

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Andrej Pečimúth
Název práce Datový generátor
Rok odevzdání 2021
Studijní program Informatika **Studijní obor** Obecná informatika

Autor posudku RNDr. Michal Kopecký, Ph.D. **Role** Vedoucí
Pracoviště KSI MFF UK

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci

	lepší	OK	horší	nevyhovuje
Obtížnost zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Komentář Práce si kladla za cíl pomoci vývojářům databázových aplikací zaplnit databázi generovanými daty, které by byly dostatečně podobné datům reálným. Měly by zachovat jak specifikaci typu, tak případně další statistiky jako procento NULL hodnot, unikátnost a podobně. Autor se rozhodl pro vytvoření webové aplikace, která by dovolila se připojit ke zdrojové databázi, načíst údaje o schématu (definice tabulek a sloupců), načíst stávající data a z nich odhadnout jejich podobu. Uživatel si může získaná data opravit podle potřeby.

Textová část práce

	lepší	OK	horší	nevyhovuje
Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analýza	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vývojová dokumentace	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Uživatelská dokumentace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Komentář Text práce je poměrně podrobný a na zhruba sto stranách popisuje analýzu stávajících aplikací, specifikaci požadavků na vytvářenou aplikaci Synthia, její design a návod k použití. Vzhledem k běhu na aplikačním serveru je potřeba aby byla databáze dostupná z aplikačního serveru. Pokud se jedná o databázi v intranetu, může být aplikace nainstalována lokálně ve stejné síti.

Implementační část práce

lepší OK horší nevyhovuje

Kvalita návrhu ... *architektura, struktury a algoritmy, použité technologie*Kvalita zpracování ... *jmenné konvence, formátování, komentáře, testování*

Stabilita implementace

Komentář Aplikace je poměrně dobře navržena. Zvolené řešení umožňuje běh nezávisle na operačním systému jak daného počítače, tak operačního systému, kde běží databázový server. Díky knihovně SQLAlchemy má aplikace teoreticky možnost pracovat jak s komerčními servery MS SQL/Oracle, tak s volně šiřitelnými databázemi jako MySQL nebo PostgreSQL. Prakticky však stávající aplikace nabízí jen PostgreSQL, což je škoda.

Pěkně je vyřešené generování tabulek s cizími klíči. Nejen že dokáže tabulky generovat ve správném pořadí, ale poradí si i s případnými cyklickými závislostmi.

Data číselných sloupců jsou generovány jen s dost omezenými možnostmi nastavení. Bylo by hezké podporovat nejen interval daný rozsahem, ale i další statistická rozložení. Existují i aplikace, které dokáží z rozsáhlých dat extrahovat podrobné statistické údaje a histogramy. Bylo by proto zajímavé mít možnost takovou aplikaci zapojit do procesu extrakce metadat, a generovat data s ohledem na další informace.

Bez ohledu na výše uvedené připomínky práce splnila jak zadání, tak požadavky kladené na práce bakalářské. Doporučuji ji proto k obhajobě.

Celkové hodnocení Výborně**Práci navrhuji na zvláštní ocenění** Ne**Datum** 19. června 2021**Podpis**