

# Oponentský posudek disertační práce

Program: *Výpočetní technika a informatika*

Jméno studenta: *Ing. et Ing. Vojtěch Mates*

Název práce: *Query-Driven Context-Aware Workflow*

## Celkové zhodnocení práce

Práce se zabývá problematikou systémů pro řízení pracovních toků (workflow) a podnikových procesů, které jsou schopny se adaptovat v závislosti na změny prostředí, jsou schopny zohlednit měnící se vlastnosti uživatelů, resp. účastníků procesu, a jsou schopny zohlednit i měnící se prostředí a cíle podniku. Téma disertační práce považuji za aktuální a odpovídající oboru disertace

Základní myšlenka využití informace o kontextu a promítnutí cílů podniku pro adaptaci a optimalizaci podnikových procesů je určitě zajímavá z hlediska výzkumu i praktického využití. V předložené práci ji autor realizuje na úrovni návrhu koncepce architektury systému pro řízení workflow (dále jen WfMS). V tom lze vidět původní přínos autora.

Práce má však řadu nedostatků. Za nejzávažnější považuji skutečnost, že návrh nebyl podle mne dostatečně ověřen. Jsou zpravidla uváděny obecně výhody navrženého řešení, ale ty nejsou nijak doloženy (experimenty, rozsáhlejší případovou studií apod.). Lze pochopit, že kompletní implementace navrženého řešení by představovala značně rozsáhlé programové dílo. Ověření postavené na pouhém obecném výčtu předpokládaných výhod a přínosů bez uvažování dopadu růstu složitosti na implementaci a použitelnost systému považuji za nedostatečné.

Za další nedostatek práce považuji, že obsah některých částí je málo srozumitelný. Důvodem je hlavně nedostatečně výstižná charakteristika popisovaného a práce s důležitými pojmy, o které se práce opírá. Pojmy jako systém řízení podnikových procesů, systém řízení workflow, sada pro řízení podnikových procesů, aktivita, úloha (task), projekt a jejich vztahy, nejsou v práci dostatečně přesně (ne nutně formálně) vymezeny. To má za následek snížení srozumitelnosti a v některých případech i dojem, že autor zachází s těmito pojmy poměrně vágně.

Některé další připomínky jsou uvedeny v podrobném hodnocení níže.

Jádro práce bylo publikováno v jedné časopisecké publikaci a třech konferenčních příspěvcích a jednom krátkém příspěvku (abstrakt). Další jedna časopisecká informace a dva konferenční příspěvky, kde je spoluautorem, s tématem disertace souvisí volněji nebo okrajově, neprezentují přímo jeho stěžejní výsledky. Podmínku publikace výsledků lze považovat za splněnou, i když v žádné z těch publikací, kde je disertant hlavním autorem, není podle mne v dostatečném rozsahu prezentováno ověření konceptu, což je moje hlavní výhrada k předložené práci.

## Podrobné hodnocení práce

Práce je členěna do 14 kapitol. První kapitolou je úvod vysvětlující motivaci a strukturu práce. Struktura je prezentována s nižší granularitou, než je seznam kapitol, ale obsah pokrývá.

Kapitoly 3 až 7 představují shrnutí současných přístupů k řešení problematiky relevantní k zaměření disertace. Každá z nich obsahuje i shrnutí a diskusi, kde autor promítá v některých případech své hodnocení. V kapitole 2 se zabývá informačními systémy zohledňujícími podnikové procesy (process-aware). Zde autor zavádí pojem sada pro řízení podnikových procesů (business proces management

suite) a stručně charakterizuje několik konkrétních existujících řešení. Také zde stručně vysvětluje vztah pojmů podnikový proces, jeho model a instance.

Kapitola 3 je věnována systémům pro řízení workflow. Autor zde především prezentuje referenční model systému pro řízení workflow, ze kterého ve své práci vychází. V obr. 3.1 referenčního modelu chybí obdélník reprezentující relevantní data workflow.

Kapitola 4 pojednává o modelování podnikových procesů, přesněji o některých existujících přístupech. Ty jsou ale charakterizovány, pokud vůbec, velice stručně, většinou ani podstata vysvětlena není, spíše motivace vzniku a výhody, které přináší. Přínos pro čtenáře je proto velmi omezený. Například obr. 4.1 není nikde odkazovaný a ani k němu není v textu žádný komentář. Ve srovnání s předchozími kapitolami je zde podstatně delší shrnutí a diskuse, kde autor již uvádí některé rysy vlastního řešení popsaného až v dalších kapitolách a prezentuje ve značně obecné rovině některá doporučení. Tato část je proto k pochopení obtížná.

Kapitoly 5 a 6 se věnují modelování podnikových procesů. V kapitole 5 ukazuje autor formální přístup k vysvětlení sémantiky některých základních prvků používaných při modelování podnikových procesů. Nejde ale o přístup vlastní, nýbrž převzatý z použité literatury. Jsou zde uvedeny formální definice procesu (nesprávně označena definice 5.1.1 jako definice meta-modelu), instance procesu a sémantika některých základních vzorů toku řízení. Jde spíše o ukázkou možné formalizace, která není v práci dále nijak využita. Zmíněno je i použití Petriho sítí pro modelování, z nichž formalismy většinou vycházejí. Vzhledem k tomu, že jsou definice a další formální zápisy převzaté z pramene, kde je k tomu uvedeno podstatně více, ne vždy je vše srozumitelné. Například diagramy událostí uvedené v obrázcích nejsou vysvětleny a není tak potom na první pohled zřejmé, proč jsou některé události u formulí pro vzory řízení označovány písmenem *e* a jiné *t*.

V kapitole 6 se autor zaměřuje na nejčastěji používané jazyky pro modelování podnikových procesů. Opět je popis v některých případech natolik stručný a nemá podobu výstižné charakteristiky, že nepřináší pro práci prakticky nic pozitivního. Očekával bych v této kapitole také výraznější charakteristiku prezentovaných jazyků z hlediska přístupů k modelování uvedených v kap. 4. Ten je zmíněn jen u systému Declare. Na str. 29 je odkazována nějaká tabulka, která ale v práci nikde není.

Kapitola 7 si všímá přístupů k analýze a optimalizaci podnikových procesů. Představuje stručný úvod do dvou oblastí – kritérií/indikátorů hodnocení výkonnosti a přístupů k analýze a optimalizaci, včetně využití dolování v datech získaných z běhů procesů.

V kapitole 8 autor shrnuje některé nedostatky existujících přístupů k modelování podnikových procesů a systémů pro jejich řízení a vymezuje zaměření vlastního přínosu prezentovaného v dalších kapitolách. Hlavní cíl disertace zde bohužel není specifikován explicitně. Jako jeden z cílů se uvádí nalezení způsobu orchestrace různých přístupů k modelování a řešení problému flexibility za běhu. Vzhledem k tomu, že autor zde specifikuje požadavky na podle jeho názoru ideální architekturu „stroje pro podnikové procesy (business system engine)“ a následně prezentuje vlastní návrh architektury systému pro řízení workflow, lze usuzovat, že právě návrh takové architektury byl hlavním cílem disertace. Architektura vychází z referenčního modelu uvedeného v kap. 3. Ukazuje ji obr. 8.1 na str. 39. Lze mu sice vytknout některé nedostatky (chybějící hrana, orientace či popis hrany, opačná orientace hrany mezi Proces Orchestrator a Dynamic Process Profiles nebo absence odlišení softwarových komponent, řídicích dat systému a externích produktů/dat snižující srozumitelnost), ale celkově dává představu o navržené koncepci.

Podrobněji jsou navržená rozšíření a zpřesnění referenčního modelu zdůvodněna a naznačena možnost jejich implementace v kapitolách 9 až 12. Stručná kapitola 9 zmiňuje zejména některá zjištění výzkumu zaměřeného na dolování v datech sbíraných při běhu podnikových procesů, která ukazují potenciál využití získaných znalostí pro optimalizaci procesů. Jde spíše o motivační úvod. Je zde i zmínka o dolování v historických datech sbíraných v průběhu výrobního procesu, jehož cílem bylo potvrdit, že taková data (zde zejména čas vykonání operace) lze využít ke zpřesnění stanovených standardizovaných hodnot.

Kapitola 10 se týká tzv. dynamických objektů. Jde spíše o obecné povídání, samotný pojem dynamického objektu není nijak přesněji vymezen, pojem kontextu, který je pro disertaci stěžejní, je vymezen poměrně vágně právě pomocí dynamických objektů.

Kapitola 11 se zaměřuje na složku kontextu, která souvisí s uživateli a jejich zkušenostmi. I zde jde v nezanedbatelné míře o značně obecné povídání týkající se nejen systému pro řízení workflow, ale dokonce sady (suite) pro řízení podnikových procesů. Za konkrétnější lze považovat zavedení profilů uživatelů, rozšíření (v práci ale jen zmíněno) ohodnocení hran formou podmínek (guard) a funkcí a naznačení možnosti modifikace podnikového procesu na základě profilu využitím vysokoúrovňových Petriho sítí (Hi-level Petri net). Použité příklady jsou triviální, možná implementace je naznačena velmi obecně, spíše jsou prezentovány očekávané výhody a přínosy.

Kapitola 12 pojednává o optimalizaci podnikových procesů generováním definice procesu na základě dat reprezentujících kontext, informací o poskytovaných vstupech a poskytovaných výstupech dílčích procesů a omezeních vyplývajících z cílů podniku. Tato kapitola poskytuje oproti předchozím kapitolám k vlastnímu řešení výrazně konkrétnější představu o navrhované koncepci. I zde se ale vyskytuje řada nedostatků, například:

- Pojem strom závislosti produktů (product dependency tree , dále jen PDT) není pořádně vysvětlen. Obr. 12. 1 je bez bližšího komentáře, v něm jsou některé uzly označeny jako (Sub)Product, jiné jako Subprocess.
- Terminologie: Připouští-li se více finálních produktů, vznikne les stromů nebo musí být nějaký zastřešující kořen (str. 64), ne více kořenů stromu.
- Zápis algoritmu na str. 64 je nevhodný (proč ne raději pseudokód?), nesprávný (přinejmenším zápis kroku 4) a není zřejmá nutnost vyloučení duplicit (duplicitní instance se asi vyskytovat mohou).
- Sémantika dotazu zapsaného v kódu blízkém SQL není úplně jasná (viz otázka k obhajobě níže).
- V kapitole 12.10 nazvané Experiments, kde by čtenář očekával popis způsobu a výsledky ověření konceptu architektury nebo alespoň PDT, je pouze obecně uvedeno, že koncept PDT je testován na dynamickém výběru nevhodnějšího dodavatele.

Kapitola 13 naznačuje možnost realizace navržené koncepce v distribuovaném prostředí SOA. Vzhledem k tomu, že se jedná až na první odstavec plnou kopii článku [30], opakuje se zde řada informací obsažených v předchozích kapitolách (např. PDT a dotazování). To nepovažuji za vhodnou formu. Nově je zde prezentován konceptuální model, resp. metamodel, představující jiný pohled na proces, než byl prezentovaný v kap. 5. Zde jde o definici, realizaci, implementaci a výkonnost procesu. PDT je tady prezentovaný v částečně formalizované podobě (jako graf). Mohlo by se zdát, že tato kapitola je tou, která má ukazovat uskutečnitelnost navržené koncepce. V podkapitole 13.5

nazvané Evaluace a Diskuse je prostřednictvím odkazu na jednu z publikací autora zmíněn nějaký implementovaný WfMS. O tom ale v předložené práci bližší informaci postrádám.

Kapitola 14 je závěrem, kde autor charakterizuje publikace obsahující výsledky prezentované v disertační práci a dále svůj pohled na přínos navržené architektury systému pro řízení podnikových procesů.

Bibliografické citace elektronických zdrojů dostupných na Internetu považuji za neúplné – chybí URL adresa.

### **Otázky k obhajobě**

1. Jakým způsobem a v jakém rozsahu byl navržený koncept architektury ověřen?
2. Podle popisu na str. 63 se vyhledává optimální řešení v již vytvořeném PDT. Dotaz na str. 66 vybírá z databáze projektů. Jakých hodnot nabývá proměnná  $p$  v dotazu? Co se vybere? Položky *term* a *cost* profilu obsahují hodnoty za celý proces včetně podprocesů? Pokud ano, pro jakou sestavu nahraditelných podprocesů, pokud ne, kde se agreguje?

### **Závěr**

Autor podle mého názoru sice ve značné míře prokázal schopnost samostatné tvůrčí práce, ale předložená disertační práce má podle mne některé výše uvedené závažné nedostatky. Proto doporučuji titul Ph.D. neudělit.

1. 8. 2016

Jaroslav Zendulka

FIT VUT v Brně