

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Jméno studenta: Martin Gold

Název práce: **Analýza a optimalizace procesu zpracování otevřených dat**

Královehradeckého kraje

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

Cíl práce: Cílem práce je analýza stávajícího řešení zpracování otevřených dat

Královehradeckého kraje a návrh řešení pro efektivnější a automatizované fungování.

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 type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input 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 checked="" 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 type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vyjádření k výsledku anti-plagiátorské kontroly

Antiplagiátorská kontrola systémem Odevzdej.cz našla shodu 0%. Práce je originální.

Dílčí připomínky a náměty:

Práce je po formální i obsahové stránce na dobré úrovni, v textu se **bohužel nachází poměrně vysoké množství pravopisných a formulačních nedostatků** (shoda, interpunkce, neodkončená slova a věty, velká/malá počáteční písmena).

Vysvětlení významu a vztahu mezi pojmy data – informace – znalosti v kap. 3.2 je nesprávné.

„3 Kapitola - Vlastní text práce“ by šla jistě pojmenovat vhodněji.

Tabulka 1 – jedná se o výšku nebo známku?

Převzaté obrázky mají různou kvalitu i formát, pro potřeby DP bylo vhodnější jejich vzhled sjednotit.

Vzorci v textu nejsou číslovány.

Celkové posouzení práce a zdůvodnění výsledné známky:

Jak jsem konstatoval v Dílčích připomínkách, práce obsahuje značný počet pravopisných chyb a formulačních nedostatků – a zatímco s chybami typu „Tyto data, společnosti začali, Královéhradecký Kraj“ se čtenář setkává víceméně běžně a tyto chyby nemají zásadní vliv na sdělovaný význam, formulace jako „Odvozením znalostí z těchto dat vzniká informace; Role datový vědec bývá často používána jako souhrnná role všech rolí v oboru; CRISP-DM (...) je rozsáhlá data-miningová meteorologie a procesní model; Návrh automatizovaného spočívá v nasazení ELT procesu do prostředí Data KHK:...” nejenom že ztěžují čtení textu, ale mohou čtenáře uvést v omyl. Po formální stránce bych tedy práci doporučil přepracovat. Samotná teoretická část představuje některé základní pojmy datové vědy a souhrn souvisejících témat včetně dvou DM metodik. Je třeba ovšem konstatovat, že toto představení je nejenom zatíženo zmíněnými formálními nedostatky, ale na poměrně malé ploše 25 stran působí velmi povrchním dojmem.

Autor v textu poněkud nekonzistentně používá příklady návštěvnosti fitness centra nebo přehled počtu osob s uděleným pobytovým oprávněním, které nekorespondují s praktickou částí, pracující s daty o proočkovánosti obyvatel Královéhradeckého kraje proti covid-19.

V praktické části (23 stran) autor na tyto teoretické základy navazuje návrhem datového skladu v prostředí MS SQL. Postup je standardní, text se víceméně skládá z komentovaných výpisů SQL příkazů podle navržené čtyřvrstvé struktury datového skladu. Představení (ne analýza) stávajícího řešení zpracování otevřených dat Královéhradeckého kraje je pak na s. 28 a 29 a na obr. 11.

Práce v části návrhu datového skladu splnila vytyčený cíl, její zpracování splňuje očekávané nároky kladené na závěrečné práce.

Práci doporučuji k obhajobě a zejména s ohledem na formální nedostatky navrhuji hodnocení stupněm D.

Otázky k obhajobě:

V práci zmiňujete tři role v datovém týmu. Proč pouze tři a proč právě tyto konkrétní role?

Jakým způsobem probíhalo testování navrženého řešení?

Jaké hw/sw nároky má MS SQL server pro navrhovaný datový sklad?

Práci **doporučuji k obhajobě.**

Navržená výsledná známka: D

V Hradci Králové, dne **27. května 2023**

podpis