

## Ontology Enrichment Based on Unstructured Text Data

**Autorka práce:** Bc. Ivana Lukšová

**Vedoucí práce:** Martin Nečaský, Ph.D.

Předložená diplomová práce řeší problematiku zpracování nestrukturovaných textů. Konkrétně se zaměřuje na extrakci slovníku pojmů, resp. ontologie (tj. slovníku doplněného o vazby mezi pojmy) z předložené množiny nestrukturovaných dokumentů (nestrukturovaných ve smyslu nepřítomnosti značek v dokumentu, např. XML, označujících strukturu dokumentu nebo sémantiku jeho částí). Tato problematika je v současné době velmi aktuální, neboť uspokojivá řešení neexistují a přitom informací zakódovaných v nestrukturovaných textech je velké množství. Autorka se zaměřila na využití metod strojového učení.

Autorka pracovala po celou dobu velmi samostatně a docházela na pravidelné konzultace s vedoucím. Taktéž aktivně konzultovala s Barborou Vidovou-Hladkou z Ústavu formální a aplikované lingvistiky. Cíle práce se podařilo splnit, včetně experimentální implementace a experimentu s pracovními nabídkami.

Jediným nedostatkem práce je množství dokumentů použitých při experimentu – pro plné ověření by bylo nutné zpracovat minimálně o jeden řád dokumentů více. Bohužel je tato činnost časově velmi náročná a nebylo tak v časových možnostech autorky takový experiment provést. Takový experiment přesahuje rozsah diplomové práce, vyžadoval by dlouhodobější výzkum v rámci Ph.D. studia.

Text práce je psán anglicky na dobré úrovni. Připomínky vedoucího práce k textu byly průběžně autorkou odstraňovány.

Práce splnila zadání a požadavky kladené na diplomovou práci na MFF UK. Proto ji **doporučuji k obhajobě.**

V Praze dne 20.5.2013

Martin Nečaský, Ph.D.