



POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Jméno studenta: Kristýna Sulíková

Název práce: Využitelnost NoSQL databází v datových skladech

Autor posudku: Ing. Karel Mls, Ph.D.

Cíl práce: Cílem práce je nalezení využitelnosti NoSQL databází v datových skladech. Cíl práce není stanoven v Zadání.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vyjádření k výsledku anti-plagiátorské kontroly

Práce byla vyhodnocena v Odevzdej.cz (celková podobnost 0%). Práce je originální.

Dílčí připomínky a náměty:

Pravopis – v textu se vyskytují překlepy, chyby v interpunkci, místy významově podstatné (netadata x metadata, nanosekundy x milisekundy) a jazykové neobratnosti dané zřejmě překladem pojmů a definic z cizojazyčných zdrojů do českého jazyka („Podle zprávy IDC objem dat do roku 2020 dosáhne 40 bajtů Zeta a zvýší se až 400 krát“).

Faktické chyby: „Pojem databázový model zavedl C. J. Date, který ho popsal jako definici entit, atributů a integritních omezení entit [16].“ – v seznamu literatury odpovídá položka č. 17, obecně se za autora pojmu datový model uvádí E. F. Codd (relační model).

V praktické části autorka chybně uvádí „Potvrzena hypotéza“ (tab. 8.2, 8.4, 8.6, 8.8, 8.10, 8.12, - hypotézu H0 buď přijímáme, nebo zamítáme ve prospěch přijetí hypotézy H1. Není jasné, jaká hypotéza se testuje v kap. 8.5 a 8.6 – viz hodnoty v tabulce 8.11 a výsledky testu v tab. 8.12, resp. tab. 8.13 a 8.14.

Hodnoty v tabulkách byly naměřeny v nanosekundách – min 0,76 max 4,14 median 0,82 avg 0,87 – tab. 8.1 (a další)? od tab. 8.10 uváděny realističtější milisekundy.

Obr. 4.1 převzatý z blogu S. Sonnelanda je spekulativní, není podložen daty (chybí měřítko na ose y)

Formátování citací neodpovídá standardu ČSN ISO 690 (viz Výnos děkana č. 1/2000), některé záznamy jsou neúplné.

Celkové posouzení práce a zdůvodnění výsledné známky:

Autorka se v práci zaměřila na NoSQL databáze a srovnání jejich vlastností při různě složitém dotazování a pro různě velké datové sady.

V teoretické části jsou představeny základní pojmy související s tématem, poněkud netradičně jsou jako referenční publikace použity nedávno publikované (studentské) závěrečné práce na téma NoSQL databáze.

Popsané výsledky dávají tušit, že autorka práci věnovala poměrně značné úsilí, provedla instalaci jednotlivých databázových systémů a celou řadu experimentů, bohužel prezentace výsledků je pouze popisná (tabulky a grafy naměřených hodnot) a chybí hlubší analýza, proč se naměřené časy pro jednotlivé dotazy a případně různě velká data liší.

Po formální stránce předkládaná závěrečná práce s výhradami splňuje požadavky Výnosu děkana FIM 1/2000 – Metodických pokynů pro vypracování bakalářských a diplomových prací. Je škoda, že zajímavé téma trpí velkým množstvím formálních ale i věcných nedostatků.

Práci doporučuji k obhajobě a s ohledem na zmiňované nedostatky navrhuji hodnocení stupněm „C“ až „D“, podle výsledku obhajoby.

Otázky k obhajobě:

- V kap. 8.2 popisujete metodu generování náhodných dat – o jaký rozsah (v kB, MB) se jednalo? V kap. 8.7 tento základní set dat desetkrát zdvojnásobujete – jak se toto zvýšení projeví na jednotlivých záznamech (2 na 10 je cca 1000)?

Práci doporučuji k obhajobě.

Navržená výsledná známka: C

V Hradci Králové, dne 27. května 2019

podpis