

1 DEBDict

DEBDict je nástroj platformy DEB, vyvinutý RNDr. Adamem Rambousekem, Ph.D., na Fakultě informatiky Masarykovy univerzity (FI MU), poskytující přístup ke slovníkům českého jazyka, morfologickému analyzátoru češtiny *ajka*, českému WordNetu (umožňující vyhledávat sémantické sítě) a CIA WorldFactbooku (v němž naleznete všechny země světa).¹ Zpočátku byl nástroj doplňkovou aplikací internetového prohlížeče Mozilla Firefox (rozšíření prohlížeče již není aktuální). Z rozšíření Firefoxu se stala webová aplikace.

Ze slovníků si můžete vybrat mezi *Akademickým slovníkem cizích slov*, *Slovníkem spisovné češtiny*, *Příručním slovníkem jazyka českého*, *Slovníkem spisovného jazyka českého*, *Slovníkem české frazeologie a idiomatiky I–IV* (rozdělený na slovesnou a neslovesnou část) a *Slovníkem českých synonym*.² Uvedené zdroje jsou dostupné pro všechny registrované uživatele aplikace.

1.1 Výkladové slovníky

1.1.1 Příruční slovník jazyka českého

Příruční slovník jazyka českého (PSJČ) je slovníkem velkého rozsahu skládající se z osmi dílů v devíti svazcích, protože čtvrtý díl je rozdělen do dvou svazků. Obsahuje 250 000 hesel a jeho rozsah je celkem 10 824 stran. Nebyl pojat jako kodifikační, šlo o deskriptivní (popisný) slovník, který se zaměřoval na upevnění spisovné slovní zásoby. PSJČ představuje podrobný lexikografický popis slovní zásoby. Stavba hesel vychází ze základního významu, zachycuje polysémii hesla a obsahuje propracované definice významů. Další podstatnou částí hesla jsou doklady (především z beletrie, která byla u dobrých autorů považována za nositele kvalitního spisovného jazyka) uváděné z velké části formou citátů. Snaha o ujasnění spisovnosti a nespisovnosti přispěla k tomu, že přinesl první významný pokus o stylistické hodnocení lexikálních jednotek.

Přípravné lexikografické práce na slovníku započaly ve 20. letech 20. století. Získané materiály byly ověřovány se slovníky jiného druhu, např. překladovými slovníky Josefa Jungmanna. Práce se zúčastnili přední čeští jazykovědci a na tvorbě hesel se podílelo 26 autorů. Přípravné práce skončily v roce 1934 a u příležitosti stého výročí vydání překladového slovníku Josefa Jungmanna vyšel v únoru 1935 první sešit Příručního slovníku jazyka českého. Dále byly vydávány

¹DEBII — Dictionary Editor and Browser: DEBDict — obecný prohlížeč slovníků [online]. Brno: Centrum zpracování přírozeného jazyka, 2010 [cit. 2017-01-05]. Dostupné z: <http://deb.fi.muni.cz/debdict/index-cs.php>

²RAMBOUSEK, Adam. *Lexikografická platforma na vývoj slovníkových aplikací* [online]. Brno, 2010 [cit. 2016-11-22]. Dostupné z:

knižní verze.

Slovník vydávalo Státní nakladatelství/SNP, Praha v letech 1935–1957. Do elektronické verze jej převedl Ústav pro jazyk český AV ČR ve spolupráci s Fakultou informatiky Masarykovy univerzity v roce 2007.

Hlavními redaktory výše uvedených svazků byly akademik Bohuslav Havránek, univ. prof. dr. Oldřich Hujer, prof. PhDr. Emil Smetánka, univ. prof. dr. Vladimír Šmilauer, prof. PhDr. Miloš Weingarta dr. Alois Získal.³

1.1.2 Slovník spisovného jazyka českého

Slovník spisovného jazyka českého (SSJČ) obsahuje 192 908 zpracovaných slov ve čtyřech svazcích, a proto se řadí ke slovníkům středního typu. Svým rozsahem je menší než PSJČ a zároveň plní funkci prakticky využitelného slovníku, tudíž jde o slovník kodifikační (jak po stránce gramatické a pravopisné, tak i výslovnostní).

Přípravné práce započaly v 50. letech 20. století. Slovník se opíral o materiály získané při práci na PSJČ a zachycuje terminologii. Menší rozsah vedl k hutnosti heslového odstavce, proto nedokládá význam slov, ale jen typická slovní spojení. SSJČ přináší podrobnou stylistickou charakteristiku daných slov. Slovník naznačuje slovesné vazby a zachycuje vidové dvojice. Na rozdíl od PSJČ je slovník hnízdován tzn., že do heslového odstavce zařazuje zdobněliny, předponová slovesa, k přídavným jménům příslovce atd.

Kromě prvního knižního vydání (1960–1971) byl zároveň vydán v sešitech (1958–1971), vyšlo 40 sešitů. Ve čtvrtém díle jsou doplňky 984 nově zařazených slov a opravy. V roce 1989 slovník vyšel ve druhém vydání tentokrát v osmi svazcích. Na elektronické podobě se podílel Ústav pro jazyk český AV ČR s Fakultou informatiky Masarykovy univerzity v roce 2002.

Hlavními redaktory byli akademik Bohuslav Havránek, PhDr. Jaromír Bělič, DrSc., PhDr. Miloš Helcl, CSc., prof. dr. Alois Jedlička, prof. dr. Václav Křístek, CSc. a prof. PhDr. František Trávníček. Rukopis slovníku připravili členové lexikografického oddělení Ústavu pro jazyk český ČSAV.⁴

1.1.3 Slovník spisovné češtiny

Slovník spisovné češtiny pro školu a veřejnost (SSČ) je slovník s menším rozsahem a kodifikačním statutem. Obsahuje 45 366 heslových slov a 62 872 vyčleněných významů. Heslo obsahuje poučení o pravopisu (také o skloňování a časování) a výslovnosti slov. Výběr materiálu se přesouvá (na rozdíl od PSJČ a SSJČ)

³ *Příruční slovník jazyka českého* (1935–1957) [online]. Praha: Ústav pro jazyk český, 2008 [cit. 2017-01-05]. Dostupné z: <http://bara.ujc.cas.cz/psjc/>

⁴ HAVRÁNEK, Bohuslav (ed.). *Slovník spisovného jazyka českého*. 2., nezměněné vyd. Praha: Academia, 1989, 8 sv.

k publicistice. Zvláštní pozornost pak věnuje dubletivním tvarům. Ke slovníku jsou připojeny přílohy podávající přehled o tvoření slov v češtině, soupisy rodných jmen a příjmení a zeměpisných jmen, jakož i běžných zkratek a značek.

Počátek práce na SSČ začaly v 70. letech. Tento slovník byl zamýšlen a také vydán jako jednosvazkový. První vydání vyšlo v roce 1978, druhé opravené a doplněné vydání pak v roce 1994, třetí opravené vydání 2003 a další vydání byla v letech 2004, 2005 a 2009. Nakladatelství LEDA, připravilo elektronické vydání v roce 2005.

Hlavními redaktory slovníku jsou PhDr. Josef Filipec, CSc., prof. PhDr. František Daneš, DrSc., PhDr. Jaroslav Machač (1. vydání) a Vladimír Mejstřík, prom. ped. (2. a 3. vydání).⁵

1.2 Slovníky současného jazyka

1.2.1 Akademický slovník cizích slov

Akademický slovník cizích slov (ASCS) je speciální normativní slovník, jenž obsahuje 100 000 významů slov, citátových spojení a běžných zkratek a značek cizího původu, je rozdělený do dvou dílů.

Jednotlivá slova uvádějí poučení o pravopisu, výslovnosti a významu, dále pak sdělují původ slova, tvaroslovnou a slovnědruhovou charakteristiku, stylové zařazení a kontextové užití. Slovník zachycuje dynamiku ve vývoji české slovní zásoby, jejího obohacování, posunu ve stylistickém hodnocení slov přejatých a významové změny. Reaguje na zásadní proměny odehrávající se v rozvíjející domácí slovní zásobě a hlavně na ty, které se vztahují ke konkrétním vědním oborům, disciplínám a řemeslům.

V roce 1995 vyšla knižní verze v nakladatelství Academia, posléze v roce 1997 vyšel dotisk. V novějších vydáních byl přejmenován na *Nový akademický slovník cizích slov* (NASCS). V roce 1999 vznikla v nakladatelství LEDA, elektronická verze 1.0 s názvem *Velký slovník cizích slov* (VSCS) a v roce 2005 vyšel slovník ještě jednou jako verze 2.0, která už není součástí nástroje platformy DEB.

ASCS napsal kolektiv autorů pod vedením doc. PhDr. Věry Petráčkové, CSc. a prof. PhDr. Jiřího Krause, DrSc.⁶

⁵MEJSTŘÍK, Vladimír, Josef FILIPEC, František DANEŠ a Jaroslav MACHAČ. *Slovník spisovné češtiny pro školu a veřejnost: s Dodatkem ministerstva školství, mládeže a tělovýchovy České republiky*. Vyd. 3., opr. Praha: Academia, 2003. ISBN 80-200-1080-7.

⁶PETRÁČKOVÁ, Věra, Jiří KRAUS a kol. *Akademický slovník cizích slov: [A-Ž]*. Dotisk. Praha: Academia, 1997. ISBN 80-200-0607-9.

1.2.2 Slovník české frazeologie a idiomatiky

Slovník české frazeologie a idiomatiky I—IV (SČFI) popisuje českou frazeologii. Slovník je dílem vypracovaným týmem lingvistů z Filozofické fakulty Univerzity Karlovy a v menší míře i Ústavu pro jazyk český AV ČR.⁷ Dílo usiluje o všestranný popis variant frazémů, které jsou užívány v současné době.

Slovníkové heslo podrobně člení kontext užívání frazému od jeho významu a funkce přes příklady, synonymní nebo opozitní řady, tak i přímý odkaz k onomaziologickému slovníku. Frazémy s vyšší frekvencí mají uvedené ekvivalenty ve čtyřech jazycích. Každý díl je vybaven odbornou studií té oblasti, již díl popisuje.

První díl slovníku se jmenuje *Přirovnání* a věnuje se nevětným výrazům, tedy materiálům mající v sobě spojkou jak/jako. Do teď přirovnání nebyla jako celek zpracována. Poprvé vyšel v roce 1983. Druhý díl nazývající se *Výrazy neslovesné* se zabývá několika oblastmi frazeologie, a to výrazům jmenným, adverbialním a všem typům frazémů gramatickým. První výtisk byl v roce 1988. Třetí díl s titulem *Výrazy slovesné* zahrnuje všechny slovesné frazémy, které jsou zároveň koločnicemi frazémy, nezahrnují frazémy, jež mají v sobě sloveso. Můžeme v něm listovat od roku 1994. Poslední čtvrtý díl *Výrazy větné* se orientuje na ustálené výrazy s podobou věty. Tento díl je založený na *Českém národním korpusu* (ČNK). Vyšel v roce 2009 a tehdy byla zároveň první tři díla doplněna o řadu hesel z ČNK. Všechna díla slovníku byla ve stejném roce vydána elektronicky nakladatelstvím LEDA, které vydávalo i tištěné verze. V DEBDictu je použita verze přímo od autora.

Autory slovníku jsou prof. PhDr. František Čermák, DrSc., doc. PhDr. Jiří Hronek, CSc. a PhDr. Jaroslav Machač.⁸

1.2.3 Slovník českých synonym

Slovník českých synonym (SČS) je praktická příručka obsahující něco málo přes 20 000 heslových slov. Je určena spíše pro rodilé mluvčí, kteří dokáží rozlišit nepatrné významové rozdíly mezi slovy.

Heslová slova jsou řazena abecedně a u každého jsou uvedeny jeho synonymní výrazy, tj. slova a slovní spojení se stejným či podobným významem, které lze libovolně nahradit. Synonymní řady začínají nejobecnějším a neutrálním slovem, vedle něho se pak řadí podobné a blízké výrazy. Tvary sloves se uvádějí s valencemi, a pokud možno v obou videch. U frekventovaných slov jsou uváděny jejich antonyma.

⁷ČERMÁK, František, Jiří HRONEK a Jaroslav MACHAČ. *Slovník české frazeologie a idiomatiky*. 2., přeprac. a dopl. vyd., V nakl. Leda vyd. 1. Praha: Leda, 2009, 507 s. ISBN 978-80-7335-215-8.

⁸Tamtéž.

Slovník poprvé vyšel v roce 1994. O dva roky později, tj. 1996, nakladatelství Lidové noviny vydalo druhé opravené vydání a 2000 vydalo třetí doplněnou verzi slovníku. V roce 2004 nakladatelství zveřejnilo elektronickou verzi tohoto slovníku. Nástroj platformy DEB používá verzi přímo od autora.

Rukopis slovníku lektoroval a k vydání doporučil doc. PhDr. František Čermák, DrSc.⁹ Autoři jsou doc. PhDr. Karel Pala, CSc. a Jan Všíanský.

1.3 Český WordNet

WordNet je lexikální databáze, která uspořádává jazykové jednotky podle sémantických vztahů. Heslem této databáze je synset neboli synonymická řada (pokud se dvě slova objeví vedle sebe, tak jsou synonymní). Jedno slovo může mít i více významů, tj. může být synonymní s více slovy, v tomto případě je dané slovo polysémické.

WordNet tvoří dvě roviny. Synsety jsou jednou a jejich provázání je druhou rovinou. Až jejich sémantické vztahy tvoří celkový WordNet. Některé příklady sémantických vztahů jsou synonymie, antonymie, meronymie a holonymie, hyperonymie a hyponymie.

Verze použitá v DEBDictu je 2.1, která obsahuje 117 597 synonymických řad, z čehož 81 426 tvoří synsety podstatných jmen, 13 650 synsety sloves, 18 877 synsety přídavných jmen a 3664 synsety příslovcí.¹⁰

1.4 Morfologický analyzátor *ajka*

Morfologický analyzátor ajka byl vyvinutý na Fakultě informatiky MU Mgr. Radkem Sedláčkem, Ph.D., podle popisu segmentace českých slov z disertační práce doc. PhDr. Kláry Osolobě, Dr. Základem algoritmu je členění slova na kmenový základ, intersegment a koncovku. Morfologický analyzátor pracuje se strojovým slovníkem kmenových základů, definičním souborem koncovkových množin a vzorů.¹¹

Analyzátor je efektivní jak při morfologické analýze, tak při generování správných gramatických tvarů. *Ajka* musí pro každou odchylku (nepravidelnost) vytvořit zvláštní vzor, což vedlo k vytvoření nového analyzátoru *majka*, který není součástí DEBDictu. Současný počet analyzovaných tvarů je 1839 vzorů.¹²

⁹PALA, Karel a Jan VŠIANSKÝ. *Slovník českých synonym*. 2. opr. vyd. Praha: Nakladatelství Lidové noviny, 1996, 439 s. ISBN 80-710-6059-3.

¹⁰ČAPEK, Tomáš. *Systém pro částečné sémantické značkování volného textu* [online]. Brno, 2006 [cit. 2016-11-22]. Dostupné z:

¹¹SEDLÁČEK, Radek. *Morfologický analyzátor češtiny* [online]. Brno, 9999 [cit. 2016-11-22]. Dostupné z: <https://nlp.fi.muni.cz/projekty/ajka/ajka.pdf>.

¹²HUSÁROVÁ, Dagmar. *Určování tvarotvorných vzorů neznámých slov* [online]. Brno, 2015

1.5 CIA World Factbook

První verze *World Factbook* byla publikována knižně v letech 1980 a obsahuje 165 zemí na 225 stranách. Od té doby vychází téměř každý rok a byl přejmenován na *The World Factbook*.¹³ První verze obsahovala informaci o počtu obyvatelstva, o jeho hospodářství, vládě, vodstvu, komunikaci a o obranných silách. Téměř při každém vydání jsou přidávány nové státy (nově vzniklé rozpadem mocností) a výše uvedené položky obměňovány. V roce 1997 společnost *Central Intelligence Agency* (CIA) představila první elektronickou verzi *Factbooku*. Použitá verze v DEBDictu je z roku 2009. V současné době obsahuje přes 250 států včetně velkých vodních ploch.

[cit. 2016-11-22]. Dostupné z:

¹³*Central Intelligence Agency: The World Factbook* [online]. Washington, D.C.: Central Intelligence Agency, 2013 [cit. 2017-01-05]. Dostupné z: <https://www.cia.gov/library/publications/the-world-factbook/>

2 Manuál pro DEBDict

2.1 Přístup k obecnému prohlížeči slovníků

DEBDict je nástroj, který umožňuje prohlížet slovníky v elektronické formě. Tato možnost vám dává příležitost podívat se do slovníků, aniž byste museli chodit do knihovny.

Pro přístup je nutné, abyste měli své vlastní přihlašovací údaje. Sice můžete použít tzv. demo verzi, ovšem nebudete mít přístup ke všem základním zdrojům, které jsou zpřístupněny registrovaným uživatelům.

Pro získání vlastních údajů, a tudíž plnému přístupu, musíte vyplnit formulář (NLPCentrum) (do kolonky: „Co budete využívat“ napište DEBDict), který následně vytisknete, podepíšete a odešlete na adresu Masarykovy univerzity. Je možné oskenovaný formulář zaslat i elektronicky e-mailem. Poté stačí počkat na přidělení uživatelských údajů a poté se směle pustit do užívání prohlížeče slovníků.

Prohlášení pro dočasné uživatele programů a dat Centra zpracování přirozeného jazyka (NLPCentrum) při Fakultě informatiky Masarykovy univerzity

 English/anglicky

Jméno a příjmení:

E-mail:

Organizace/pracoviště:

Adresa:

Co budete využívat:

Po vyplnění formuláře stiskněte , podepíšte jej a pošlete na adresu uvedenou níže.

Potvrzuji svým podpisem, že souhlasím s níže uvedenými podmínkami užívání programů a dat (korporusů, slovníků, ...) NLPCentra:

- Při využívání dat získaných z NLPCentra nebo z výstupů programů NLPCentra uvedu jako jejich zdroj příslušný produkt NLPCentra.
- Citovaný výstup nesmí být delší než 200 slov souvislého textu.
- Kopii (nebo elektronickou podobu) práce, ve které byla použita data či výstupy z produktů NLPCentra, zařadu po jejím uveřejnění k archivaci do NLPCentra.
- Přístup k programům a datům je umožněn pouze pro vědecké a výzkumné účely a nelze je bez předchozí smlouvy používat ke komerčním účelům.
- Právo užívat programy a data NLPCentra je nepřenosné.
- NLPCentrum při Fakultě informatiky Masarykovy univerzity si vyhrazuje právo kdykoliv přístup k programům a datům omezit či zrušit.

v dne

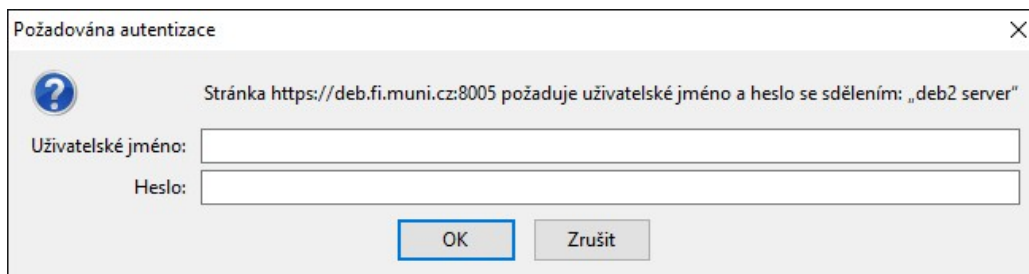
Obrázek 1: Formulář pro přístup k programům z NLPCentra

2.2 Přihlášení

Původní rozšíření prohlížeče Mozilla Firefox již nefunguje, s touto změnou odpadá nutnost instalace. K přístupu k prohlížeči slovníků je třeba použít webovou aplikaci, do níž se přihlásíte pod svými přihlašovacími údaji, získanými zasláním prohlášení, viz kapitola 2.1.

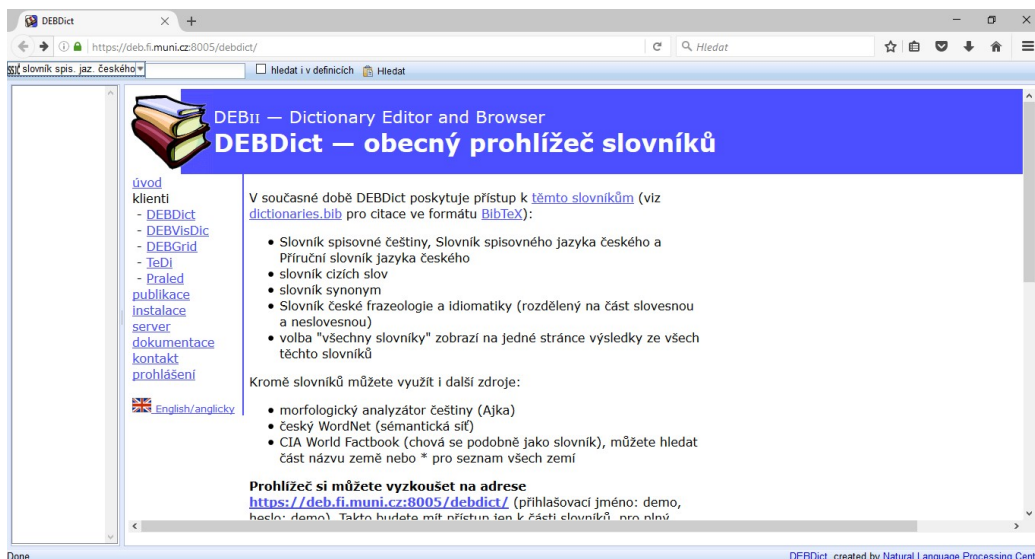
Otevřete si kterýkoli internetový prohlížeč (nemusí se jednat pouze o Mozillu Firefox). Najed'te na stránky nástroje DEBDict (Rambousek, 2010c). Při načítání

bude aplikace požadovat napsání uživatelského jména a hesla. V demo verzi je uživatelské jméno i heslo *demo*. Po jejich zadání se objeví titulní strana.



Obrázek 2: Přihlášení

Po přihlášení se zobrazí hlavní strana programu. V horním levém rohu, pod lištou záložek, máte rozbalovací nabídku s aktuálním vybraným elektronickým zdrojem, vedle něhož se nachází dotazovací řádek (textové pole) k hledání požadovaného slova. Pod nabídkou se slovníky je připraveno pole pro seznam nalezených slov. Zbytek místa v nástroji je pro vyhledaná heslová slova.

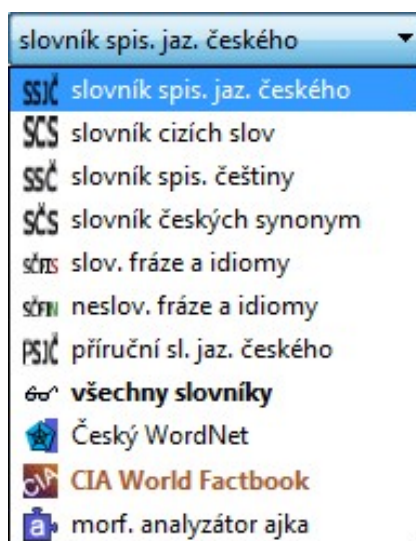


Obrázek 3: Hlavní strana programu

2.3 Dotaz

2.3.1 Volba elektronických zdrojů

Při přihlášení je hned na začátku vybrán jako výchozí *Slovník spisovného jazyka českého*. Pro volbu jiného slovníku musíte rozkliknout rozbalovací nabídku v levém horním rohu a vybrat si jiný slovník, popřípadě jiný dostupný elektronický zdroj.

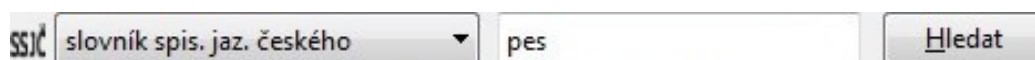


Obrázek 4: Výběr slovníku

2.3.2 Typ dotazu

Dotaz do dotazovacího řádku napište v základním tvaru, tedy jako byste jej hledali ve slovníku, tzn. jména budou v prvním pádě jednotného čísla a slovesa v infinitivu. To platí nejen pro slovníkovou část, ale i pro zbytek dostupných databází.

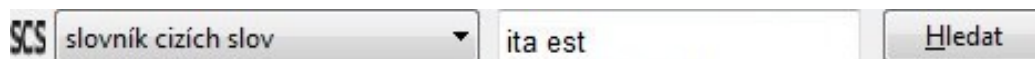
Veškeré slovníky (vyjma SCS), WordNet a morfologický analyzátor *ajka* mají zcela stejné vyhledávání. Například potřebujete-li vyhledat ve slovníku SSJČ význam slova *pes*, napište hledané slovo do dotazovacího řádku následovně:



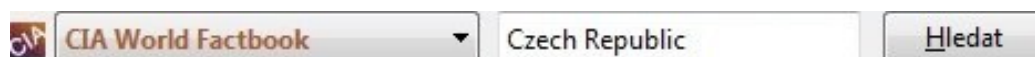
Obrázek 5: Ukázka dotazu v SSJČ

Při výběru *Slovníku cizích slov* i *CIA World Factbooku* bude dotazované slovo taktéž v prvním pádě čísla jednotného. Navíc výraz musí být v příslušném jazyce,

příčemž země hledáte pouze v angličtině (protože se jedná o anglický program). U SCS jako příklad vyzkoušíte vyhledat latinské sousloví *ita est* a u *Factbooku* Českou republiku pod anglickým názvem *Czech Republic*.



Obrázek 6: Ukázka dotazu v SCS



Obrázek 7: Ukázka dotazu v CIA World Factbook

Poté klikněte na políčko *Hledat* a aplikace pošle dotaz serveru¹⁴, který vám na výstup zobrazí hledaný výraz. U všech zdrojů probíhá vyhledávání stejně.

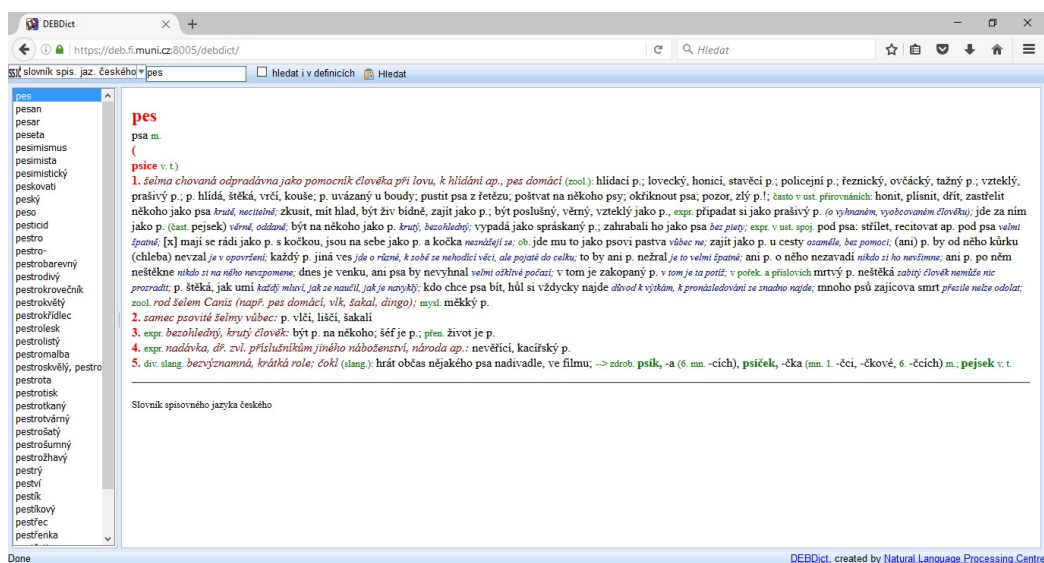
2.4 Výstup

2.4.1 Výstup hledání ve slovnících

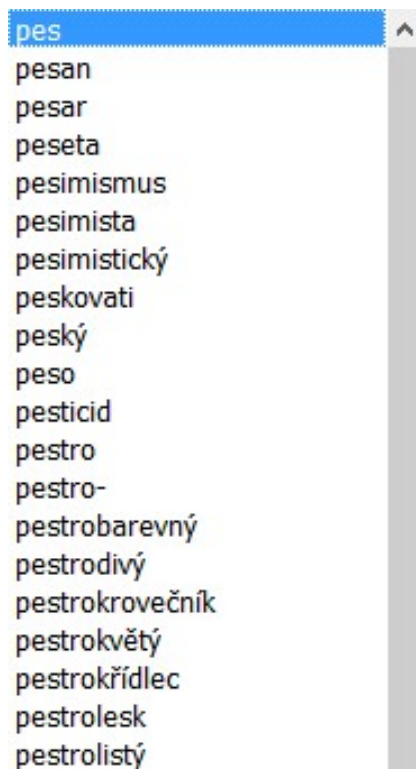
Jednotlivé slovníky

Výstup ve slovníkové části vypadá totožně u kteréhokoli slovníku vyjma obsahu. Obrázek č. 8 ukazuje pohled na celou obrazovku. Pod nabídkou se slovníky je obdélníkové pole se seznamem nalezených slov, která mají na začátku výrazu hledaný pojem. (Se seznamem se v doplňkové aplikaci mohlo dále pracovat, uložit a následně provádět analýzy. Ve webové aplikaci to již není možné.) Zbytek místa je určen pro definice ze slovníku.

¹⁴Každý slovník má jinou strukturu a server za vás vyřeší, pro který slovník je slovo určeno. Proto si nemusíte pamatovat jednotlivá dotazování ke každému slovníku.



Obrázek 8: Výstup vyhledávání slova *pes* v SSJČ



Obrázek 9: Seznam slov v levém poli

pes

psa m.

(

psice v. t.)

1. *šelma chovaná odpradávná jako pomocník člověka při lovu, k hlídání ap., pes domácí (zool.):* hlídací p.; lovecký, honiči, stavěcí p.; policejní p.; řeznický, ovčácký, tažný p.; vztekdlý, prašivý p.; p. hlídá, štěká, vrčí, kouše; p. uvázaný u boudy; pustit psa z řetězu; poštvat na někoho psy; okřiknout psa; pozor, zlý p.!, často v ust. přirovnáních: honit, plísňit, dít, zastřelit někoho jako psa *krutě, nechtělně*; zkusit, mít hlad, být živ bičně, zajít jako p.; být poslušný, věrný, vztekdlý jako p., *expr. připadat si jako prašivý p. (o vyhlazeném, vyobcovaném člověku)*; jde za ním jako p. (*čast. pejsek*) *věrně, oddaně*; být na někoho jako p. *krutě, bezohledně*; vypadá jako spráskaný p.; zahrabali ho jako psa *bez piety*; *expr. v ust. spoj. pod psa: stílet, recitovat ap. pod psa velmi špatně*; [x] mají se rádi jako p. s kočkou, jsou na sebe jako p. a kočka *nasnájejí se*; ob. jde mu to jako psovi *pastva vůbec ne*; zajít jako p. u cesty *osaměle, bez pomoci*; (ani) p. by od něho klůrku (chleba) *nevzal je v opovržení*; každý p. jiná ves *jde o různé, k sobě se nehodící věci, ale pojeté do celku*; to by ani p. *nežral je to velmi špatně*; ani p. o něho *nezavádí někdo si ho nevíme*; ani p. po něm *neštěkne někdo si na něho nevzpomene*; dnes je venku, ani psa by *nevyhnal velmi ošklivé počasí*; v tom je *zakopaný p. v tom je ta potě*; v pořek. a příslovích *mrtvý p. neštěká zabíjí člověk nemůže nic prozradit*; p. štěká, jak umí *každý mluví, jak se naučil, jak je navyklý*; kdo chce psa bit, hůl si *vždycky najde důvod k výtčám, k pronásledování se snadno najde*; mnoho psů *zajícova smrt přestě nelze odolat*; zool. *rod šelem Canis (např. pes domácí, vlk, šakal, dingo); myst. měkký p.*
2. *samec psovitě šelmy vůbec: p. vlčí, liščí, šakal*
3. *expr. bezohledný, krutý člověk: být p. na někoho; šéf je p.; přen. život je p.*
4. *expr. nadávka, dř. zvl. příštušníkům jiného náboženství, národa ap.: nevěřící, kacírský p.*
5. *div. slang. bezvýznamná, krátká role; čokl (slang): hrát občas nějakého psa nadivadle, ve filmu; -> zdrob. psík, -a (6. mn. -cích), psiček, -čka (mn. 1. -čci, -čkové, 6. -čcích) m.; pejsek v. t.*

Slovník spisovného jazyka českého

Obrázek 10: Definice slova *pes*

ita est

(1) kniž. tak jest, souhlasí, ano

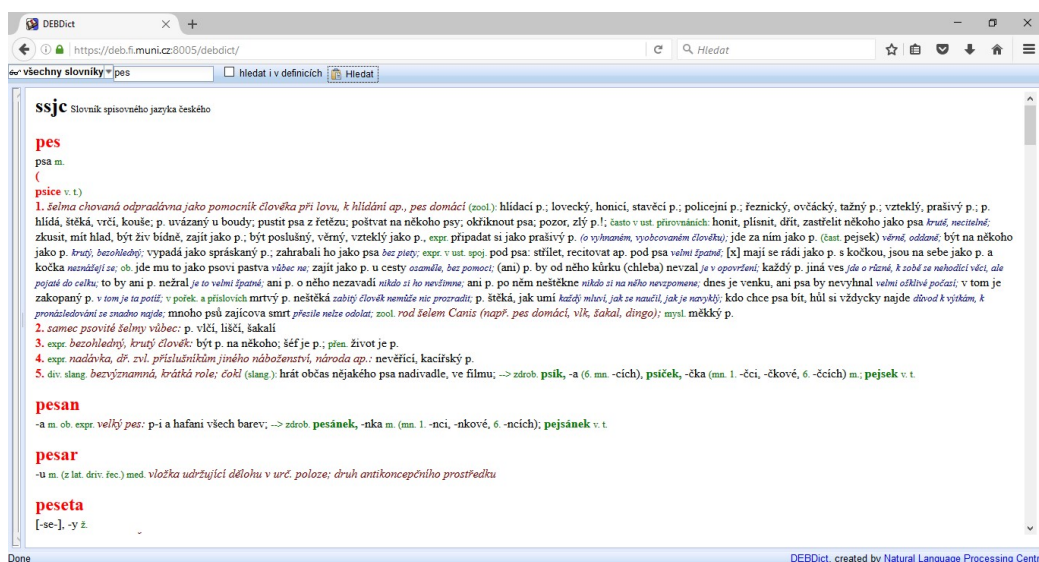
Slovník cizích slov

Obrázek 11: Definice latinského sousloví *ita est*

Všechny slovníky

U zvolení nabídky *všechny slovníky* server vypíše ze všech slovníků a českého WordNetu nejen hledaný výraz, ale i další slova, která v jednotlivých slovnících naleznete v seznamu na levém okraji.

Např. jste hledali význam slova *pes* ve všech slovnících a na výstup jste dostali u SSJČ kromě hledaného výrazu i pojmy jako *pesan*, *pesar*, *peseta* a *pesík*. Aplikace nevytvořila seznam, místo toho je vypsala do stejné úrovně s hledaným slovem. Pokud je v některém slovníku umožněn výběr z několika slov utvořených z hledaného výrazu, dostanete na výčet pouhých pět.



Obrázek 12: Výstup vyhledávacího slova *pes* ve všech slovnících

2.4.2 Výstup hledání v dalších dostupných zdrojích

Český WordNet

Výstup WordNetu je poněkud jiný než výstup slovníkové části. Nedostanete klasickou definici hledaného pojmu. Na obrázku č. 13 můžete vidět výsledek hledání v českém WordNetu.

Nahoře máte síť slov, kterou si můžete představit jako strom synsetů, v jehož kořeni je nejobecnější pojem a dále v listech ty nejspecifičtější (Pala a Ševeček, 1999). Při vyhledávání slova *pes* je nejspecifičtější pojem *psovitá šelma:1* a přes jednotlivé výrazy se postupně dostáváte k nejobecnějšímu pojmu *entita:1*. Síť slov může obsahovat i synonymní synset. Obrázek č. 13 obsahuje synonymní synset

organismus:1, bytost:1, forma života:1, což znamená, že *organismus* ve významu *1* je synonymní s výrazem *bytost* ve významu *1* i s *formou života* ve významu *1*.

Pod sítí slov jsou obecné informace ke slovu, jako jsou synonyma a slovní druh. Novinkou pro WordNet bylo udělení jedinečného identifikačního čísla (ID), které sloužilo pro vytváření ekvivalencí mezi synsety cizích jazyků. ID se začalo používat až ve WordNetu 1.5. Dalším vylepšením bylo rozdělení synsetů do sad základních pojmů (BCS – base concept set). Základní pojmy jsou vysoko v hierarchii a mají hodně potomků. Prvotní členění bylo na 3 skupiny. První skupinu tvořily subjekty vyznačující se podle způsobů conceptualizace (původ, forma, složení a funkce), druhá skupina se klasifikovala pomocí situačního typu a situační složky a třetí skupinou byly nepozorovatelné problémy, výroky či duševní subjekty (Vossen et al., 1998). Toto členění poprvé použili v EuroWordNetu a později v bylo BalkaNetu) doplněné o necelých 3 700 synsetů. Přesné rozčlenění naleznete na stránkách The Global WordNet Association (Weisscher, 2013). Rozšířené členění se používá v jakémkoliv cizojazyčném WordNetu.

Sémantickou relaci chápeme jako přiřazení hledaného výrazu k hyperonymii (slovo nadřazené) a popř. označuje i holonymii (vztah mezi slovem označujícím celek a slovem označujícím část celku). Hyperonymum k hledanému výrazu *pes* je *psovitá šelma* a holonymum je *smečka* a *rod Vlk* (latinsky *Canis*).

Po sémantické relaci následuje seznam hyponym (slova významově podřízená) a v tomto případě i meronym (části celku). Hyponyma jsou slova významově podřazená. V příkladu jste hledali slovo *pes*, jehož hyponyma jsou *puhl*, *mopslík* atd. Meronyma jsou části celku, slovo *prut* je zde míněno jako ocas, což je část každého psa. *Prut* je tedy část psa.

entita:1>objekt:1>živoucí entita:1>organismus:1, bytost:1, forma života:1>tvor:1, zvíře:1,
živočich:1>strunatec:1>obratlovec:1>savec:1>placentál:1>masožravec:1>psovitá šelma:1

Synonyma: pes:1, psice:1

Slovní druh: **podstatné jméno (n)**

ID: ENG20-02001223-n

BCS: 3

Sémantické relace:

-->> [hypernym] psovitá šelma:1

-->> [holo_member] rod Vlk:1, rod Canis:1

-->> [holo_member] smečka:1

hyponym: 10

<<-- [mero_part] prut:1

<<-- [hyponym] mopslík:1

<<-- [hyponym] hafan:1

<<-- [hyponym] špic:1

<<-- [hyponym] puď:1

<<-- [hyponym] trpasličí pes:1, trpasličí psík:1

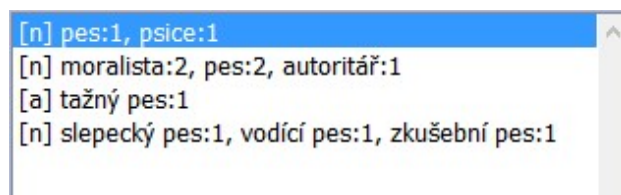
<<-- [hyponym] lovecký pes:1

<<-- [hyponym] pracovní pes:1

<<-- [hyponym] psisko:1, čokl:1

Obrázek 13: Výstup z českého WordNetu

Seznam v levé části nabízí všechny synsety mající hledaný výraz. Tento seznam má na prvním místě zmíněno, o jaký slovní druh se jedná, a není členěn podle sítí slov, nýbrž podle synonym.



Obrázek 14: Seznam v levém poli

Morfologický analyzátor *ajka*

Ajka umožňuje segmentovat slova. Tato funkce je efektivní jak při morfologické analýze, tak při generování správných gramatických tvarů. Výsledek vyhledávání se zobrazí v tabulce, kterou můžete dále rozkliknout na segmentaci slova, kategorii slova a skloňování (značené *).

Při hledání slova *pes* morfologický analyzátor na výstupu nabídl kromě hledaného výrazu i slovo *peso*. Důvod tohoto výběru je popsán níže v odstavci zabývajícím se skloňováním.

Analyzovaný tvar: pes			
Základní tvar	Segmentace	Číslo vzoru	Kategorie
pes (*)	=p=es==	1136-pes	k1gMnSc1
peso (*)	=pes===	232-barvivo	k1gNnPc2

Obrázek 15: Výstup v morfologickém analyzátoru *ajka*

Segmentace výrazu se dělí na prefix, kmenový základ, intersegment, koncovku a postfix. U základního tvaru je segmentace p-es, kde prefix, koncovka a postfix je nulový.¹⁵

¹⁵Kmenový základ je ten, který zůstává u všech tvarů stejný. Např. u slova *pes* jsou skloňované tvary: *pes*, *psa*, *psovi*, ..., proto kmenový základ je pouze *p-*, které nezměnilo svoji pozici v žádném skloňování.

Popis segmentace

=p=es==

Prefix	Kmenový základ	Intersegment	Koncovka	Postfix
-	p	es	-	-

Segmentuj tvar:

[Morfologická analýza - interaktivní režim](#) [Morfologická analýza - dávkový režim](#)

Obrázek 16: Segmentace slova *pes*

Kategorie je složena ze zkratk gramatických kategorií (GK). Při rozkliknutí dostanete podrobný popis kategorie, kde zkratky (nazvané *Hodnoty*) jsou podrobněji popsány. Malé písmeno vždy značí GK a velké písmeno, popřípadě číslo, určuje realizaci GK. Zkratky jsou použity kvůli úspornosti a přehlednosti. Kdyby tabulka obsahovala celistvý výpis gramatických kategorií, nebyla by přehledná.

Jak můžete vidět na obrázku č. 17, podrobný popis zkratky *k1* je následující. Písmeno *k* označuje GK slovní druh a *1* je realizací prvního slovního druhu, a to podstatného jména (substantivum). Totéž platí i u nečíselné hodnoty *nS*, kde malé *n* přiřazuje slovo ke GK číslo a velké *S* realizuje singulár, tedy jednotné číslo.

Popis značky

k1gMnSc1

Hodnota	Gramatická kategorie	Popis
k1	Slovní druh	Substantiva
gM	Rod	Mužský životný
nS	Číslo	Singulár
c1	Pád	Nominativ

Popiš značku:

[Morfologická analýza - interaktivní režim](#) [Morfologická analýza - dávkový režim](#)

Obrázek 17: Značka slova *pes*

Na obrázku č. 18 a č. 19 je rozkliknuté skloňování obou slov, které demonstruje, proč analyzátor uvedl výběr mezi slovem *pes* a *peso*. Tvar *pes* se nachází jak v prvním pádě jednotného čísla u výrazu *pes*, tak i v druhém pádě množného čísla u slova *peso*. Tudíž pokud se hledaný tvar objeví i v některém odvozeném tvaru jiného slova, automaticky jej *ajka* vydá na výstup.

Odvozené tvary ke slovu "pes"

Rod: Mužský životný			
Pád	Singulár	Pád	Plurál
1	pes	1	psi, psové
2	psa	2	psů
3	psu, psovi	3	psům , psům
4	psa	4	psy
5	pse	5	psové , psi
6	psu, psovi	6	psech, psách
7	psem	7	psy, psama

Obrázek 18: Skloňování slova *pes*

Odvozené tvary ke slovu "peso"

Rod: Střední			
Pád	Singulár	Pád	Plurál
1	peso	1	pesa
2	pesa	2	pes
3	pesu	3	pesům
4	peso	4	pesa
5	peso	5	pesa
6	pesu	6	pesech, pesách
7	pesem	7	pesy, pesama

Obrázek 19: Skloňování slova *peso*

CIA World Factbook

Elektronický zdroj CIA slouží jako atlas informací k jednotlivým státům. Do dotazovacího panelu jste napsali *Czech Republic* (česky Česká republika) a server vyhledal veškeré známé informace o České republice. Jelikož se musíte ptát anglicky, jsou i vyhledaná data v angličtině. Kromě informací, které jsou popsány v kapitole 1.5, můžete vpravo nahoře vidět vlajku dané země a dole mapu, kde se stát geograficky nachází.

2.5 Rozšířené vyhledávání

U obecného prohlížeče slovníků máte dvě možnosti rozšířeného vyhledávání. První možnost umožňuje vyhledat přesně hledaný výraz a druhá použije retrográdní seřazení.

Volbu, kdy potřebujete vyhledat přesné znění hledaného výrazu, většinou použijete při hledání ve výběru *všechny slovníky*. Jak již bylo zmíněno, při této volbě se vám budou mezi hledaný pojem plést i další slova, která mají na začátku výrazu hledané spojení. Pokud se chcete těchto slov zbavit, stačí před slovo v dotazovacím řádku napsat !.

Jak můžete vidět, na obrázku č. 22 je vyhledané slovo *pes* ve *všech slovnících* bez rušivých elementů. Tento výraz můžeme použít i na jednotlivé slovníky, kde je výklad stejný, jen s malým rozdílem, a to tím, že v levém postranním sloupci bude vypsán pouze hledaný výraz.

Retrográdní řazení je takové řazení, v němž se slova neřadí podle abecedy od začátku, ale naopak od konce slova. Museli byste procházet celý slovník, a to by zabralo mnoho času. Ovšem v DEBDictu taková slova můžete snadno nalézt jednoduchým způsobem. Stačí před slovo napsat znaménko -.

Vyhledaná slova jsou *herpes*, *ostropes* a původně hledané slovo *pes*. U jednotlivých slovníků to funguje stejně jako u obyčejného vyhledávání, jen jsou slova na výstupu seřazena od konce.

Czech Republic

[map](#)



Introduction

Background Following the First World War, the closely related Czechs and Slovaks of the former Austro-Hungarian Empire merged to form Czechoslovakia. During the interwar years, the new country's leaders were frequently preoccupied with meeting the demands of other ethnic minorities within the republic, most notably the Sudeten Germans and the Ruthenians (Ukrainians). After World War II, a truncated Czechoslovakia fell within the Soviet sphere of influence. In 1968, an invasion by Warsaw Pact troops ended the efforts of the country's leaders to liberalize Communist party rule and create "socialism with a human face." Anti-Soviet demonstrations the following year ushered in a period of harsh repression. With the collapse of Soviet authority in 1989, Czechoslovakia regained its freedom through a peaceful "Velvet Revolution." On 1 January 1993, the country underwent a "velvet divorce" into its two national components, the Czech Republic and Slovakia. The Czech Republic joined NATO in 1999 and the European Union in 2004.

Geography

Location Central Europe, southeast of Germany
Geographic coordinates 49 45 N, 15 30 E
Map references Europe
Area total: 78,866 sq km
land: 77,276 sq km
water: 1,590 sq km
Area - comparative slightly smaller than South Carolina
Land boundaries total: 1,881 km
border countries: Austria 362 km, Germany 646 km, Poland 658 km, Slovakia 215 km
Coastline 0 km (landlocked)

Obrázek 20: Výstup z CIA World Factbook



Obrázek 21: Rozšířené vyhledávání přesného tvaru

SSC Slovník spisovné češtiny

pes

- psa m

1. *šelma ohočená k hlídání, lovu ap.* hlídač, lovecký, ovčácký pes, pozor, zlý pes!, pustit psa z řetězu,
2. *samoc psovitě šelmy liščí pes, je to pes, nebo fenka?;*
3. *expr. bezohledný, krutý člověk být (na někoho) pes, život je pes, [x] ani pes po něm nešetke, expr. je zapomenut; být na někoho jako pes, zlý; být poslušný, věrný jako pes, honit, zastítelit někoho jako psa, bez citu; každý pes jiná ves, (pořek.) (o nesourodých věcech); (ani) pes by od něho kúrku neval, expr. je v opovržení; pes, který štěká, nekouše, (pořek.); práce ap. pod psa, hanl. velmi špatná; vypadat jako spráskaný pes, být schliplý n. unavený; (hledat, kde je) zakopaný pes, hlavní potíž, překážka; (venku je,) že by ani psa nevyhnal, velmi špatně počasí;*

psice -e ž *fena (svno)*

psík -a

psíček -čka

pejssek -ska m (mn. 1. -ci, -kově, 6. -cich) zdrob. expr.

psjc Příruční slovník jazyka českého

pes, psa m. *domácí šelma chovaná člověkem k hlídání, k lovu a pod.* Za zavřenými vraty ozval se štěkot psa. Jir. Děti i dva velicí psové, všecko vyběhlo přede dveře, vítat babičku. Něm. Kmet staropražský, Přivítan jménem, musel třikráte nositi prašivého psa po celém náměstí. Pal. Tu se Pyskáček rozzlobil a pošval na něj domácího velkého psa. Erb. Dělati pány chodice s puškou na rameni a se psem stavěcím v patách. Baar. Potkal zuřivého ovčáckého psa. Herb. Při kramářských karách zaštěkal hlídač pes. Win. *Ustál sp.* Karabiny byly rezavé jako pes. Ros. Ale cakra, cakra to nádobí, zednický nádobí, mám já už taky hlavu, aby ji hodil psu *nestojí za nic.* Čap. Ch. Život je krušný, lidé v něm zkusi jako psi. Šmil. Na mou duši, pracovat budu jako pes, dostanu-li se tam. Klost. Mám hlíd jako pes *velký.* J. Šir. Ohlíží se jako mlsný pes. Prasek. Jako pes věrná [je Katka]. Jir. Pravice jeho se zdvihla a rozkazně ukázala ku dveřím. Poslušná jako pes zdvihla jsem se z pohovky. Ner. Všichni jej znali mírného, pokorného jako psa po výprasku. R. Svob. Bartoň jako zbitý pes šoural se ke dveřím *bázlivě.* Hol. Musíš pamatovat na sebe, jako pes živ být nemůžeš *bídně.* Rais. Zahynu tu někde jako pes, sipal jsem, skřípaje zuby *bídně, bez pomoci.* Svob. Zabit ho chtěli jako psa. Jir. Zastítelte mě jak psa, jestli věrným nebudu. Don. Řekněte mi, proč jsou ti šefové na člověka jako psi *krutí.* Mah. Což mohu jít, když na mne voláš jako na psa *hrubě, neslušně?* R. Svob. Lítáš celý den po závodě, od dělníka k dělníkovi, a honíš je do práce jako psy *krutě, necitelně.* Svob. Někdy si na něj také houknu, a to jde, jakby ho pes kousl, ani nemukne. Něm. Hube sem dostal jak

Obrázek 22: Výstup vyhledávání přesného tvaru



Obrázek 23: Retrogradní řazení

herpes

-etu m. (6. j. -u) (z řec.) med. opar

ostropes

-psu m. statná bylina s listy ukončenými silným ostnem a s nafialovělými květy; bot. rod *Onopordon*: o. trubil

pes

psa m.

(

psice v. t.)

1. šelma chovaná odpradávná jako pomocník člověka při lovu, k hlídání ap., pes domácí (zool.): hlídací p.; lovecký, honič, stavěcí p.; policejní p.; řeznický, ovčácký, tažný p.; vztekly, prašivý p.; p. hlídá, štěká, vrčí, kouše; p. uvázaný u boudy; pustit psa z řetězu; poštvat na někoho psy; okřiknout psa; pozor, zlý p.!, často v ust. přirovnáních: honit, plisnit, dřít, zastřelit někoho jako psa *krutý, necitelný, zkusit, mít hlad, být živ bídně, zajít jako p.*; být poslušný, věrný, vztekly jako p., expr. připadat si jako prašivý p. (o vyřknutém, vybočovaném člověku); jde za ním jako p. (čast. pejsek) *věrné, oddané*; být na někoho jako p. *krutý, bezohledný*; vypadá jako spráskaný p.; zahrabali ho jako psa *bez písky*; expr. v ust. spoj. pod psa; střílet, recitovat ap. pod psa *velmi špatně*; [x] mají se rádi jako p. s kočkou, jsou na sebe jako p. a kočka *neznátejší se*; ob. jde mu to jako psovi pastva *vábeň ne*; zajít jako p. o cesty *osaměle, bez pomoci*; (ani) p. by od něho kůrku (chleba) nevezal *je v opovržení*; každý p. jiná ves jde o řízný, k sobě se nehodí věci, ale *pojaté do celku*; to by ani p. nežral *je to velmi špatné*; ani p. o něho nezavadí *nikdo si ho nevšímá*; ani p. po něm neštěkne *nikdo si na něho nevzpomene*; dnes je venku, ani psa by nevyhnal *velmi ošklivý počasí*; v tom je zakopaný p. *v tom je ta potřeba*; v pořek. a příslovích mrtvý p. neštěká *zabitý člověk nemůže nic prozradit*; p. štěká, jak umí *každý mluví, jak se naučil, jak je navyklý*; kdo chce psa bít, hůl si vžďycky najde *důvod k výtěm, k přeháňedování se snadno najde*; mnoho psů zajícova smrt *přesille nelze odolat*; zool. rod šelem *Canis* (např. pes domácí, vlk, šakal, dingo); mysl. měkký p.

2. samec psovitě šelmy *vůbec*: p. vlčí, liščí, šakalí

3. expr. *bezohledný, krutý člověk*: být p. na někoho; šéf je p.; *přen.* život je p.

4. expr. *nadávka, dř.* zvl. příslušníkům jiného náboženství, národa ap.: nevěřící, kacírský p.

5. div. slang. *bevýznamná, krátká role*; *čokl* (slang): hrát občas nějakého psa nadvadle, ve filmu; --> zřáb. *psík*, -a (6. mn. -cích), *psiček*, -čka (mn. 1. -čci, -čkové, 6. -čcích) m.; *pejsek* v. t.

Obrázek 24: Výsledek retrográdního řazení