



Posudek oponenta závěrečné práce

Student: Bc. Radomír Ort
Oponent práce: Ing. Zdeněk Rybola, Ph.D.
Název práce: Anonymizace osobních údajů v dokumentech
Obor: Webové a softwarové inženýrství

Datum vytvoření: 31. 8. 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
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: Cílem práce bylo analyzovat, navrhnout a implementovat anonymizaci souborů a dat ve strukturovaných souborech. Tohoto cíle se povedlo dosáhnout pouze pro některé z plánovaných způsobů anonymizace a pouze některé typy souborů.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	80 (B)
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 je na dobré úrovni. Struktura práce je přehledná a obsah dobře pochopitelný. Text je místy velmi podrobný, bohatě popisující každé rozhodnutí a problém, každou část řešení i způsobu jeho ověření. Práce popisuje různé alternativy řešení anonymizačního modelu a srovnává jejich výhody a nevýhody, popisuje podrobně architekturu řešení a jednotlivé komponenty, balíčky a třídy. Na druhou stranu práce často v detailních popisech opakuje již několikrát uvedené informace, text se tak protahuje a nabývá na hutnosti, takže je těžko stravitelný. V mnoha místech by pomohlo informace vizualizovat formou diagramu. Mezi největší nedostatky písemné práce patří chybné použití třídních diagramů v návrhu. Diagramy popisují vztahy mezi konkrétními třídami formou významových vazeb (např. obrázek 5.2, 5.3 nebo 5.4) nebo jako "flow" aktivity nebo dat (např. obr. 5.8). Pro účely zachycení referencí mezi třídami by měly sloužit "směrové asociace" s označením proměnných, pro účely komunikace mezi třídami pak spíše sekvenční nebo komunikační diagramy (místo třídních diagramů). Nedokončené části zadání a následná zamýšlená rozšíření aplikace a zlepšení kódu jsou v práci podrobně rozepsány, takže jejich realizace by byla již jednoduchá, leč studentovi na ně již nezbyl dostatek času.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	85 (B)
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 kompletní zdrojový kód rozšiřujícího balíčku pro aplikaci Winch pro anonymizaci souborů a dat ve strukturovaných souborech. Kód obsahuje pouze obchodní a datovou vrstvu řešení bez jakéhokoli uživatelského nebo programového rozhraní (to odpovídá zadání práce, na prezentační vrstvě pracuje jiný tým). Veškerý kód aplikace je však nadstandardně podrobně testování jednotkovými a integračními testy.

V aplikaci bohužel není implementována kompletní očekávaná funkčnost. Implementace anonymizačních funkcí pro rodné číslo, křestní jména a příjmení je prázdná, implementace anonymizace pro CSV soubory a ODS formát také nebyla realizována. Stejně tak nebyla implementována podpora pro jiná úložiště než lokální. Chybí také podpora pro wildcards a konfiguraci anonymizace pro skupiny souborů na obchodní vrstvě aplikace.

Zdrojové kódy jsou verzovány v repositáři zadavatele. Součástí je také podrobná dokumentace kódu.

Hodnotící kritérium:

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

4. Hodnocení výsledků, jejich využitelnost

90 (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:

Výsledkem práce je funkční první verze rozšíření systému Winch společnosti GEM System pro anonymizaci souborů a dat ve strukturovaných souborech. Ačkoli část plánovaných funkcionalit nebyla dokončena, aplikace je použitelná a je dobře připravena na další vývoj a rozšiřování. Součástí písemné práce je i podrobný popis způsobu implementace dalších funkcí a rozšíření.

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:

Z jakého důvodu byla zvolena použita reprezentace mapování anonymizačních tříd a funkcí/dekorátorů?

Hodnotící kritérium:

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

6. Celkové hodnocení

85 (B)

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:

Celkově hodnotím práci velmi dobře. V písemné části student velmi dobře popisuje požadavky na aplikaci, alternativy způsobu uložení anonymizačního modelu, všechna dílčí rozhodnutí a samotnou realizaci. Výsledkem je kvalitní a dobře dokumentovaný a otestovaný kód. Výsledná aplikace umožňuje anonymizaci souborů a dat v souborech XLS. Byť se nepodařilo implementovat všechny plánované funkce, práce obsahuje rozsáhlý a podrobný popis způsobu realizace těchto i dalších rozšíření.

Zejména z důvodu chybného použití třídních diagramů v písemné práci a nedotaženému rozsahu implementace navrhuji hodnotit stupněm B - velmi dobře.

Podpis oponenta práce: