

Jazykové zdroje slovenčiny v rámci projektu CESAR

Adriána Žáková

Slovenský národný korpus, Jazykovedný ústav L. Štúra Slovenskej akadémie vied, Bratislava

V príspevku sa venujeme najnovším výsledkom v oblasti spracovania prirodzeného jazyka, počítačovej a korpusovej lingvistiky na Slovensku. Sústreďíme sa na prístupné jazykové zdroje a databázy vybudované s podporou projektu CESAR, ktorý značnou mierou prispieva k tvorbe jednotného digitálneho trhu v Európe. Predstavíme jazykovú bielu knihu pre slovenčinu, ktorá porovnáva stav jazykových nástrojov a technológií na Slovensku a zároveň sumarizuje úroveň podpory týchto nástrojov a technológií v Európe v štyroch kategóriách: strojový preklad, textová analýza, spracovanie reči a dostupnosť jazykových zdrojov.