



Posudek oponenta závěrečné práce

Oponent práce:	Mgr. Vladislav Severa, Ph.D.
Student:	Filip Sikora
Název práce:	Modelování a řízení dat v doméně online vzdělávání
Obor / specializace:	Webové a softwarové inženýrství, zaměření Softwarové inženýrství
Vytvořeno dne:	6. června 2021

Hodnotící kritéria

1. Splnění zadání

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

Předložená práce dobře splnila zadání, které spočívalo ve vytvoření modelové datové ontologie pro definici datových domén a zdrojů společnosti Learneron, dále pak ve vytvoření celkového datového modelu a konverzí z různých izolovaných částí, a konečně také v konkrétní vzorové implementaci. V případě Learneronu se konkrétně jednalo o zadání modelovat data jako propojené entity v rámci grafu - jedná se totiž o navzájem úzce "interagující" datové entity typu studenti a jejich vzdělávací položky.

Projekt se zabýval zadáním nejenom z pohledu prostého (ovšem i tak komplexního) ontologického zachycení struktury dat prostředky OntoUML a konzistentní, vskutku uspořádávající vzorovou implementací konverze dat z různorodých, převážně relačních zdrojů do grafové databáze Neo4J, ale také vnesl do datové struktury řadu prvků, které lze nazvat prozíravou a nejlepší praxí. Jako příklad může sloužit modelování některých vztahů mezi entitami v grafu nikoliv toliko prostřednictvím hran (k čemuž by sváděla přímočarost takového postupu), ale prostřednictvím kombinace hran a nových entit - taková volba implementace v budoucnosti umožní i zachycení a modelování vztahů "vyššího řádu", například "vztahů mezi vztahy".

Ocenit je nutné také metodickou a důslednou práci zpracovatele s ontologií a ontologickým modelováním: výsledný datový model, který v sobě propojuje řadu domén světa vzdělání, světa práce i uživatelů, je poměrně komplikovaný a jeho komplexita v čase spolu s dalšími zapojovanými datovými doménami jenom poroste. Ontologický model - narozdíl například od schématu grafové databáze, který postupně ztrácí přehlednost - pomůže i při rostoucí komplexitě udržet konzistenci, přehled a pořádek.

Kromě výše zmíněného se pak zpracovatel také dobře vyrovnal s vlastní praktickou implementací datových konverzí: propojil svět relačních databází (MySQL) s grafovým světem (Neo4J), prostřednictvím vhodného dotazovacího jazyka (Cypher), zautomatizovaného jazykem současné datové vědy (Python), a to vše v prostředí cloudu (AWS). Velmi dobrá ukázka kombinace tzv. DevOPS.

2. Písemná část práce

95 /100 (A)

Písemná část práce je dobře provedena, kapitoly jsou přiměřeného rozsahu i obsahu a na sebe logicky navazující. Příspěvek samotného zpracovatele je od čerpání z literatury a citací dobře odlišený, literatura je použita a citována příkladně. Výkladový text je pochopitelný, byť místy ne bez znalosti kontextu problematiky - což je ovšem automatický předpoklad kladený na čtenáře odborných textů.

Ačkoliv nejde (a ani nemá jít) o vyčerpávající monografii pokrývaných témat, na řadě míst lze ocenit autorovy poznámky k existujícím alternativám či aspektům daných témat, které jdou za rámec bezprostředního zadání - například v kapitole o ontologickém modelování pasáž o vrstvě UFO-B o modelování událostí a procesů. Jako čtenář, který není odborníkem na danou oblast, ale který intenzivně hledá užitek dané oblasti pro svůj hlavní předmět zájmu (v tomto případě grafové modelování komplexních sociálních jevů a interakcí) právě takové naznačení "co je dál" velmi vítám a oceňuji.

K rovině jazykové: práce pojednává o odborném tématu v oborech softwarového a datového inženýrství a český jazyk, přes všechnu svou bohatost, není těmto oborům ideálně uzpůsoben. Zpracovatel se s jazykovými úskalími vyrovnává většinou se ctí, bez větších ztrát na jasnosti výkladu.

Typograficky je práce zpracována přehledně. Co se týče licenčních a autorských podmínek používaných nástrojů, jsou v textu uváděny jasně a podle mých vědomostí v souladu se skutečností.

3. Nepísemná část, přílohy

100 /100 (A)

Podle mého soudu jsou všechny nepísemné výstupy práce - OntoUML model, Cypher i Python skripty a další - zpracovány zcela bezvadně a replikovatelně. Jsou i vhodným způsobem dokumentovány (či samy slouží jako dokumentace zpracovávané problematiky).

4. Hodnocení výsledků, jejich využitelnost

100 /100 (A)

Výstupy předložené práce jsou v současné době implementovány, testovány nebo nasazeny v provozu společnosti Learneron a navíc slouží jako východisko pro další významové a funkční vrstvy Learneronu - především pro sémantické nástroje a pro obsahový doporučovací systém. Práce co do své využitelnosti beze zbytku splnila počáteční očekávání. Jde o příkladný přenos ne zcela běžného ontologického a grafového výzkumu a vývoje do aplikační praxe.

Celkové hodnocení

99 /100 (A)

Předložená práce je co do obsahu písemné části velmi dobře zpracována a z mého pohledu jsou zvláště hodny pozornosti body, které jdou za rámec přímočarého zpracování

a přinášejí svého druhu nejlepší a sofistikovanější praxi - například výše zmíněné modelování vztahů mezi grafovými entitami způsobem, který později umožní i zachycení "vztahů mezi vztahy". Za zásadní pak považuji nasaditelnost a užitečnost daného řešení v prostředí společnosti Learneron.

Otázky k obhajobě

1. Jak se ontologický (a grafový) model dat v principu liší od schematu relační databáze? V jakých situacích je vhodnější ten či onen přístup?
2. Ontologická vrstva UFO-B pro události a procesy - byla by vhodná například pro modelování dynamických změn / časového vývoje grafu (dat Learneronu)?
3. Proč a v jakých případech je vhodné modelovat i relace jako svého druhu uzly na grafu? Jaký by mohl být příklad "relací mezi relacemi" v nějakém obecném případě, či v případě dat a grafu Learneronu?

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. 26/2017, 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.