

Název práce:

Integrace a konzumace důvěryhodných Linked Data

Autor:

RNDr. Tomáš Knap

Katedra:

Katedra Softwarového Inženýrství

Vedoucí disertační práce:

RNDr. Irena Holubová, PhD., Katedra Softwarového Inženýrství

Abstrakt:

V posledních letech celá řada jedinců a společností (od univerzitních výzkumníků, přes soukromé společnosti, až po vládní úřady) začíná publikovat na Webu svá data s využitím otevřených standardů. Linked Data nabízí konzumentům dat úroveň agregace a integrace dat, která nebyla až doposud možná. Uživatelé si nyní mohou sestavit potřebná data zcela dle svých požadavků. S tímto procesem výběru dostupných dat z široké palety zdrojů souvisí řada problémů, zejména nízká kvalita dat, nepřesnost dat, nízká relevance dat pro účely daného konzumenta a také data, která jsou záměrně pozměněna. Všechny tyto faktory pak přináší další náklady související s použitím dat, které negativně ovlivňují přínos principů Linked Data pro jejich uživatele a rozšíření Linked Data aplikací.

V této práci proto navrhujeme ODCleanStore, nástroj pro správu a dotazování Linked Data, který je schopen poskytovat konzumentům dat Linked Data, která jsou pročištěná, prolinkovaná a důvěryhodná dle daných subjektivních požadavků konzumenta. Důvěryhodnost dat znamená, že tato data jsou asociovaná s informací o jejich původu, který odpovídá požadavkům konzumenta, dále tato data mají určitou kvalitu požadovanou konzumentem a také jsou poskytována důvěryhodnými uživateli. V této práci navrhujeme vlastní komponentu pro integraci dat implementovanou do nástroje ODCleanStore, která řeší konflikty mezi heterogenními daty a doplňuje takto integrovaná data informacemi o kvalitě dat a původu těchto dat. Abychom umožnili vyjádřit informace o původu dat na Webu, navrhujeme v práci vlastní model pro zachycení původu dat na Webu – W3P. V práci také diskutujeme důvěryhodnostní modely a jejich užitečnost pro výpočet důvěryhodnosti agentů (poskytovatelů a konzumentů Linked Data) v sociálních sítích. Nástroj ODCleanStore je k dispozici pod otevřenou licencí a je plánováno jeho nasazení ve výzkumné skupině Agile Knowledge Engineering and Semantic Web (AKSW) na Lipské univerzitě, na univerzitě v Miláně, pro zákazníky společnosti Semantic Web Company v Rakousku a v rámci projektu <http://opendata.cz>.

Klíčová slova:

konzumace Linked Data, integrace Linked Data, původ dat na Webu, důvěryhodnost v sociálních sítích

