

Posudek oponenta bakalářské práce

Autor/autorka práce: **Martin Kryl**

Název práce: **Benchmark RDF úložišť**

Obsah práce

Cílem bakalářské práce bylo popsat teorii RDF metadatového modelu a způsoby testování RDF úložišť při zpracování rozsáhlých dat s ohledem na rychlost plnění daty a rychlost vyhledávání. Autor vybírá metodu testování Berlin SPARQL Benchmark (BSBM) a vytvořil sadu SPARQL dotazů nad medicínskými daty na základě používaných ontologií (DASTA, DICOM, NIHSS, SITS). Dále provedl testy na zvolených úložištích (Virtuoso OpenSource, Mulgara, OpenRDF Sesame , Apache Jena TDB, Apache Jena SDB, Oracle).

Kvalita řešení a dosažených výsledků

Množství testovacích medicínských dat (50 milionů RDF trojic) je pro zátěžové testy plně dostačující. V dosažených výsledcích vychází, že Jena TDB je nejrychlejším persistentním RDF úložištěm při načítání medicínských dat a Jena SDB nad relačními databázemi je nevhodná pro velké množství dat. Autor neoznačuje nejlepší RDF úložiště, ale pouze poukazuje na slabé a silné stránky daných úložišť.

Formální úroveň

Po formální stránce je kvalita práce vyhovující. Práce je logicky strukturovaná. Počet překlepů je v toleranci vzhledem k rozsahu práce (33 stran textu práce). V bakalářské práci je 19 zdrojů a čtyři přílohy. Několik stran bakalářské práce je nekvalitně vytištěno.

Doplňující informace k práci

Tištěná příloha obsahuje BSBM SPARQL dotazy, MRE SPARQL dotazy, první graf zobrazující načítání x-tého milionu trojic včetně indexování na prvním testovacím počítači a druhý graf zobrazující dobu importování dat do úložišť. Elektronická příloha obsahuje text bakalářské práce, tabulky s naměřenými daty a zpracovanými výsledky, Bash skripty pro ovládání úložišť, logovací soubory z testování medicínských dat, logovací soubory z testování podle BSBM a SPARQL dotazy.

Dotazy k práci

Proč jste nezahrnul RDF úložiště BigOwlím do testování?

Které z testovaných RDF úložišť byste vybral pro data o velikosti 5 milionů RDF trojic?

Splnění zadání

K bakalářské práci nemám výhrady. Všechny body zadání byly splněny.

Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.

V Plzni 20. 5. 2013


Ing. Petr Brůha