



Posudek oponenta závěrečné práce

Oponent práce:	Ing. David Bernhauer
Student:	Matěj Karpíšek
Název práce:	Modul pro chytré vyhledávání v LearnShellu
Obor / specializace:	Webové a softwarové inženýrství, zaměření Webové inženýrství
Vytvořeno dne:	7. 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

Student měl navrhnout pro aplikaci LearnShell rozšíření, které by umožňovalo lepší vyhledávání. Student ve své práci nasadil v rámci systému Learnshell vyhledávací engine Elasticsearch a zaintegroval tuto službu do existujícího kódu. Zadání je tak spíše menšího rozsahu, jelikož 3 body zadání sloučil na "nasazení systému Elasticsearch". Při takto přímočarém zadání mi chybí důkladná analýza variant a jejich fungování (proto "s menšími výhradami").

2. Písemná část práce

70/100 (C)

Písemná část je psána poměrně příjemným stylem. První kapitola popisuje základní technologie používané v rámci aplikace LearnShell, zde mi přijde lehce nevyvážený popis technologie Docker, která má oproti ostatním celkem 3 strany. Druhá kapitola se zabývá analýzou a návrhem modulu pro vyhledávání, z analýzy zde nalezneme pouze "Požadavky", následuje popis full-textového vyhledávání a možnosti vyhledávacích enginů se zaměřením na Elasticsearch. Tady bych očekával detailnější srovnání, zcela chybí produkty jako Sphinx, Terrier nebo Xapian. Z popisu mi není jasné, zda se autor zabýval i možností využití embeddingů nebo jak jednotlivé technologie fungují na pozadí. Z návrhu se autor omezuje pouze na model dat a jednoduchý náskok globální architektury aplikace. Tato část by si zasloužila větší pozornost. Stejně tak některá rozhodnutí vyžadují kvalitnější zdůvodnění, př. 2.5.1, proč zrovna tento druh denormalizace bude působit kladně na výkon aplikace? Třetí kapitola popisuje realizaci, čtvrtá kapitola popisuje jednotkové testování, k těm nemám příliš výhrad.

Rozsah práce je dostatečný, a práce obsahuje minimum chyb. Ale chybí mi srovnání s aktuálním řešením, příp. využitím PostgreSQL. A dále již jednou zmíněná hloubková rešerše popisovaných systémů. Student cituje primárně internetové zdroje, v mnoha případech by bylo vhodnější najít kvalitnější zdroje uvádějící skutečné experimenty či nějaké základní provést v rámci své práce.

3. Nepísemná část, přílohy

80/100 (B)

Zdrojový kód dodaný v přílohách vypadá rozumně, ačkoli je spíše menšího rozsahu. Většina práce byla vyřešena nasazením Elasticsearch a jeho "pouhým" nastavením a napojením na již existující systém. Samotné testování je pouze jednoduché, skládající se z jednotkových testů. Testování QueryBuilderu je dle mého názoru umělé, test porovnává na rovnost data knihovnicí funkcí, které se v budoucnu mohou měnit a pravděpodobně nijak neovlivní zbytek vyhledávání (vnějšího chování). Naopak musím pochválit testování autorizace. Testování "NF1 - Search performance" jsem v práci nikde nenašel, ačkoli v požadavcích je bod požadující vyhledávání pod 1s.

4. Hodnocení výsledků, jejich využitelnost

80/100 (B)

Text práce slouží spíše jako pěkný průvodce nasazením systému s vysvětlením základních pojmů. Chybí mi zde alespoň srovnání s existující variantou, případně standardními indexy nad PostgreSQL. Samotný výsledek prezentovaný v přílohách D a E vypadá použitelně. Hodnocení zda a jakým způsobem bude modul nasazen nechám na vedoucím práce, jelikož jsem integraci modulu neviděl.

Celkové hodnocení

75/100 (C)

Hodnocení zohledňuje, že práce je spíše jednoduššího charakteru a popis by si zasloužil rozšířit o standardní přístupy softwarového inženýrství. Samotný text práce je pěkný, ale chybí kvalitnější zdroje, detailnější rešerše či experimentální vyhodnocení. Vzhledem k rozsahu nechávám na uvážení komise, zda odvedená práce svým rozsahem odpovídá rozsahu bakalářské práce, v případě pozitivního rozhodnutí (ke kterému se i osobně přikláním) navrhuji 75 bodů.

Otázky k obhajobě

Jak vypadá diagram nasazení?

Bylo provedeno srovnání PostgreSQL vs Elasticsearch?

Jaká je skutečná responzivita vyhledávání?

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.