih kolokatorjev, prepoznavanje značilne ali pravilne oblike kolokatorjev (množina, vrstnost, posamostaljeni pridevniki) in brisanje deiktčnih kolokatorjev. Pri tem smo ravnali inkluzivno, se pravi, da smo se brisanju raje odrekli, kot da bi že v tej fazi izpustili podatek, ki bi vendarle lahko bil informativen (in ki vsekakor ni napačen). Napake in nedoslednosti, ki so po prvem pregledu še ostale, je mogoče razrešiti predvsem z ogledom zgledov rabe, ki se neposredno odprejo s klikom na povezavo (gl. tudi razdelek IV. *Dostop in vizualizacija*).

V navezavi na projekt KOLOS bodo načrtovane prihodnje redakcije zajele še problematiko kolokacij z manjkajočim členom, glagolskih iztočnic s *si/se*, daljših, vsaj tričlenskih struktur pri prislovih, pomensko (precej) praznih kolokatorjev ipd.

III. Mikrostruktura

Geselski članek ima naslednje elemente:

- Vsaka splošnostrokovna beseda je prikazana v samostojnem geselskem članku. Beseda iztočnica ima osnovno obliko (samostalnik v imenovalniku ednine, pridevnik v obliki za moški spol ednine, glagol v nedoločniku, prislov je nestopnjevan) in je opremljena z enim, dvema ali tremi rombi, ki ponazarjajo njeno pogostost v korpusu KAS (o vrednosti rombov gl. razdelek V. *Krajšave in simboli*). Pripisan je še podatek o besedni vrsti (samostalnik, pridevnik, glagol, prislov).
- Vse iztočnice so enopomenske (čeprav bi bile pri nekaterih možne tudi drugačne rešitve).
- Sledi del geselskega članka, ki prikazuje besedilno okolje iztočnice, in sicer v obliki kolokacij in zgledov rabe. Kolokacije so členjene po slovnični zgradbi in pomenski povezanosti. Pri vsaki skupini kolokacij je najprej navedena njihova struktura, npr. *pbz0 SBZ0, sbz0 SBZ2, SBZ1 gbz* (gl. pojasnilo v seznamu krajšav v nadaljevanju). Vsaka kolokacija ima po tri korpusne zglede. Vsi zgledi so iz korpusa KAS in so avtentični, torej v ničemer spremenjeni.

- V zadnjem delu je povezava na korpus KAS v konkordančniku NoSketch Engine. S klikom nanjo se prikaže celotni konkordančni niz iztočnice v navedenem korpusu. Preklop v korpus torej omogoča še precej širšo analizo besedilnega pojavljanja iztočnice, in sicer v obsegu od enega odstavka pa do celega besedila.

IV. Dostop in vizualizacija

Zbirka ALEKS v1.0 je za iskanje prosto dostopna na portalu [Termania](#).

Prikaz na portalu Termania ima tri ravni:

A. Na vstopnem mestu portala Termania je iskalno okence, ki omogoča iskanje po vseh slovarjih oz. podatkovnih zbirkah, ki so vključene v ta portal. Tu se iskalni zadetki zbirke ALEKS pokažejo v naslednjem obsegu: iztočnica in podatek o besedni vrsti. Npr.:

 **analiza** samostalnik

B. S klikom na iztočnico se odpre druga raven, na kateri so prikazane kolokacije, poskupinjene po strukturi ter znotraj te po slovnični in pomenski povezanosti. Npr.:

analiza*** samostalnik

SBZ0 sbz2

analiza [podatkov, rezultatov, stroškov, odgovorov, tveganja, virov, vsebine, vzorca] [več...](#)

analiza [stanja, vplivov, dejavnikov, prednosti] [več...](#)

analiza [poslovanja, okolja, trga, procesa, dela, podjetja, primera, sistema] [več...](#)

analiza [variance] [več...](#)

analiza [vprašalnikov, ankete] [več...](#)

analiza [uspešnosti] [več...](#)

analiza [občutljivosti] [več...](#)

gbz na SBZ4

[osredotočiti se, nanašati se, osredotočati se] na analizo [več...](#)

[omejiti se, vplivati] na analizo [več...](#)

gbz z SBZ6

[ukvarjati se, pomagati si] z analizo [več...](#)

[ugotoviti, preveriti, ugotavljati, preverjati, določiti, analizirati, potrditi, primerjati, dokazati, prikazati] z analizo [več...](#)

[začeti, dobiti, pridobiti, nadaljevati, pričeti, zaključiti, priti, doseči] z analizo [več...](#)

[obdelati] z analizo [več...](#)

Na koncu geselskega članka je na voljo še povezava na Korpus akademskih besedil KAS. S klikom na povezavo se v novem okencu prikažejo konkordančne vrstice iskane iztočnice v imenovanem korpusu, dostopnem na naslovu <https://www.clarin.si/noske/>. Npr.:

gbz k SBZ3
[pristopiti] k analizi več...

Korpus akademskih besedil KAS

C. Na tretjo raven pride uporabnik s klikom na *več...* Ta odpre zglede rabe, in sicer po tri za vsako kolokacijo. Zgledi se zaprejo s klikom na *manj...* Npr.:

sbz0 iz SBZ2

[podatek, rezultat] iz analize *manj...*

- S pomočjo pridobljenih podatkov iz *analize* lahko za konec potrdimo oziroma zavrnemo raziskovalne tematike in odgovorimo na raziskovalno vprašanje.
- Spremljanje podatkov iz *analize* anket je nujen pogoj za uspeh.
- Tako podatki iz različnih *analiz* in virov kot tudi mnenja o njeni uspešnosti so različna.
- Rezultate iz *analize* konsistentnosti smo dobili nekoliko drugače.
- Predvideni rezultati iz *analize* ankete bodo pokazali pomembne razlike med tistimi z višjo stopnjo izobrazbe in tistimi z nižjo.
- Pogostost vseh neželenih dogodkov po dnevih je primerljiva z rezultati iz prejšnje *analize* v opazovanem obdobju.

[ugotovitev] iz analize *več...*

Iskanje na portalu v zbirki ALEKS v1.0 poteka po polju iztočnica (ne pa tudi drugih).

V. Krajšave in simboli

gbz	glagolska besedna zveza
GBZ	glagolska besedna zveza z iztočnico
inf-gbz	glagolska besedna zveza v nedoločniku
inf-GBZ	glagolska besedna zveza z iztočnico v nedoločniku
Neg-GBZ	glagolska besedna zveza z iztočnico v zanikani obliki
pbz	pridevniška besedna zveza
PBZ	pridevniška besedna zveza z iztočnico
pbz0, pbz1, pbz2 ...	pridevniška besedna zveza, številka označuje njen sklon: 0 = katerikoli sklon, 1 = imenovalnik, 2 = rodilnik itd.
rbz	prislovna besedna zveza
RBZ	prislovna besedna zveza z iztočnico
sbz	samostalniška besedna zveza

SBZ	samostalniška besedna zveza z iztočnico
sbz0, sbz1, sbz2 ...	samostalniška besedna zveza, številka označuje njen sklon: 0 = katerikoli sklon, 1 = imenovalnik, 2 = rodilnik itd.
Vez-	besedna zveza s pomožnim glagolom
◆	pogostost v korpusu KAS v2.0: samostalniki do 200.000 pojavitev; glagoli do 60.000 pojavitev; pridevniki do 60.000 pojavitev; prislovi do 80.000 pojavitev
◆ ◆	pogostost v korpusu KAS v2.0: samostalniki od 200.001 do 800.000 pojavitev; glagoli od 60.001 do 300.000 pojavitev; pridevniki od 60.001 do 300.000; prislovi od 80.001 do 200.000 pojavitev
◆ ◆ ◆	pogostost v korpusu KAS v2.0: samostalniki 800.001 in več pojavitev; glagoli 300.001 in več pojavitev; pridevniki 300.001 in več pojavitev; prislovi 200.001 in več pojavitev

Literatura

- ERJAVEC, Tomaž, FIŠER, Darja, LJUBEŠIĆ, Nikola, LOGAR, Nataša, OJSTERŠEK, Milan (2016). Slovenska akademska besedila: prototipni korpus in načrt analiz. ERJAVEC, Tomaž, FIŠER, Darja (ur.): *Zbornik konference Jezikovne tehnologije in digitalna humanistika*. Ljubljana: Znanstvena založba Filozofske fakultete. 58–64. Dostopno prek: <http://www.sdjt.si/wp/wp-content/uploads/2016/09/ITDH-2016-Erjavec-et-al-Slovenska-akademska-besedila.pdf>.
- KILGARRIFF, Adam, HUSÁK, Milos, MCADAM, Katy, RUNDELL, Michael, Rychlý, Pavel (2008). GDEX: Automatically Finding Good Dictionary Examples in a Corpus. BERNAL, Elisenda, DECESARIS, Janet (ur.): *Proceedings of the 13th EURALEX International Congress*. Barcelona: Institut Universitari de Linguística Aplicada, Universitat Pompeu Fabra. 425–432. Dostopno prek: https://www.academia.edu/2862022/GDEX_Automatically_finding_good_dictionary_examples_in_a_corpus.
- KOSEM, Iztok, HUSAK, Milos, MCCARTHY, Diana (2011). GDEX for Slovene. KOSEM, Iztok, KOSEM, Karmen (ur.): *Electronic Lexicography in the 21st Century: New Applications for New Users*. Ljubljana: Trojina, Institute for Applied Slovene Studies. 151–159. Dostopno prek: <http://elex2011.trojina.si/Vsebine/proceedings/eLex2011-19.pdf>.
- KOSEM, Iztok, GANTAR, Polona, KREK, Simon (2013). Avtomatizacija leksikografskih postopkov. *Slovenščina 2.0* 1/2. 139–164. Dostopno prek: <https://doi.org/10.4312/slo2.0.2013.2.139-164>.
- KOSEM, Iztok (2015). Slovarki zgledi. GORJANC, Vojko, GANTAR, Polona, KOSEM, Iztok, KREK, Simon (ur.): *Slovar sodobne slovenščine: problemi in rešitve*. Ljubljana: Univerza v Ljubljani, Filozofska fakulteta. 320–339.
- KOSEM, Iztok, KREK, Simon, GANTAR, Polona, ARHAR HOLDT, Špela, ČIBEJ, Jaka, LASKOWSKI, Cyprian (2018). Collocations Dictionary of Modern Slovene. ČIBEJ, Jaka, idr. (ur.): *Proceedings of the 18th EURALEX International Congress: Lexicography in Global Contexts*. Ljubljana: Ljubljana University Press, Faculty of Arts. 989–997. Dostopno prek: <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/118/211/3000-1>.
- KOSEM, Iztok, KOPPEL, Kristina, ZINGANO KUHN, Tanara, MICHELFEIT, Jan, TIBERIUS, Carole (2019). Identification and Automatic Extraction of Good Dictionary Examples: The Case of GDEX. *International Journal of Lexicography* 32/2. 119–137. doi: [10.1093/ijl/ecy014](https://doi.org/10.1093/ijl/ecy014).
- LOGAR, Nataša, ERJAVEC, Tomaž (2017). Slovene Academic Writing: A Corpus Approach to Lexical Analysis. *Forum für Fachsprachen-forschung (FFF)*. Berlin: Frank & Timme. (V tisku.)

- LOGAR, Nataša, ERJAVEC, Tomaž (2018). Strokovno-znanstvena slovenščina: besednovrstne in oblikoskladenjske značilnosti. FIŠER, Darja, PANČUR, Andrej (ur.): *Zbornik konference Jezikovne tehnologije in digitalna humanistika*. Ljubljana: Znanstvena založba Filozofske fakultete. 175–180. Dostopno prek: http://www.sdit.si/wp/wp-content/uploads/2018/09/ITDH-2018_Logar-et-al_Strokovno-znanstvena-slovenscina-besednovrstne-in-oblikoskladenjske-znacilnosti.pdf.
- LOGAR, Nataša, GANTAR, Polona, KOSEM, Iztok (2014). Collocations and Examples of Use: A Lexical-semantic Approach to Terminology. *Slovenščina 2.0* 2/1. 41–61. Dostopno prek: http://www.trojina.org/slovenscina2.0/arhiv/2014/1/Slo2.0_2014_1_03.pdf.
- LOGAR, Nataša (2013). *Korpusna terminografija: primer odnosov z javnostmi*. Ljubljana: Trojina, zavod za uporabno slovenistiko; Fakultete za družbene vede.
- PORI, Eva, KOSEM, Iztok (2018). V iskanju slovarsko relevantne kolokacije na primeru struktur s prislovi. *Slovenščina 2.0* 6/2. 154–185. doi: [10.4312/slo2.0.2018.2.154-185](https://doi.org/10.4312/slo2.0.2018.2.154-185).

Priloga

Priloga 1: ALEKS v1.0: Leksikalno-skladenjska podatkovna zbirka slovenskega strokovno-znanstvenega jezika – geslovník

1. aktualen	36. določati	71. izmeriti
2. analiza	37. določen	72. izmerjen
3. analiziran	38. določitev	73. izpostaviti
4. analizirati	39. določiti	74. izpostavljati
5. anketa	40. dolžina	75. izpostavljen
6. anketiranec	41. doseči	76. izračunan
7. anketiran	42. dosednji	77. izračunati
8. avtor	43. doseganje	78. izračun
9. besedilo	44. dosežen	79. izrazito
10. bistveno	45. dostopen	80. izrazit
11. bistven	46. dostopno	81. izražati
12. celota	47. element	82. izvajanje
13. celoten	48. empiričen	83. izvajati
14. celovit	49. enak	84. izvedba
15. cilj	50. faza	85. izveden
16. časovno	51. formulacija	86. izvesti
17. definicija	52. glaven	87. izziv
18. definiranje	53. grafikon	88. jedro
19. definiran	54. graf	89. kakovost
20. definirati	55. hipoteza	90. kasen
21. dejanski	56. hkrati	91. kategorija
22. dejavnik	57. identificirati	92. kazalec
23. delež	58. informacija	93. kazalnik
24. delno	59. interpretacija	94. kazati
25. delo	60. intervju	95. ključen
26. delovanje	61. intervjuvanec	96. količina
27. del	62. izbira	97. kombinacija
28. diagram	63. izbor	98. koncept
29. disertacija	64. izbran	99. končen
30. dobljen	65. izdelava	100. konkreten
31. dodatek	66. izhajati	101. kontekst
32. dodaten	67. izhodiščen	102. kriterij
33. dodatno	68. izhodišče	103. lahko
34. doktorski	69. izkazovati	104. lastnost
35. določanje	70. izkušnja	105. leto

106. literatura	146. neustrezen	186. opisovati
107. majhen	147. nivo	187. opis
108. maksimalen	148. nizek	188. opozarjati
109. manj	149. nov	189. opravljen
110. medsebojen	150. obdelava	190. opredelitev
111. medsebojno	151. obdobje	191. opredeliti
112. mehanizem	152. običajno	192. opredeljen
113. mentor	153. objavljen	193. opredeljevati
114. mera	154. oblika	194. optimalen
115. meritev	155. oblikovanje	195. osnova
116. meriti	156. oblikovati	196. osrednji
117. merjenje	157. območje	197. osredotočati
118. metoda	158. obravnavanje	198. ostal
119. metodologija	159. obravnavan	199. ovrednotenje
120. mnenje	160. obravnava	200. ovrednotiti
121. morebiten	161. obravnavati	201. označen
122. možen	162. obseg	202. označevati
123. možnost	163. obstajati	203. označiti
124. nabor	164. obstoječ	204. paradigma
125. način	165. ocena	205. parameter
126. nadalje	166. oceniti	206. podajati
127. nadaljevanje	167. ocenjevanje	207. podan
128. nadaljnji	168. ocenjevati	208. podatek
129. največ	169. odgovor	209. podlaga
130. nakazovati	170. odločanje	210. podobno
131. naloga	171. odločitev	211. podpoglavje
132. namen	172. odnos	212. podroben
133. nanašati	173. odražati	213. podrobno
134. naslednji	174. odstavek	214. področje
135. nastanek	175. odstopanje	215. poenostavljen
136. nastati	176. odstotek	216. poglavje
137. naštet	177. odvisnost	217. pogoj
138. natančen	178. okvir	218. pogosto
139. natančnost	179. omejen	219. pojasnjen
140. navajati	180. omenjen	220. pojavljati
141. naveden	181. omogočati	221. pojav
142. navesti	182. omogočiti	222. pojem
143. negativen	183. opazovanje	223. pojmovanje
144. negativno	184. opazovan	224. pomanjkljivost
145. neposredno	185. opisan	225. pomemben

226. pomembno	266. praktičen	306. primer
227. pomembnost	267. pravilen	307. pripomoči
228. pomeniti	268. praviloma	308. priprava
229. pomen	269. predhoden	309. prisoten
230. pomoč	270. predhodno	310. prispevati
231. populacija	271. predlagan	311. prispevek
232. posamezen	272. predmet	312. pristop
233. posebej	273. prednost	313. problematika
234. poseben	274. predpostaviti	314. problem
235. poskus	275. predpostavka	315. proces
236. posledica	276. predpostavljati	316. projekt
237. posledično	277. predstavitev	317. proučevanje
238. posredno	278. predstaviti	318. proučevan
239. postavljen	279. predstavljati	319. proučevati
240. postopek	280. predstavljen	320. raven
241. potekati	281. predvidevati	321. razdeliti
242. potek	282. preglednica	322. raziskava
243. potrditi	283. pregled	323. raziskovalec
244. potreben	284. preizkus	324. raziskovalen
245. potrebno	285. prenos	325. raziskovanje
246. potrjevati	286. prepoznavanje	326. različen
247. poudarek	287. presoja	327. različno
248. poudariti	288. pretek	328. razlika
249. poudarjati	289. preučevanje	329. razlikovanje
250. povečanje	290. preučevan	330. razlikovati
251. povečati	291. preverjanje	331. razlog
252. povečevati	292. prevladujoč	332. razmerje
253. povezanost	293. pričakovan	333. razprava
254. povezan	294. pridobiti	334. razširjen
255. povezava	295. pridobljen	335. razumevanje
256. povezovati	296. prikazan	336. razviden
257. povprečje	297. prikazovati	337. razviti
258. povzetek	298. prikaz	338. razvit
259. povzet	299. prilagojen	339. razvoj
260. povzročati	300. priloga	340. razvrščanje
261. povzročiti	301. primaren	341. relativno
262. pozitiven	302. primeren	342. reševanje
263. pozitivno	303. primerjati	343. rešitev
264. poznavanje	304. primerjava	344. rezultat
265. praksa	305. primerljiv	345. sekundaren

346. sestava	386. teoretičen	426. visok
347. sestavljen	387. teorija	427. vključen
348. shema	388. termin	428. vključevati
349. sinteza	389. testiranje	429. vključiti
350. sistematičen	390. teza	430. vloga
351. sistem	391. tovrsten	431. vplivati
352. skladno	392. trditev	432. vpliv
353. skladnost	393. trend	433. vprašalnik
354. sklep	394. učinek	434. vprašanje
355. sklop	395. učinkovitost	435. vrednost
356. skupen	396. učinkovit	436. vrednotenje
357. skupina	397. ugotavljanje	437. vrsta
358. slednji	398. ugotavljati	438. vsebina
359. slika	399. ugotovitev	439. vsebinski
360. smernica	400. ugotoviti	440. vsebovati
361. smiseln	401. ugotovljen	441. vzorec
362. smisel	402. uporaba	442. vzrok
363. sodelavec	403. uporabiti	443. zadosten
364. sodelovanje	404. uporabljati	444. zagotavljanje
365. sodelovati	405. uporabljen	445. zahtevati
366. sodelujoč	406. upoštevanje	446. zajemati
367. sodoben	407. upoštevati	447. zajet
368. sposobnost	408. usmerjen	448. zaključek
369. sprememba	409. uspešno	449. zaključen
370. spremenljivka	410. uspešnost	450. zapisati
371. spreminjanje	411. ustrezen	451. zaporeden
372. spreminjati	412. ustrezno	452. zaslediti
373. stališče	413. ustreznost	453. zasnova
374. stanje	414. uvesti	454. zastavljen
375. stolpec	415. uvod	455. zbiranje
376. stopnja	416. uvrščati	456. zmanjšanje
377. struktura	417. večina	457. zmanjšati
378. širok	418. velikost	458. značilen
379. številen	419. velik	459. značilno
380. število	420. veljati	460. značilnost
381. študija	421. veljavnost	461. znanstven
382. tabela	422. verjetnost	462. znašati
383. tematika	423. vezan	463. želen
384. temeljiti	424. vidik	
385. temelj	425. vir	

