

Projekt Český národní korpus spravovaný Filozofickou fakultou UK se stal oficiálně uznaným K-centrem CLARIN

Síť CLARIN je evropskou výzkumnou infrastrukturou zaměřenou na jazykové zdroje a nástroje pro práci s nimi v humanitních a společenských vědách. Smyslem K-centra ČNK je poskytovat všem zájemcům informace, konzultace a technickou podporu v oblasti korpusové lingvistiky zaměřené na češtinu.

Oficiální udělení statutu K-centra je potvrzením kvalit Českého národního korpusu, o jehož všestranný rozvoj již od roku 1994 pečuje Ústav Českého národního korpusu FF UK. Hlavním cílem činnosti ústavu je mapování různých podob psané i mluvené češtiny prostřednictvím budování a zveřejňování jazykových korpusů. Kromě sběru dat se ÚČNK věnuje také rozvoji metodologie korpusové lingvistiky a vývoji uživatelských rozhraní pro práci s korpusy.

Projekt ČNK je od roku 2011 zařazen na tzv. Cestovní mapu ČR velkých výzkumných infrastruktur, jejímž cílem je integrovat tyto infrastruktury na národní úrovni a tím zefektivnit jejich využívání, financování a rozvoj. Zařazení na Cestovní mapu se projevilo výrazným posílením servisní složky ČNK, zejména spuštěním portálu korpus.cz společně s řadou prvků uživatelské podpory, jako je např. uživatelská Poradna či korpusová wiki.

V současné době má ČNK více než 6 500 registrovaných aktivních uživatelů z ČR i ze zahraničí a průměrný počet dotazů položených přes některé z rozhraní činí přes 2 500 denně. Na zdrojích ČNK vzniká řada vědeckých a studentských závěrečných prací, jen za rok 2017 jich projekt eviduje téměř 300.

Jazykový korpus je elektronický soubor autentických textů,

který je sestaven tak, aby co možná nejlépe odrážel celek jazyka. Korpus je zpravidla bohatě anotován, a to jak na úrovni celých textů (bibliografické údaje a další metadata), tak jednotlivých slov (lemmatizace, morfologická, příp. také syntaktická anotace). Korpusy slouží jako rozsáhlá datová základna pro empirický výzkum, zejména v lingvistice, ale i v dalších disciplínách jako literární věda, psychologie, sociologie či historie.

Petr Kukař, mluvčí FF UK