

Švartová Ildikó mác Posudek docentské habilitační práce

Název práce: Ontology Tool Benchmarks and Their Construction

Autor / uchazeč: Ing. Ondřej Zamazal

Oponent: Doc. Ing. Zdeněk Kouba, CSc.

Předložená habilitační práce se zabývá úlohou vytváření standardizovaných ontologických korpusů, tzv. ontology benchmarks, pro porovnávání nástrojů pro zpracování ontologií. Konkrétně se zaměřuje na rozbor ontologického korpusu OntoFarm, jehož je autor předložené práce správcem a hlavním autorem, a dále na návrh systému a metodiky pro automatické sestavování takových ontologických benchmarků.

Práce je psaná anglicky. Jako nerodilý mluvčí nejsem schopen důsledně posoudit její jazykovou úroveň. Z mého pohledu je, až na několik drobných gramatických prohřešků, dobrá. Co se týče srozumitelnosti předloženého textu, musím konstatovat, že se práce čte poměrně obtížně. Bylo by vhodné, kdyby autor na počátku práce stručně formuloval konkrétní vědecký cíl práce nejlépe formou hypotézy nebo výzkumného problému, na který se práce snaží odpovědět. Takovou jednoznačnou explicitní formulaci cíle v práci postrádám. Čtenář je nucen sám cíl během čtení práce hledat.

Další připomínka, kterou k práci mám, je, že autor předpokládá vysokou specializaci čtenáře. Například věta na str. 14 „*The profile OWL 2 DL corresponds to ontologies with SROIQ description logic semantic*“, je zcela zbytečná. Čtenář zabývající se sémantickými technologiemi to ví, čtenáři s méně specializovaným profilem věta nic neřekne, protože práce sémantiku SROIQ práce ani nedefinuje, ani neodkazuje na pramen obsahující její definici. Podobně obr. 3.1 na str. 61 obsahuje přehled počtu výskytů jednotlivých logických konstruktů v korpusu OntoFarm, aniž by jednotlivé konstrukty byly v práci vysvětleny. Tato informace je opět určena vysoce specializovanému čtenáři. Obě tyto připomínky by odpadly, kdyby práce obsahovala stručnou kapitolu poskytující přehled logických konstruktů a jejich příslušnost k jednotlivým sémantikám deskripčních logik.

Téma práce je velmi aktuální. Existuje celá sada nástrojů pro zpracování ontologií, pro jejichž porovnání a volbu nástroje vhodného pro daný účel je zcela nezbytné mít k dispozici standardizované sady ontologií. Vědecká komunita se jejich sestavováním systematicky zabývá. Problému se věnuje iniciativa OAEI (*Ontology Alignment Evaluation Initiative*), která má kromě intenzivní spolupráce v kyberprostoru i pravidelná každoroční zasedání v rámci mezinárodní konference ISWC (*International Semantic Web Conference*). Aktivita týmu VŠE a jmenovitě autora předložené práce v této komunitě je nepřehlédnutelná – viz stránky iniciativy <http://oaei.ontologymatching.org>.

Oceňuji podrobnou rešerši využití ontologického korpusu OntoFarm v praxi, kterou autor zpracoval v kapitole 3.4. Nejenom, že z ní vyplývá, že výsledek autorova výzkumu je využíván a citován mezinárodní vědeckou komunitou, podrobný rozbor, pro jaké typy úloh je korpus využíván (str. 71), a rozbor dat získaných od

uživatelů OntoFarm na základě autorova dotazníku jsou základem pro další cílený rozvoj korpusu OntoFarm.

Druhým z autorových vědeckých výsledků prezentovaných v habilitační práci je nástroj pro automatickou tvorbu ontologických benchmarků nazvaný OOSP (*Online Ontology Set Picker*), který umožňuje ohodnotit ontologie pomocí množiny metrik a vybrat do benchmarku nevhodnější z nich. Zde bych chtěl vyzvednout, že autor propojil svůj výzkum se svým pedagogickým působením, když zapojil studenty bakalářského programu *Umělá inteligence a reprezentace znalostí* do testování nástroje OOSP.

Na uchazeče mám následující otázky:

Uchazeč v kapitole 1 zavádí kategorizaci ontologických nástrojů. V tabulce 1.1 zavádí pět kategorií ontologických nástrojů, které se navzájem liší podmnožinami tzv. aktivit, přičemž celkem autor zavádí osm různých aktivit. Rád bych, aby uchazeč v rozpravě vysvětlil, jak dospěl k výběru těchto osmi aktivit. Proč mezi ně nezařadil např. sémantické porovnání ontologií (diff), popř. verzování ontologií.

V sekci 4.2.2 uvádí „přístup založený na exaktních intervalech metrik“. V práci jsem nepostřehnul, o které konkrétní metriky se jedná. Mohl by uchazeč toto vysvětlit?

Dále by bylo vhodné, aby uchazeč v rozpravě vysvětlil podrobně svou osobní roli správce (viz „maintainer“ na str. 5) korpusu OntoFarm.

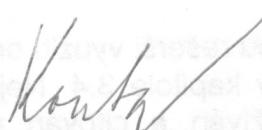
Závěr:

Konstatuji, že předložená habilitační práce obsahuje dva nové vědecké výsledky – (i) ontologický korpus OntoFarm a (ii) nástroj OOSP a související metodiku automatického sestavování ontologických benchmarků. Splňuje tak požadavek §72, odst. 1, písmeno a).

Chci zdůraznit autorovo aktivní působení v mezinárodní vědecké komunitě, zejména jeho zapojení do aktivity OAEI. Třebaže nemám k dispozici seznam uchazečových publikací, ze seznamu literatury uvedeného v habilitační práci vyplývá, že publikoval celou řadu prací a v řadě případů se zahraničními spoluautory, což dále dokládá jeho úspěšné zapojení do mezinárodní vědecké komunity. Přes moje výhrady ke srozumitelnosti textu práce, které zmiňuji v úvodní části posudku, konstatuji, že autor během svého dosavadního pedagogického působení prokázal, že umí vhodně propojit svůj výzkum s výukou.

Z výše uvedených důvodů doporučuji předloženou habilitační práci k obhajobě.

V Praze dne 21.ledna 2019



Doc. Ing. Zdeněk Kouba, CSc.