

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Bc. Jan Valášek
Oponent práce: Ing. Michal Valenta, Ph.D.
Název práce: Efektivní dotazování nad databází PubChem
Obor: Znalostní inženýrství

Datum vytvoření: 27. 1. 2018

Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 5:
1. Náročnost a další komentář k zadání	1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
Komentář: Pro splnění zadání bylo nutné, aby se autor práce seznámil s velmi specifickou a komplikovanou doménou ukládání a dotazování bio-chemických dat, dále pak hybridním databázovým strojem Virtuoso a další implementací mapování realčních dat na RDF. Přípravná fáze práce byla tedy nadprůměrně obsáhlá. Další výzvou práce bylo množství dat, které způsobilo, že jednotlivé experimenty měly dlouhou přípravu i běh.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
2. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
Komentář: Všechny dílčí úkoly zadání jsou v práci splněny.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
3. Rozsah písemné zprávy	1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
Komentář: Rozsah textu odpovídá požadavkům na diplomovou práci. Text neobsahuje žádné nadbytečné části.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
4. Věcná a logická úroveň práce	70 (C)
Popis kritéria: Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře.	
Komentář: Práce je dobře strukturovaná. Samotný text se již tak dobře nečte. V práci nejsou nepravdivá tvrzení, některé pojmy jsou však zavedeny nepřesně nebo nevhodně vysvětleny. Občas dochází při výkladu k promíchání pojmů, které spolu nesouvisí - například relační databáze a IRI. Tyto nedostatky se týkají zejména úvodní části a první kapitoly práce. V úvodní části je pak uvedeno několik schématických obrázků, které nejsou v textu dostatečně vysvětleny. Od druhé kapitoly, kde autor popisuje vlastní postupy a výsledky se text již čte o poznání lépe.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
5. Formální úroveň práce	80 (B)
Popis kritéria: Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3.	

Komentář:

V textu se občas vyskytují gramatické prohřešky - typicky nesoulad v rodech při skloňování. Místy utíká formátování při zalamování řádků, některé obrázky přetékají textovou oblast, je to ovšem zjevně kvůli lepší čitelnosti obrázků, což nepovažuji za chybu.

Text práce též obsahuje větší množství anglicismů, které autor skloňuje, což působí poněkud nevhodně: "počty joinů" apod.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

6. Práce se zdroji

95 (A)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a uvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:

Jedná se o implementační práci. 24 citovaných zdrojů považuji proto za zcela dostatečné.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

7. Hodnocení výsledků, publikační výstupy a ocenění

90 (A)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:

Přínos práce spočívá v tom, že skutečně prozkoumala současné možnosti pro ukládání a dotazování biochemických dat.

Testování bylo provedeno na reálných datech.

Výsledkem práce je zjištění, že pro specifické potřeby dotazování není dosud používaný systém Virtuoso úplně vhodným.

Podobné výsledky přineslo i testování alternativního řešení - vrstvy D2RQ.

Dosažené výsledky mají hodnotu jednak pro zdůvodnění vlastního řešení, na kterém UOCHB pracuje a jednak by bylo možné je publikovat jako určitý benchmark, vůči kterému by se mohly vztahovat budoucí verze nástrojů Virtuoso, D2RQ i vlastního řešení.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků**Popis kritéria:**

Uveďte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uveďte možnosti využití výsledků ZP v praxi.

Komentář:

Výsledky práce jsou přímo využitelné v praxi.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

9. Otázky k obhajobě**Popis kritéria:**

Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

Otázky:

1. Můžete upřesnit metodiku měření výsledků? Bylo měření několikrát opakováno nebo se jedná o jednorázové výsledky?

Vzhledem k reálné velikosti dat je i druhá varianta přípustná. Pro publikace výsledků dále by ovšem bylo lepší měření vícekrát opakovat a metodiku doplnit.

2. Vysvětlete podrobněji rozdíl mezi měřeními označenými Virtuoso s klazulí FROM a bez ní. Z textu práce jsem je příliš nepochopil.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

10. Celkové hodnocení

80 (B)

Popis kritéria:

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nesmí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Práce splnila zadání v dostatečném rozsahu i kvalitě. Autor odvedl velký kus implementační práce, která může být dále využita. Snížené hodnocení, které navrhuji, je způsobené nižší kvalitou textové části práce.

Podpis oponenta práce: