

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Jakub Tuček
Oponent práce: Ing. Michal Kusý
Název práce: InfoWeb - Nástroj získávání informací z webů
Obor: Softwarové inženýrství

Datum vytvoření: 11. 6. 2017

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: Tato práce navazuje již na 2 etapy vývoje aplikace v rámci předmětů BI-SP1 a BI-SP2. Zadání považuji za náročnější z následujících důvodů: 1) jedná se o implementační řešení a práci s cizím kódem, 2) součástí zadání je i nastavení a konfigurace infrastruktury pro dalšího rozvoj.	
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: Zadání se podařilo splnit. Nebyly vyřešeny všechny identifikované problémy aplikace, ale to v rámci hodinové dotace na BP není reálné.	
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: Menší výhradou je rozsah práce, který je dle mého názoru až příliš dlouhý. Je to způsobeno obsahem, který se netýká přímo zadání. Částečně dále rozvinuto v bodě 4.	
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: Věcná stránka je v pořádku. Vytkl bych až zbytečné zaměření se na týmový projekt jehož byl student členem. V principu bych v této práci kapitoly 2 až 4 shrnul v krátkosti pod kapitolou Analýza stávající řešení (AS IS). Tato kapitola by přiblížila pouze současný stav projektu a popsala procesy / stories, které má systém naplňovat a zhodnocení jejich funkcionality. Kapitolu 1.6 popis konkrétních existujících služeb, bych si s ohledem na stručnost představil ideálněji jako tabulku srovnání jednotlivých služeb.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
5. Formální úroveň práce	90 (A)

Popis kritéria:

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.

Komentář:

Práce obsahuje drobné překlepy, jinak nemám připomínky.

Hodnotící kritérium:

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

6. Práce se zdroji

75 (C)

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 úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:

Bohužel v seznamu literatury nenalezneme žádnou literaturu týkající se téma web data mining, na němž jsou tato práce a projekt založeny.

Hodnotící kritérium:

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

7. Hodnocení výsledků, publikační výstupy a ocenění

85 (B)

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:

Testováním pouze nad vzorkem 10 produktů a 22 webů byla potvrzena funkčnost, ale skutečné problémy by vyvstaly při stovkách produktů a webů. Např. první zřejmá optimalizace je rozřídění e-shopů dle oblasti trhu, tak aby se nehledalo na webech, které daný segment zboží vůbec nenabízí.

Ve shrnutí výsledků mne překvapilo, že tabulky obsahují ne co se v práci studentovi podařilo vyřešit, ale ty co dále zbývají. Bylo by hezké mít 2 sloupce pro srovnání před a po, tak aby byly hned zřejmé výsledky práce.

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:

Práce pomohla vylepšit a uzavřít proces již existující aplikaci InfoWeb pomocí praktik "čistého kódu" a . Díky zvýšení pokrytí aplikace testy je další rozvoj pod lepší kontrolou zda neovlivní funkcionalitu celé aplikace.

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 odřázkami).

Otázky:

Skutečně bude administrátor řešit všechny problémy a přiřazení? Nedospěl jste při analýze k přesvědčení, že je nutná další role s menší odpovědností, která by řešila obecné problémy?

Poměrně mne překvapilo, že DataProviderService obstarává jak stahování, tak parsování odpovědi. Tyto služby bývají často odděleny, protože samotné stažení stránky je triviální a lze řešit nízkouúrovňově (např. C kód s curl, případně využít AWS Lambda). Na druhou stranu párování s heuristikami může nabrat značné složitosti. Zaujalo mne, že při refactoringu jste k tomuto závěru nedošel. Proč?

Hodnotící kritérium:

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

10. Celkové hodnocení

85 (B)

Popis kritéria:

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

Text hodnocení:

Tato práce je převážně implementačního charakteru a moje připomínky jsou převážně k formě struktury práce. Zadání bylo splněno, nové části kódu jsou přehledné a vzorně okomentované.

Podpis oponenta práce: