

NA INTERNETU JE DOSTUPNÝ KORPUS MEDIÁLNÍCH TEXTŮ O NÁRODNÍM PARKU ŠUMAVA

Na webové adrese <http://sumava-corpus.narra.eu> byl zveřejněn rozsáhlý korpus téměř 7000 textů týkajících se Šumavy, a zejména jejího národního parku. Snadno prohledávatelný soubor převážně mediálních textů z let 1883 – 2010 poskytne svým uživatelům – novinářům, studentům a ostatním zájemcům – dostupný a přehledný nástroj pro orientaci ve vývoji myšlení a psaní o Šumavě.

Na základě jednoduchého třídění pomocí kategorií „autor“, „respondent“, „zdroj“, „žánr“, „datum“, „titulek“ a za pomoci fulltextového vyhledávání může uživatel nejen rychle najít texty vyjadřující se k dílčímu tématu, ale také získat informace o jejich širším kontextu.

Centrálním korpusem je MediaBank, který sestává z textů vzniklých v letech 1991 až 2010. Převážná většina sebraných textů

pochází z českých tištěných médií nebo jejich internetových verzí. Korpus je doplněn přepisy televizních a rozhlasových zpráv a reportáží. MediaBank obsahuje necelých 6800 textů o rozsahu 2,9 miliónu slov.

Druhým (referenčním) korpusem je Pref-Bank, který je souborem předmluv, doslovů a částečně také kapitol týkajících se obecné charakteristiky pohoří vypsanych z průvodců, fotografických publikací, memoárů a dalších pramenů zabývajících se Šumavou. Sestává z téměř 80 textů vzniklých mezi roky 1883 a 2013.

Třetím (referenčním) korpusem je KlosterBank, který je souborem předmluv a doslovů k beletristickým dílům Karla Klostermanna. Celkem 19 textů odráží vývoj vnímání Šumavy od počátku 20. století.

Sumava-corpus by měl sloužit nejen jako snadno dostupný materiál pro vysokoškolské studenty věnující se (například ve svých diplomových pracích) analýze diskurzu, mediální komunikaci ap., ale také novinářům, kteří se za pomoci korpusu mohou rychle zorientovat v dílčích rysech dlouholeté debaty o Národním parku Šumava. Korpusy by však měly sloužit také těm, které jednoduše zajímá, jak se o Šumavě v minulosti psalo a mluvilo.

Michal L. Hořejší

