

## Posudek oponenta diplomové práce

*Název diplomové práce:* Ontology Enrichment Based on Unstructured Text Data

*Autor diplomové práce:* Bc. Ivana Lukšová

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

Diplomová práce se věnuje tématu obohacování ontologií na základě nestrukturovaných textů. Jako testovací data byly vybrány pracovní nabídky z oblasti IT. Téma je moderní a dobře zapadá do současného výzkumu v oblasti sémantických technologií.

Diplomová práce je dobře strukturovaná. Autorka se v úvodní části (Kapitola 1) věnuje představení obecných pojmů jako je ontologie apod. Následující část (Kapitola 2) se již věnuje zavedení problému obohacování ontologií pomocí strojového učení (klasifikátory apod.). Pojmy jsou zavedeny dostatečně formálně, což oceňuji. Kapitola 3 shrnuje přístupy k tématu obohacování ontologií, výčet existujících přístupů je dostatečný a svědčí o dobré rešeršní práci. V kapitole 4 autorka popisuje problém, který by měla řešit, a v kapitole 5 se věnuje návrhu řešení. Kapitola 6 popisuje techniky řešení, kapitola 7 implementaci a kapitola 8 závěr.

Autorka k řešení předloženého problému využívá existujících metod zpracování přirozeného jazyka (nástroj TreeX z ÚFAL MFF UK) a strojového učení (klasifikátor Weka), pro něž navrhuje vlastní výběr atributů a jejichž výpočet tvoří důležitou část implementace. Anotovaný vzorek textů pro testování byl relativně malý (32 nabídek práce), nicméně je nutné přihlídnout ke značné časové náročnosti jeho tvorby. Implementace řešení je netriviální a její funkčnost jsem společně s autorkou ověřil na datech mimo testovací vzorek.

Práce je napsaná anglicky a je stylisticky na průměrné úrovni. Vyskytuje se v ní nepříliš velké množství překlepů, které nemají zásadní vliv na porozumění textu. Nicméně by je bylo možné pečlivější kontrolou odstranit. Výtku ovšem zaslouží chybějící odkazy na obrázky z textu.

Práce je ve shodě se zadáním a splňuje požadavky kladené na diplomovou práci a proto ji **doporučuji k obhajobě**.

V Praze dne 17. 5. 2013

Mgr. Jakub Kozák