

**MUNI**  
**LAW**

# **Projekt ‚Exaktní hodnocení aplikační relevance judikatury‘**

Jakub Harašta

# Motivace

- Základní popsání toho, jak se cituje a nastínění metody automatického zjišťování v průběhu studentského projektu podpořeného PrF MU
- Snaha zjistit, jaké jsou citační tendence v rozhodnutích vrcholných soudů ČR
- Lepší pochopení citační praxe jako cesta k lepší práci s právními informacemi
- Podklady (datasets) pro další kvalitativní/kvantitativní výzkumy

# Jednoduchý pohled

- Vytěžíme citace
- Uspořádáme dataset
- Zanalyzujeme dataset
- Zjistíme, jestli a jak lze data využít pro rešeršní činnost

# Složitější cesta

- Vytěžíme citace
  - Reference recognition pomocí ML/NLP: trénovací dataset, vytvoření modelu
  - Segmentace pomocí ML/NLP: trénovací dataset, vytvoření modelu
  - Pipeline (propojení obou technik): evaluace, vytěžení dat
- Uspořádáme dataset
  - Odstraňování falešných pozitiv, manuální odstraňování typických chyb
- Zanalyzujeme dataset
  - Vlastnosti, statistiky
- Zjistíme, jestli a jak lze data využít pro rešeršní činnost (co vlastně citace znamenají?)
  - Experimenty

# Výsledky projektu

- Dílčí texty popisující jednotlivá zjištění
- Zpřístupněné datasety obohacených i prostých textů (Novotná)
- Citační dataset (Harašta)
- Teorie citační analýzy (Smejkalová)
- Případové studie a individuální experimenty (Harašta)
  
- Aktuálně připravovaná monografie vyjde v nakl. Wolters Kluwer v příštím roce

# Tým

- Harašta, Novotná, Smejkalová, Šavelka
- Polčák
- Kasl, Míšek, Loutocký, Myška
- Kotková, Procházková, Pullmannová, Semenišín,  
Šejnová, Šimková, Vosinek, Zavadilová, Zibner

Tento projekt je podpořen Grantovou agenturou ČR v rámci projektu GA17-20645S.

[jakub.harasta@law.muni.cz](mailto:jakub.harasta@law.muni.cz)