

# Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

**Autor práce** Bc. Jakub Maroušek  
**Název práce** Automatická klasifikace smluv pro portál HlidacSmluv.cz  
**Rok odevzdání** 2020  
**Studijní program** Informatika      **Studijní obor** Softwarové systémy

**Autor posudku** Doc. Mgr. Martin Nečaský, Ph.D.      **Role** Vedoucí  
**Pracoviště** KSI

## Text posudku:

Práce se zabývala automatickou klasifikací smluv zveřejňovaných ve státním a veřejně dostupném registru smluv. V něm jsou dostupné všechny smlouvy uzavírané subjekty veřejné správy, na které se vztahují povinnosti v oblasti poskytování informací veřejnosti. V současnosti je v registru smluv velké množství smluv, ale možnosti práce se smlouvami jsou omezeny tím, že smlouvy lze vyhledávat pouze na úrovni základních metadat (název, zadavatel, dodavatel, apod.). Možnosti práce rozšíří klasifikace smluv do věcných kategorií. To ale není možné vzhledem k množství smluv dělat ručně. Tématem této práce je tedy nalezení vhodné automatické metody klasifikace do předem daného seznamu kategorií, které byly navrženy jako vstup práce zástupci portálu HlídačStátu.cz.

Zadání práce bylo řešitelem naplněno. Podařilo se otestovat několik existujících metod aplikovaných na problematiku klasifikace smluv a srovnat je. Na základě srovnání byla vybrána nejvhodnější metoda a bylo vytvořeno softwarové řešení, které je přímo použitelné na neziskovém portálu HlidacStatu.cz.

Vypracovaná textová část diplomové práce má všechny potřebné náležitosti diplomové práce. Autor analyzuje zadání problému a dostupné datové sady. Během řešení přišel na to, že jako vhodnou trénovací datovou sadu by mohl zvolit datovou sadu o veřejných zakázkách, která obsahuje zadávací dokumentaci, což jsou texty blízké textům smluv, a jsou již klasifikovány do potřebných kategorií. Provedl tudíž i analýzu těchto dat. Na základě analýzy navrhl řešení a několik možných způsobů automatické klasifikace textů smluv. Provedl pak jejich experimentální srovnání. Zhodnotil jeho výsledky, vybral nejvhodnější metodu a na jejím základě vytvořil softwarové řešení pro klasifikaci smluv pro portál HlídačStátu.cz.

V průběhu řešení práce řešitel průběžně komunikoval s vedoucím a reflektoval všechny jeho připomínky k textu i ke způsobu řešení práce.

**Práci doporučuji k obhajobě.**

**Práci navrhuji na zvláštní ocenění.**

*Pokud práci navrhuje na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).*

Práce má významný celospolečenský dopad. Výsledné softwarové řešení je nasazeno na portálu [hlidacstatu.cz](http://hlidacstatu.cz), kde klasifikuje smlouvy navrženým způsobem. Iniciativa Hlídač státu bude klasifikaci využívat i při komplexnějších analýzách prostoru veřejných smluv, které bude nabízet veřejnosti.

**Datum** 2. September 2020

**Podpis**