

Cílem této práce je identifikovat a vhodně označit složené gramatické tvary v datech z projektu Universal Dependencies. Data z projektu Universal Dependencies jsou anotována na morfologické i syntaktické rovině, nicméně tyto anotace se vztahují pouze k jednotlivým slovům, složené tvary tedy nelze snadno vyhledat. Výstupem této práce je program, který na základě nastudovaných gramatických pravidel slovanských jazyků tyto složené gramatické tvary v datech slovanských jazyků objeví a označí. Dále program, jenž načte pravidla z konfiguračního souboru a na základě těchto pravidel v datech označí složené tvary. Konfigurační soubor je navržen ve formátu YAML, lze ho snadno upravovat v klasickém textovém editoru a je dostatečně jednoduchý na to, aby s ním dokázal bez větších obtíží pracovat i uživatel bez zkušeností s programováním.