

# Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

**Student:** Miroslav Špak  
**Oponent práce:** Mgr. Jiří Toušek  
**Název práce:** Nástroj pro extrakci metadat z databáze DB2  
**Obor:** Softwarové inženýrství

**Datum vytvoření:** 30. 5. 2017

<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 5:</b>
<b>1. Náročnost a další komentář k zadání</b>	<b>1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání</b>
<b>Popis kritéria:</b> Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
<b>Komentář:</b> Práce se zabývala extrakcí metadat z databázového systému DB2.  Autor se musel detailně seznámit s tímto DBMS a se způsobem, jak ukládá a zpřístupňuje metadata o definovaných objektech (nad rámec získání základních metadat pokrytých standardem SQL). Práce ve velké míře vyžadovala provádění vlastních experimentů s tímto databázovým systémem a některými jeho nedokumentovanými vlastnostmi.  DB2 je transakční databáze, má tak bohatší sadu typů objektů než databáze používané pro datové sklady.  Databázi DB2 lze přepnout do režimu kompatibility se skripty v jazyce Oracle PL/SQL. Práce odhalila, že uložení metadat Oracle PL/SQL objektů není v DB2 řešeno příliš koncepčně, pro jejich získání tak bylo často potřeba použít odlišné postupy než u zbytku databáze.  Protože výsledek práce měl být přímo použitelný jako součást již existujícího softwarového systému, byl autor nucen aktivně sbírat technické požadavky a omezení kladená tímto systémem a následně se s nimi vypořádat.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 4:</b>
<b>2. Splnění zadání</b>	<b>1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</b>
<b>Popis kritéria:</b> Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
<b>Komentář:</b> Práce splnila zadání.  V práci mi chybí srovnání struktury databáze s databázemi Oracle a Microsoft SQL Server, našel jsem pouze srovnání způsobů získání metadat (viz kapitolu 2.4).  Implementované prototypové řešení je dostatečně kvalitní a splňuje všechny technické požadavky softwarového systému Manta, je tak možné jej přímo využít v praxi.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - následující škálou 1 až 4:</b>
<b>3. Rozsah písemné zprávy</b>	<b>1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky</b>
<b>Popis kritéria:</b> Zhodnotte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
<b>Komentář:</b> Rozsah práce je přiměřený. Práce je informačně bohatá.	
<b>Hodnotící kritérium:</b>	<b>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</b>
<b>4. Věcná a logická úroveň práce</b>	<b>80 (B)</b>

**Popis kritéria:**

Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře.

**Komentář:**

Práce je dobře strukturovaná a srozumitelná.

Informace o možných způsobech extrakce v kapitole 3 jsou poněkud nepřehledné, velmi by pomohla tabulka shrnující pro každý typ objektu možné způsoby extrakce metadat a DDL skriptu.

Popis implementace rozhraní Db2DependencyManager v kapitole 6.5 je poněkud zmatečný. Dle mého názoru je to důsledkem špatně rozmyšlené a zbytečně komplikované implementace:

- Popisovaný "graf závislostí" je ve skutečnosti pouze mapa již seřazených objektů; graf závislostí je reálně držen v seznamu "notSortedList".
- Velká složitost implementace jde na vrub práce s touto mapou - ve skutečnosti třemi do sebe vnořenými mapami. Místo takové struktury lze však ekvivalentně (a výrazně snáze) použít jednu mapu s kompozitním klíčem.
- Hodnota v mapě není nikde dále využita (jen informace o její přítomnosti v mapě), takže by šla použít místo mapy množina (Set).
- Domnívám se, že jako klíč pro tuto množinu by se dala snadno využít přímo hodnota nyní ukládaná do mapy jako hodnota (typ DependentObject) - v takovém případě by bylo možné uvažovat o možnosti sloučit seznam seřazených objektů a množinu již seřazených objektů (např. použitím LinkedHashSet).
- Zvolené řešení má kvadratickou časovou složitost. Jde přitom o standardní problém topologického uspořádání, který je řešitelný v lineárním čase.

**Hodnotící kritérium:**

*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*

**5. Formální úroveň práce**

95 (A)

**Popis kritéria:**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.

**Komentář:**

Typografická a stylistická úroveň práce je vyhovující. Práce obsahuje jen malé množství gramatických chyb.

**Hodnotící kritérium:**

*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*

**6. Práce se zdroji**

100 (A)

**Popis kritéria:**

Vyjáďřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

**Komentář:**

V pořádku.

**Hodnotící kritérium:**

*Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):*

**7. Hodnocení výsledků, publikační výstupy a ocenění**

95 (A)

**Popis kritéria:**

Vyjáďřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

**Komentář:**

Práce popisuje málo zdokumentovanou oblast získávání metadat z databáze DB2. Popisuje mj. i funkcionální nepopsanou v oficiální dokumentaci.

Implementované řešení prokazuje teoretickou správnost i praktickou použitelnost navržených postupů.

Vytvořený prototyp je dobře navržený, zohledňuje všechny funkční i nefunkční požadavky.

**Hodnotící kritérium:**

*Způsob hodnocení - nehodnotí se*

**8. Komentář o využitelnosti výsledků**

**Popis kritéria:**

Uvedte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uvedte možnosti využití výsledků ZP v praxi.

**Komentář:**

Výsledky práce jsou přímo uplatnitelné v praxi.

Vzhledem ke kvalitě prototypové implementace a dodržení technických požadavků systému Manta lze bez větších problémů nejen použít teoretických závěrů, ale bude i možné implementovaný prototyp převzít jako základ pro praktické řešení.

**Hodnotící kritérium:**

*Způsob hodnocení - nehodnotí se*

**9. Otázky k obhajobě**

**Popis kritéria:**

Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřázkami).

**Otázky:**

- Podařilo se získat všechna potřebná metadata? Pokud ne, která metadata se nepodařilo získat?
- V kapitole 5.3 uvádíte: "Aplikační server není blíže specifikován, nástroje na něm nejsou závislé (díky použití frameworku Spring)." - jakým způsobem použití Springu implikuje nezávislost na aplikačním serveru?
- V kapitole 3.4 uvádíte, že v katalogu se v některých případech objevují chyby. Lze tyto chyby nějak automaticky detekovat a případně obejít?

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů  
(známka A až F):

**10. Celkové hodnocení**

90 (A)

**Popis kritéria:**

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nesmí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

**Text hodnocení:**

Práce řeší všechny podstatné aspekty problematiky získání metadat o objektech z databáze DB2. Analyzuje několik možných postupů.

Autor se musel při analýze často spolehnout na vlastní experimenty s nedokumentovanými funkcemi databáze. Musel se též vypořádat s nekoncepčním uložením metadat o objektech v případě použití DB2 v módu kompatibility s databází Oracle.

V písemné zprávě chybí důkladnější porovnání hierarchické struktury objektů v databázi s ostatními databázovými systémy, toto porovnání je přitom potřebné pro správné namapování objektů na obecný model cílového softwarového systému. Implementované řešení používá vhodné mapování, autor tak zřejmě toto porovnání provedl a vyvodil správné závěry. V písemné zprávě však tato diskuze chybí.

Implementované prototypové řešení je technicky na vysoké úrovni, až na některé detaily dobře navržené a splňuje všechny na něj kladené požadavky. Bylo provedeno ověření korektnosti implementace.

Teoretické závěry práce i implementované řešení jsou přímo využitelné v praxi.

Podpis oponenta práce: