

# Abstrakt

Cílem diplomové práce je podat na základě strojové analýzy publikační činnosti České akademie císaře Františka Josefa pro vědy, slovesnost a umění přehled témat, která byla předmětem bádání této instituce v letech 1890–1910. Práce tak doplňuje a rozšiřuje dosavadní stav poznání této problematiky, který se dosud omezoval na analýzy konkrétních publikací, oborů nebo vědních tříd akademie, ale nedovoloval pohled na oblasti zájmu akademie jako celku. Úvodní část práce představuje kontext vzniku České akademie, shrnuje dosavadní stav poznání témat, kterým se její badatelé věnovali a popisuje, jak lze digitalizované tištěné publikace akademie využít pro analýzu pomocí počítačích metod v rámci digitálních humanitních věd. Ve výzkumné části byly za účelem dosažení stanoveného cíle získány z Digitální knihovny Akademie věd ČR digitalizované svazky vědeckého časopisu *Rozpravy*, který představoval stěžejní publikační platformu všech tří vědeckých tříd České akademie. Získaná data byla upravena pomocí volně dostupných nástrojů a použita jako vstup pro tematické modelování provedené metodou LDA (Latentní Dirichletova alokace). Výsledkem práce je ucelený přehled 35 konkrétních témat, kterým se Česká akademie na počátku svého vzniku věnovala. Jsou popsány počty publikací spadajících do jednotlivých témat a jejich proměny v průběhu sledovaného období. Prezentována je také sémantická podobnost jednotlivých témat a jejich vztah k současné klasifikaci věd. Práce splňuje svůj cíl a umožňuje nový pohled na publikační činnosti České akademie. Krátce nastiňuje možné další směry výzkumu spočívající v rozšíření korpusu zkoumaných publikací anebo podrobnější evaluaci tematického modelu. Práce také ukazuje, že dokumenty dostupné v digitálních knihovnách lze použít k výzkumu v digitálních humanitních vědách, přestože toto použití není jejich primárním smyslem.

## Klíčová slova

České akademie císaře Františka Josefa pro vědy, slovesnost a umění; tematické modelování; LDA; digital humanities