

NABÍDKOVÝ LIST

Ústav Českého národního korpusu
Filozofická fakulta UK v Praze
www.korpus.cz



Odborné zaměření

Český národní korpus je akademický projekt při FF UK, který se dlouhodobě soustředí na sběr, zpracování a zpřístupňování jazykových dat pro výzkum jazyka a komunikace. Primárním cílem jeho činnosti je vytvářet rozsáhlé a reprezentativní sbírky textů (tzv. jazykové korpusy), které jsou dostupné jazykovědcům i zájemcům z řad veřejnosti pro zkoumání češtiny v celém spektru jejích podob a forem. Na základě korpusů tak vznikají slovníky či gramatiky, ale také různé aplikace založené na analýze jazyka (spell-checkery, slovníkové databáze, rozpoznávače řeči, automatická extrakce terminologie, extrakce klíčových slov apod.). Sběr se soustředí na tyto oblasti:

- současný psaný jazyk (beletrie, odborná literatura, publicistika);
- současný neformální mluvený jazyk (komunikace v rodině nebo mezi přáteli), včetně propojení přepisu se zvukem;
- čeština v kontrastu s jinými jazyky (český text propojený s jeho překlady do různých jazyků);
- starší vývojové fáze jazyka (od 14. století po současnost).

Nabízené služby

Český národní korpus vyvíjí pro svoji činnost nástroje pro zpracování přirozeného jazyka, aplikace pro jazykovědný výzkum a provádí analýzy textů. Konkrétně jde o:

- automatickou **lemmatizaci textů**, tj. přiřazení slovníkového tvaru ke slovům v textu (např. *pes* k tvarům *pes*, *psa*, *psům* apod.);
- automatické **morfologické značkování**, tj. přiřazení informací o gramatických vlastnostech každému slovu v textu (např. slovní druh, pád, číslo, rod, osoba, způsob apod.);
- lingvistickou **analýzu textů** (např. frekvenční zastoupení jednotlivých slov spolu s jejich srovnáním s referenčním korpusem češtiny, analýza spojitelnosti zadaných slov atp.);
- vytváření **seznamů lexikálních jednotek** češtiny (na základě frekvence a/nebo v kombinaci s různými gramatickými či stylovými charakteristikami);
- zkoumání běžných **překladových řešení** a automaticky generované seznamy pravděpodobných **ekvivalentů v různých jazycích**.

Dosažené výsledky, reference a příklady spolupráce

Projekt Český národní korpus poskytuje svoje data a nástroje všem registrovaným uživatelům prostřednictvím portálu www.korpus.cz. Kromě obecných nástrojů na práci s rozsáhlými jazykovými korpusem (**KonText**), zde zájemce najde i nástroje specializované, např. **SyD** pro analýzu variant nebo **KWords** pro zjišťování klíčových (prominentních) slov v textech.

Český národní korpus úzce spolupracuje např. s nakladatelstvím Fraus na vytváření pomůcek pro výuku češtiny a pravidelně organizuje praktické semináře a workshopy zaměřené na práci s korpusem pro veřejnost i odborníky z řad učitelů, překladatelů nebo redaktorů.

Kontaktní osoba

Jméno: Václav Cvrček

Název instituce: Ústav Českého národního korpusu, FF UK

E-mail: vaclav.cvrcek@ff.cuni.cz

Tel.: (+420) 221 619 837

Web: www.korpus.cz



FILOZOFICKÁ FAKULTA
UNIVERZITY KARLOVY
V PRAZE

NABÍDKOVÝ LIST

Studia nových médií - Ústav informačních studií a knihovnictví, FF UK v Praze

Odborné zaměření

Information Science
New Media
Social Network Analysis
Geodata Analysis
Text Analysis
Social Media Marketing
Critical TypoGraphic Design
User Experience/Centered Design and Usability
Social media marketing
Digital PR
Data mining

Nabízené služby

měření prezentace společnosti v internetových komunikačních kanálech a následná analýzy jejich postavení v porovnání s konkurencí
navržení komunikační strategie v prostředí sociální sítě
školení a koučink zaměstnanců pro potřeby prezentace společnosti v nových médiích
audit využívání informačních technologií a optimalizace procesů
data mining firemních dat
výzkum definované cílové skupiny (např. klientů, určité věkové skupiny apod.)
výzkum konkrétního tématu
průzkumy trhu
zákaznická spokojenost
spokojenost zaměstnanců
analýza a návrhy marketingových strategií a komunikačních strategií
analýza tržních trendů v zájmových segmentech trhu
analýza a návrhy aktuálního bezpečnostního prostředí firmy ve vztahu k interním/externím procesům
vývoj prototypu, nového produktu - software, aplikace, databáze

Výčet není vyčerpávající, reprezentuje doménovou oblast znalostí. Rádi pro vás připravíme konkrétní nabídku na míru potřebám vaší firmy.

Dosažené výsledky, reference a příklady spolupráce

Z významných projektů v posledních letech, které jsou základem našeho odborného zaměření a související s námi nabízenými službami uvádíme například:

Znalostní báze pro obor organizace informací a znalostí (NAKI MK ČR).

Využití mezinárodních zkušeností s business ontologiemi pro návrh a vývoj platformy propojení nezaměstnaných s poskytovateli přímých nabídek práce prostřednictvím hyperlokálních služeb (OPLZZ ESF ČR).

Příběhy z dějin československého státu: výzkum a experimentální vývoj softwarových simulací pro výuku historie českých zemí ve 20. století (NAKI MK ČR).

Studium informační vědy a znalostního managementu v evropském kontextu (OPVK ESF ČR).

Z nejvýznamnějších publikací v posledních letech uvádíme například:

Research Information Systems Through The Linked Open Data Web.

Socializing on the Internet: Case Study of Internet Use Among University Students in the United Arab Emirates.

The Research and Development Information System of the Czech Republic.

Research Information Systems Through The Linked Open Data Web.

A National CRIS Infrastructure as the Cornerstone of Transparency in the Research Domain.

Introduction to management for information specialists.

Z vyvinutého softwaru uvádíme například:

Software Rullus, který slouží ke statistickým analýzám česky psaných textů a pro měření čitelnosti česky psaných textů. Software byl vyvinut pro použití pro výzkum měření přístupnosti.

Aplikace Corpus Viewer ukazující možnosti jazyka R pro analýzu textových korpusů. Aplikace běží v prostředí Shiny, které umožňuje spouštět R jako web server. Byla vytvořena například pro vypočet n-gramů.

HTML5 aplikace Dáme práci, která přímo propojuje uchazeče s poskytovateli práce skrze online aplikaci, kterou je možné používat na počítačích, tabletech i chytrých telefonech.

Simulace Evropa 2045, první výuková hra v České republice, simulátor Evropy na dalších několika desítkách let, která funguje přímo v prohlížeči s využitím technologie Flash.

Geolokační hra Veltrusy: Ostrov pokladů, která využívá chytrých telefonů s operačním systémem Android; vychází z konceptu m-learning a location based gaming (známému zejména díky geocachingu).

Kontaktní osoba

Jméno: Mgr. Jakub Fiala

Název instituce: Studia nových médií - Ústav informačních studií a knihovnictví, FF UK v Praze

E-mail: jakub.fiala@ff.cuni.cz

Tel.: +420 251 080 384

Web: <http://novamedia.ff.cuni.cz>

FF UK v Praze, Ústav jazyků a komunikace neslyšících



Odborné zaměření

Deaf Studies, komunikace neslyšících, vzdělávání neslyšících, tlumočení a překlady pro neslyšící

Nabízené služby

- *Konzultace k tématu komunikace neslyšících*
- *Překlady čeština-český znakový jazyk (např. webových stránek)*
- *Kurzy českého znakového jazyka*

Dosažené výsledky, reference a příklady spolupráce

Spolupráce s:

Pevnost – České centrum znakového jazyka

Česká komora tlumočnicků znakového jazyka

Institut Neslyšících pro specializované vzdělávání

Ústav translologie FF UK v Praze

Ústav českého národního korpusu FF UK v Praze

Ústav českého jazyka a teorie komunikace FF UK v Praze

Školy pro sluchově postižené – všechny v ČR, dvě z nich jsou naše fakultní školy – máme smlouvu o vzájemné spolupráci

Federace rodičů a přátel sluchově postižených

Pedagogická fakulta UK

Obor Výchovná dramatika pro neslyšící, Divadelní fakulta JAMU v Brně

Teiresiás – Středisko pro studenty se specifickými nároky MU v Brně

Asociace organizací neslyšících, nedoslýchavých a jejich přátel

Centrum zprostředkování tlumočnicků pro neslyšící

Post Bellum – Paměť národa

Česká unie neslyšících

Kontaktní osoba

Jméno: Mgr. Andrea Hudáková, Ph.D.

Název instituce: FF UK v Praze, Ústav jazyků a komunikace neslyšících

E-mail: andrea.hudakova@ff.cuni.cz

Tel.: (+420) 221 619 238 (sekretariát Ústavu jazyků a komunikace neslyšících)

Web: <http://ujkn.ff.cuni.cz/>