

systemech, text by měl obsahovat jejich (alespoň krátký) přehled s přihlédnutím jak k normám, tak ke zvolenému databázovému systému.

- Přestože práce obsahuje několik experimentů, které byly použity pro účely volby nejhodnějších variant uložení, v závěru schází očekávané experimentální srovnání dvou navržených metod mapování. Autor pouze uvádí, že druhá metoda je 20x rychlejší než první, ale není úplně jasné jakým způsobem k tomuto závěru došel.
- Seznam použité literatury je překvapivě malý – práce obsahuje pouze 7 položek (2 odborné články a 5 webových stránek), navíc s poměrně malým množstvím informací (schází roky vydání, autoři, čísla verzí specifikací apod.). Větší množství odkazů na související přístupy, technologie, specifikace apod. s podrobnějším popisem by výrazně zjednodušily čtení textu a začlenily text do souvislostí.

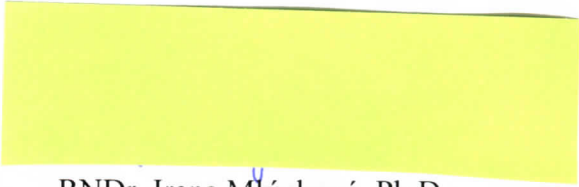
K méně zásadním připomínkám patří následující postřehy:

- Tvzení, že nevýhodou uložení XML dat v LOB je velká redundance, není úplně přesné. Z textu se zdá, že autor pojmu „redundance“ nerozumí úplně správně.
- V úvodu autor definuje „relační databáze s podporou/rozšířením XML“ jako databáze které podporují datový typ XML. Jako jiný typ uvádí databáze, které mapují XML data do složitějších databázových struktur. To může být poněkud zavádějící – pojem „databáze s podporou/rozšířením XML“ totiž zastřešuje oba tyto typy, které se liší tím, jak jsou data logicky uložena.
- Poměrně matoucí je také použití pojmu „nativní XML úložiště“ v kapitole 6. Datový typ XML není obvykle považován za variantu nativního ukládání XML dat.

I přes výše uvedené připomínky, které jsou míněny spíše jako rady pro zkvalitnění díla, splňuje práce původní zadání i cíl. Student dostatečně pronikl do zadané problematiky a prokázal schopnost detailně zanalyzovat a řešit zadaný problém. Práce obsahuje několik zajímavých analýz, poznatků a nápadů, které by bylo jistě vhodné (s přihlédnutím k výše uvedeným připomínkám) opublikovat na tematicky související vědecké konferenci.

Práce Petra Uhera podle mého názoru bezpochyby splňuje podmínky na diplomovou práci kladené, a proto ji **doporučuji** jako práci diplomovou **uznat**.

V Praze, 2.1. 2008



RNDr. Irena Mlýnková, Ph.D.  
KSI MFF UK

## **Ukládání XML dat popsaných modelem XSEM v databázi**

Předkládaná diplomová práce se zabývá možnostmi mapování XML dat popsaných modelem XSEM, umožňujícím jejich konceptuální modelování, do relačních databází. Pro tyto účely se autor inspiroval metodami schématem řízeného mapování XML dat do relačního modelu a navrhl (a implementoval) jeho modifikaci pro model XSEM.

Práce je rozdělena do několika částí. V první části je čtenář uveden do problematiky, výhod a nevýhod možných typů ukládání XML dat a seznámen s modelem XSEM. Další část práce popisuje vlastní návrh dvou variant mapování jednotlivých prvků modelu XSEM do relačního modelu. V závěru práce je popsána experimentální implementace navrženého přístupu, zhodnoceny jeho výhody a nevýhody a nastíněna další možná rozšíření.

Hlavními klady práce jsou především:

- Zaměření na poměrně novou a ne příliš známou problematiku modelování XML dat, kterou autor evidentně velmi dobře pochopil a dále rozpracoval.
- Jasný a přehledný výklad modelu XSEM.
- Diskuze možných způsobů uložení prvků modelu XSEM s přihlédnutím k jejich výhodám a nevýhodám demonstrovaných na příkladech a experimentech.
- Detailní návrh dvou možných variant mapování XML dat a přehled jejich výhod a nevýhod.

Práce má kvalitní úpravu a je napsána poměrně srozumitelně. Navíc obsahuje množství vysvětlujících příkladů a schémat, které pochopení ne úplně známé problematiky zjednodušují. V práci je minimální množství gramatických chyb.

Z věcného hlediska má následující nedostatky:

- Práce by zasloužila podrobnější analýzu existujících přístupů. V předloženém textu jsou nastíněny obecné přístupy k ukládání XML dat (LOB, nativní databáze a databáze s XML rozšířením) a jako výchozí přístup je zvoleno schématem řízené mapování XML dat do relačního modelu. Hlubší analýza konkrétních zástupců těchto metod chybí, stejně jako analýza metod které mapují do relací např. klasická ER schémata (a jejich různé varianty).
- Práce má poměrně zvláštní členění – ještě před vysvětlením problematiky, kterou se bude zabývat a souvisejících základních pojmů, probíhá volba databázového systému v němž bude přístup implementován a jsou popsány související experimenty, u nichž se předpokládají poměrně netriviální znalosti. To může navíc budít dojem, že přístup není navržen obecně, ale pouze pro jeden určitý databázový systém a jeho speciální, nestandardní funkce.
- Autor v celém textu předpokládá, že má čtenář velmi detailní znalosti relačního modelu a konstruktů jazyka SQL (jako např. datového typu XML a jeho použití, uživatelsky definovaných typů, vestavěných funkcí apod.). Vzhledem k tomu, že se obvykle liší specifikace těchto konstruktů a jejich podpora v různých databázových