



Posudek oponenta závěrečné práce

Student: Oleksandr Chmel
Oponent práce: Ing. Ivan Halaška
Název práce: Anonymizace osobních údajů pro testovací prostředí
Obor: Webové a softwarové inženýrství

Datum vytvoření: 15. 6. 2020

Hodnotící kritérium:	Způsob hodnocení – následující škálou 1 až 4:
1. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
<p><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.</p> <p><i>Komentář:</i> Zadání je členěno do 5 bodů: bod1: splněn, bod 2: splněn, bod 3: rešerše stávajících řešení je v pořádku, analýza funkčních a nefunkčních požadavků je zmatená a neúplná, bod 4. nelíbí se mi, že uživatelské testování provedl pouze autor programu, testovací zdroj dat má pouze 10 záznamů (řádků), jak se bude program chovat nad rozsáhlým zdrojem dat. bod 5: splněn.</p>	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	65 (D)
<p><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.</p> <p><i>Komentář:</i> Práce obsahuje 41 stran textu. Text se mi čte obtížně. Student i v obecných partiích neustále odbíhá k programátorským technikám. Text student pojal střídavě v osobním nebo neosobním módu. Cituji: " K testování jsem použil ...", "Tento problém byl vyřešen". Jakoby si autor nebyl jistý autorstvím toho, co ta která pasáž popisuje. V kap. 2.9 student používá termín "pseudonymizace", v kap. 4.1.1 student používá termín "pseudo-anonymizace". Chtělo by to ujednotit terminologii. Student nemá jasno v procesu identifikace funkčních a nefunkčních požadavků. Např. studentův funkční požadavek na modulární řešení programu je typický nefunkční požadavek. Postrádám funkční požadavky ohledně identifikace a struktury zdroje a cíle transformace, identifikace osobních a citlivých údajů ve zdroji a jak s nimi bude naloženo? Tomu pak odpovídá hodně nekonkrétní kapitola 6.3. Obrázky obsahují příliš malé texty, které nejsem schopen přečíst. Ani v pdf verzi dokumentu nejsem schopen (například pomocí Acrobat reader) si obrázek zvětšit do čitelné podoby. V kapitole Literatura je uvedeno 30 literárních zdrojů, které jsou v textu všechny řádně citované.</p>	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	65 (D)

Popis kritéria:

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:

Nemám představu, jaké má vytvořený program výkonové vlastnosti, student to neotestoval na rozsáhlejších souborech dat. Přílohy jsou zveřejněny na github.com, a jsou mi dostupné. Zdrojový text programu je úsporně komentován, nejsem si jistý tím, že to je dostatečný podklad pro rozšiřování programu v budoucnu.

Hodnotící kritérium:

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

4. Hodnocení výsledků, jejich využitelnost

65 (D)

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:

Analýzu problému student provedl nekvalitně. Testování prováděl pouze student sám. Osobně jsem si program nedokázal vyzkoušet. Chybí uživatelská dokumentace, konkrétně návod pro sestavení konfigurace, co má obsahovat, jakou má mít strukturu. Jak se bude program spouštět. Student se nezmiňuje o tom, zda byl program u zadavatele použit, či bude použit. Snad toto zhodnotí vedoucí práce, který pracuje ve firmě zadavatele.

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 odřádkami).

Otázky:

1. ČVUT používá jednak testovací instanci systému KOS (aplikace+databáze s testovacími daty), určenou pro uživatele dat v ostré databázi, jednak vývojovou instanci systému, kterou používají programátoři a testeři. Objem dat je v jednotkách terra. Dal by se zde použít váš program?
2. Plánuje zadavatel použití vašeho programu v konkrétních projektech? Případně byl již použit?

Hodnotící kritérium:

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

6. Celkové hodnocení

65 (D)

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:

Zadání posuzované práce patří k méně náročným. Student vytvořil program pro transformaci dat z ostré databáze do testovací, při které je aplikována anonymizace osobních a testovacích dat. Analýzu problému student provedl nekvalitně, textová část práce má průměrnou kvalitu. Program má dle zadání modulární strukturu, což oceňuji. K vytvořenému programu chybí uživatelská dokumentace. Uživatelské testování provedl pouze student sám. Výkonové testování neprovedl vůbec.
Není mi známo, zda výsledný program je pro zadavatele použitelný a zda jej plánuje využít.

Podpis oponenta práce: