



Posudek oponenta závěrečné práce

Oponent práce: Ing. Zdeněk Rybola, Ph.D.
Student: Karel Ševčík
Název práce: Generátor testovacích dat
Obor / specializace: Webové a softwarové inženýrství, zaměření Softwarové inženýrství
Vytvořeno dne: 23. srpna 2021

Hodnotící kritéria

1. Splnění zadání

- [1] zadání splněno
- ▶ [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Zadání práce považuji za splněné. Výhradou je absence komplexnějšího návrhu řešení z pohledu architektury řešení a především popis implementace a spouštění přes API.

2. Písemná část práce

60/100 (D)

Písemná práce je po formální stránce vyhovující. Splňuje požadovaný rozsah a obsahuje pouze příležitostné gramatické a typografické nedostatky.

Obsahově je práce logicky členěna, ovšem postrádá některé důležité části.

- Úvod je vcelku stručný a zcela postrádá jakékoli vysvětlení souvislosti se systémem Winch, jeho účelem (anonymizace) a obecně problematikou osobních údajů, na které je celý systém zaměřen.

- V analytické části jsou sice analyzovány 4 existující nástroje, ale vyvozené požadavky jsou popsány velmi stručně a nekonkrétně. Postrádám například podrobnosti o požadovaném způsobu práce s konfigurací či generátorem, jakoukoli souvislost se strukturou databáze (načtení struktury, omezení, apod.) i vlastnosti generovaných dat (nenulovost, unikátnost, závislosti mezi tabulkami, apod.).

- Návrhová část postrádá popis architektury řešení i souvislost se stávající architekturou nástroje Winch a začlenění nové aplikace/funkčnosti do ní. V doménovém modelu pak chybí celá řada podstatných informací - ručně vkládané záznamy, primární klíče a cizí klíče mezi tabulkami, datové typy sloupců s unikátností, nenulovostí či omezením délky hodnot, podpora více dekorátorů pro jeden sloupec.

- V kapitole věnované implementaci je podrobněji vysvětlen postup generování a řešení pomocí rozdělení na jednotlivé kroky a úlohy a jejich řazení Kahnovým algoritmem. Popis

však vychází z vysvětlení závislostí mezi kroky, z čehož lze výsledné pořadí těžko identifikovat. Postrádám nějakou formu vizualizace jednotlivých kroků (např. Activity diagram). Také je v textu zmiňována řada balíčků a implementačních tříd, postrádám však jejich přehledné zobrazení formou diagramu, aby se čtenář mohl snáze zorientovat. Z popisu také není zřejmé, jestli se zpracovávají a řadí úlohy pro každou tabulku zvlášť nebo dohromady a také není zřejmé ošetření porušení jakýchkoli integritních omezení v rámci vygenerovaných dat, čímž může být část záznamů neplatná a způsobit chybu při vkládání.

- Kapitola zaměřená na testování je stručná a bez konkrétnějších informací. Oceňuji však alespoň shrnující tabulku s pokrytím kódu testy.

- V práci pak celkově postrádám jakékoli informace o možnosti použití aplikace přes API v rámci CI/CD, jak bylo požadováno i v zadání. Popis se zaměřuje především na GUI.

- V závěru implementační kapitoly autor zmiňuje řadu nedostatků, které nebyly v práci adresovány. Řada z nich by však jistě byla v práci snadno realizovatelná a jejich absence výrazně snižuje použitelnost (např. podpora unikátnosti).

3. Nepísemná část, přílohy

75 /100 (C)

Součástí bakalářské práce je vytvořený prototyp pro generování testovacích dat splňující zadání práce. Aplikace se zdá být funkční, ačkoli jsem neměl možnost ji vidět v praxi.

Zdrojové kódy jsou rozčleněny do řady balíčků, ke kterých však chybí dokumentace (model). Zdrojové kódy nejsou nijak komentovány, s výjimkou anotace generátorů. Pro implementaci byl zvolen jazyk Groovy a framework Swing, avšak bez podrobnějšího zdůvodnění v rámci textové části práce.

Zdrojové kódy jsou částečně pokryty jednotkovými testy. Některé testy mají však pramalou přidanou hodnotu - např. test na generátor hodnoty typu "příjmení" pouze porovná, že vygenerované příjmení pro ženu a muže se liší.

4. Hodnocení výsledků, jejich využitelnost

80 /100 (B)

Výsledná aplikace je použitelná pro generování testovacích dat do jednoduchých databází. Bohužel v případě složitějších databází s řadou integritních omezení bude narážet na problémy s porušením těchto omezení a nemožností vygenerovaná vložit. Výsledek tak beru jako prototyp vhodný k rozšíření před reálným nasazením do produkčního užití.

Celkové hodnocení

68 /100 (D)

Celkově hodnotím práci jako lehce podprůměrnou. Výsledná práce sice splňuje zadání, ale jak písemná, tak programová část má řadu nedostatků zmiňovaných výše. Jako největší nedostatky vidím absenci architektonického návrhu a dokumentace a také podporu integritních omezení existujících v cílové databázi, kam mají být testovací data generována.

Otázky k obhajobě

1. Proč byl k implementaci zvolen framework Swing? Zdůvodnění v práci zmiňuje pouze existenci podpory v Groovy. JavaFX nemá podporu? V čem by se implementace v modernějším JavaFX lišila?

2. Jak by bylo třeba logiku generování upravit pro podporu integritních omezení v databázi ovlivňujících množinu validních dat (unikátnost, CHECK omezení, příp. trigger)?

Instrukce

Splnění zadání

Posudte, 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, posudte 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.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, 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. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, 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.

Nepísemná část, přílohy

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.

Hodnocení výsledků, jejich využitelnost

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.

Celkové hodnocení

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.