

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Kryštof Hrubý

Název práce Semi-automatizované mapování dat do ontologií

Rok odevzdání 2024

Studijní program Informatika **Studijní obor** Softwarové a datové inženýrství

Autor posudku Petr Škoda **Role** vedoucí

Pracoviště Katedra softwarového inženýrství

Text posudku:

Cílem práce je návrh a implementace semiautomatizovaného řešení pro převod (lifting) dat do 5-star podoby. Z hlediska uživatele jsou data reprezentována datovým schématem, které může uživatel upravovat (transformovat). Jednotlivé transformace kromě úpravy schématu transformují také samotná data. Transformace provádí uživatel ve webovém editoru. Jednou z hlavních částí funkcionality je možnost využít doporučených transformací. Cílem těchto doporučení je zpřístupnit převod širšímu okruhu uživatelů.

Z hlediska textové části práce autor využívá motivační příklad, na kterém demonstruje jednotlivé pod-úlohy převodu. Ty zahrnují nejen technickou úpravu struktury, ale zejména mapování na doménově specifické ontologie a zdroje, například číselníky. Po poměrně extenzivním představení related-work pro jednotlivé pod-úlohy autor v sekci 2.4 představuje navržený přístup.

V další sekci autor popisuje design řešení a implementaci. Z hlediska architektury je popsáno nejen výsledné řešení, ale také uvažované alternativy. Zvolená architektura obsahuje 10 komponent, které společně podporují požadovanou funkcionalitu. Poměrně výraznou částí softwarové části jsou zdroje doporučení v editoru, ty se skládají z analyzátorů a doporučovačů. Analyzátoři jsou schopni analyzovat uživatelem vložená data a připravit data pro doporučovače. Data k analýze jsou vkládána pomocí metadata s využitím DCAT slovníku. Tato skutečnost, nejen že demonstruje respekt autora k doméně, ale také umožňuje nasazení na existujících datových katalozích. Jako úvodní naplnění autor implementuje 6 analyzátorů a 5 doporučovačů.

Textová část práce dále obsahuje uživatelskou a administrátorskou dokumentaci. V 5. kapitole autor, v podobě tutoriálu, demonstruje použití implementovaného softwaru pro vyřešení motivačního příkladu.

Softwarová část práce je poměrně rozsáhlá a využívá soudobé technologie a přístupy jak z hlediska správy projektu, implementace, tak i nasazení.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 31.5.2024

Podpis: