



## Posudek oponenta závěrečné práce

**Student:** Bc. Juraj Polačok  
**Oponent práce:** Ing. Michal Valenta, Ph.D.  
**Název práce:** Využití grafových databází pro pokročilou analýzu dat  
**Obor:** Znalostní inženýrství

**Datum vytvoření:** 2. 6. 2019

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – následující škálou 1 až 4:</i>
<b>1. Splnění zadání</b>	<b><u>1=zadání splněno,</u> 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<i>Popis kritéria:</i> Posuďte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.	
<i>Komentář:</i> Zadání považuji za splněné ve všech dílčích bodech.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>2. Písemná část práce</b>	<b>95 (A)</b>
<i>Popis kritéria:</i> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.	
<i>Komentář:</i> Jednotlivé kapitoly práce víceméně následují dílčí úkoly v zadání. Kapitoly na sebe logicky dobře navazují. Práce se dobře čte, po formální i typografické stránce mi přijde zcela v pořádku, gramatickou stránku nejsem schopen hodnotit (práce je napsaná slovensky). Všechny kapitoly jsou rozsahem i obsahem přiměřené. Nejpřínosnější částí práce je, dle mého soudu, kapitola věnovaná rešerši technologií pro zpracování dat, které jsou dobře horizontálně škálovatelné, a grafovým databázím, které lze distribuovat mezi více výpočetních uzlů. Tyto technologie zažívají v posledních letech velmi bouřlivý rozvoj, vzniká a zaniká mnoho projektů, neexistuje mnoho byt jen komunitou akceptovaných standardů a vládne terminologický chaos. O to více je třeba tuto část práce ocenit. Je napsaná velmi věcně a je konsistentní. Čtenář neznalý problematiky by s touto kapitolou zřejmě měl i tak problémy. Pro lepší orientaci by nejspíše pomohlo ještě hlubší členění v části 3.1. Další potíž spočívá v tom, že diskutované technologie se různě prolínají a navazují na sebe, takže je prakticky nemožné je seřadit lineárně do textu práce, proto je v této partii hodně zpětných a dopředných odkazů.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
<b>3. Nepísemná část, přílohy</b>	<b>90 (A)</b>
<i>Popis kritéria:</i> Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů	

**Komentář:**

Součástí závěrečné práce je též experiment - benchmark - provedený na zvolené technologii. Data pro benchmark jsou dobře popsána a vhodně zvolená.

Celá implementační část práce mi byla autorem prezentována živě, přičemž jsme rovněž diskutovali o konfiguraci celého experimentu.

Experiment považuji za vhodně nastavený, samotná konfigurace a orchestrace prostředí je velmi pracná a náročná. Autor ji zvládl výborně.

Skutečnost, že pro experiment byly použity existující implementace analýzy grafů, nijak nesnižuje výsledky práce ani není v rozporu se zadáním.

*Hodnotící kritérium:*

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

**4. Hodnocení výsledků, jejich využitelnost**

95 (A)

*Popis kritéria:*

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

**Komentář:**

Výsledky práce jsou, dle mého soudu, dobře využitelné. Autor samozřejmě mohl věnovat více prostoru rozboru výsledků. Já jsem však těžiště práce vnímal hlavně v diskusi a výběru existujících technologií a jejich implementaci. Rozsah práce i její pracnost je i v této podobě nad běžným průměrem.

*Hodnotící kritérium:*

*Způsob hodnocení – nehodnotí se*

**5. Otázky k obhajobě**

*Popis kritéria:*

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

**Otázky:**

V rámci benchmarku jste použil algoritmy Peer pressure clustering a PageRank. Měla by tato analýza nějakou speciální interpretaci v použité datové sadě (za předpokladu, že by byla reálná a nikoliv částečně syntecká)?

*Hodnotící kritérium:*

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

**6. Celkové hodnocení**

92 (A)

*Popis kritéria:*

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.

**Text hodnocení:**

Téma práce je aktuální, technologie, které bylo třeba nastudovat, konfigurovat a nasadit, jsou netriviální a většinou velmi špatně dokumentované. Autor odvedl výbornou práci zejména v rešeršní části. Na jejím základě správně zvolil, nastavil a realizoval experiment. Interpretaci experimentu by mohlo být věnováno více prostoru, ale to už je, myslím, mimo rozsah zadání. Práce splnila zadání ve vysoké kvalitě.

Podpis oponenta práce: