



Posudek oponenta závěrečné práce

Student: David Skalský
Oponent práce: Ing. Zdeněk Rybola, Ph.D.
Název práce: Vyhledávání osobních údajů v relačních databázích
Obor: Webové a softwarové inženýrství

Datum vytvoření: 5. 6. 2018

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
Popis kritéria: 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.	
Komentář: Zadání je splněno s menšími výhradami, zejména z hlediska problematiky ochrany osobních údajů a jejich definicí, které jsou popsány velmi stručně. Taktéž provedené testování výsledného řešení nemá valnou vypovídací hodnotu.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	55 (E)
Popis kritéria: 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.	
Komentář: Písemná část práce působí nepřehledně a nesrozumitelně. Obsahuje především fakta a závěry, ale postrádá logické vysvětlení důvodů pro jednotlivé části, vysvětlení důvodů pro ochranu osobních údajů, bližší popis jednotlivých klíčových pojmů, vysvětlení pointy identifikace osobních a citlivých údajů. Jednotlivé kapitoly nejsou vhodně členěny (samostatná kapitola jen pro cíle práce, specifikace požadavků v rámci kapitoly Návrh řešení, stručný popis navrženého způsobu řešení v kapitole Implementace). V práci zcela chybí návrh architektury řešení a vnitřní struktury kódu. Chybí jakékoli grafické znázornění struktury i algoritmu ověřování existence osobních údajů v datech. Chybí vysvětlení způsobu vyhodnocení provedené analýzy dat. Popsané testování má velmi malou vypovídací hodnotu. Jednotlivé testy pokrývají minimum funkčnosti, manuální testování zcela postrádá specifikaci testovaných scénářů. Vyhodnocení testů je založené na velmi minimalistickém případě se zkrácenými předpoklady a neúplnými výsledky. Písemná část práce navíc obsahuje velké množství překlepů a hlavně gramatických chyb (především interpunkce).	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	60 (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: Součástí práce je jen velmi omezená část zdrojových kódů. Součástí nejsou žádné automatizované testy popisované v písemné části práce! Ve struktuře kódu se duplikují třídy pro vyhledávání ve strukturovaných datech a pro vyhledávání v binárních datech. Oproti požadavkům je implementována podpora jen pro PDF soubory. Zdrojové kódy nejsou komentované a obsahují "TODO" části.	

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
4. Hodnocení výsledků, jejich využitelnost	55 (E)
<i>Popis kritéria:</i> 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.	
<i>Komentář:</i> Výsledkem práce je funkční prototyp nových vyhledávacích tříd schopných vyhledávat několik nových typů osobních údajů v databázi a v binárních souborech uložených v databázi nebo formou reference. Vzhledem k nízké kvalitě kódu však nelze výsledek považovat za plně použitelný a může sloužit spíše jako základ pro další vývoj.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – nehodnotí se</i>
5. Otázky k obhajobě	
<i>Popis kritéria:</i> Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřázkami).	
<i>Otázky:</i> Jaký je smysl celé práce? K čemu je implementované řešení určeno a v čem může pomoci? Jaká je softwarová architektura řešení? Byly použity nějaké návrhové vzory pro optimalizaci kódu?	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
6. Celkové hodnocení	55 (E)
<i>Popis kritéria:</i> 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.	
<i>Text hodnocení:</i> Výsledkem práce je rozšíření existující aplikace pro vyhledávání osobních údajů a anonymizaci o vyhledávání nových typů osobních údajů v databázi a PDF souborech v databázi uložených. Kvalita kódu je však pouze na úrovni prototypu. Také písemná část práce má značné nedostatky. Vzhledem k výše uvedeným nedostatkům tak hodnotím práci stupněm E - dostatečně.	

Podpis oponenta práce: