

# IDENTIFIKACE DATOVÉHO OBSAHU PRO VÝVOJ DATOVÉHO MODELU

## IDENTIFICATION OF DATA CONTENT FOR DEVELOPMENT OF DATA MODEL

**Stanislava Šimonová, Martin Kořínek**

**Abstract:** *Each organization or company wants to reach good output in a long-term, whether output products are commodities or services. The goals are improving quality and efficiency of production. Concept and norms of quality feature recommendation, of which process approach and efficient information / data administration take significant place. Business processes and business data constitute linked unit, because business processes need relevant data for work and business data should fully serve to business processes. Similarly, linked unit is constituted by process and data modeling, which needs to be done by cooperation of processes actors and data users, thus by cooperation of managers. Role of managers is un-substitutable, for they create requirements for improving and for higher quality via their models. The main topic of this paper is identification of business data content with a view to management.*

**Keywords:** *Data Content, Quality of Data, Data Model, Information System.*

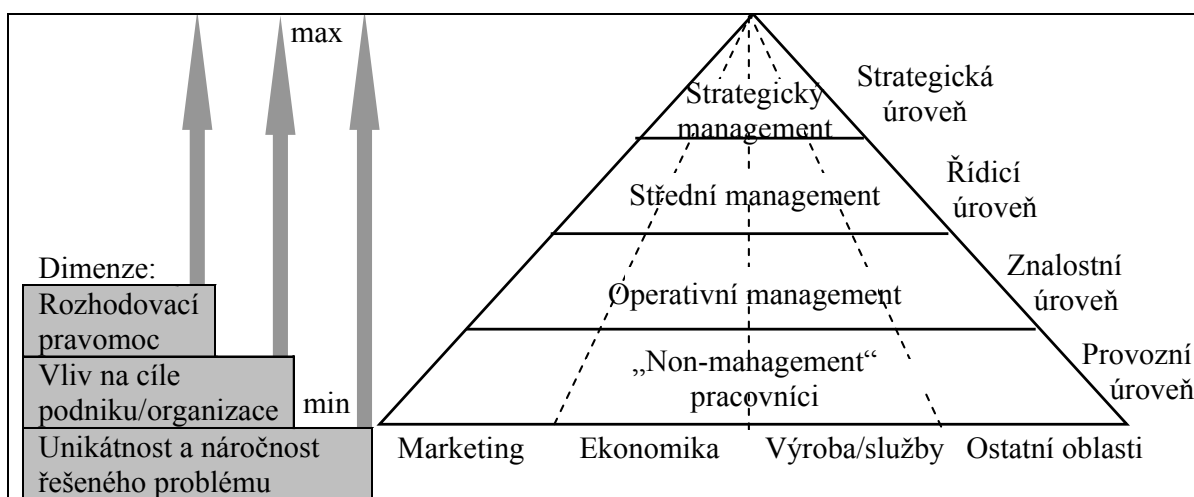
**JEL Classification:** *L15, M11, M15.*

### 1 Úvod

Řízení podnikových / podnikatelských procesů probíhá na základě rozhodování a působení managementu. Manažeři uplatňují své znalosti a zkušenosti, zároveň využívají i nezbytné podpory, ke kterým patří informační prostředí firmy. To je tvořeno informačním systémem, resp. několika propojenými informačními systémy. Ty jsou jednak využívány pro každodenní chod podniku, jednak poskytují podklady pro strategické rozhodování ve smyslu dlouhodobého směřování podniku. Informační systém je realizován sadou propojených informačních technologií, které se společně podílí na zpracování, uchování, vybírání, a distribuci informací [10] a slouží k řízení podnikových procesů, manažerskému rozhodování a správě podnikové agendy [18]. Aktualizace informačního systému je vyvolána výhradně procesní změnou [21], tzn. změnou ve výkonu daných podnikatelských aktivit. Různé organizační úrovně podniku vyžadují specifický druh informací a specifický způsob zpracování informací. Obrázek č. 1 vyjadřuje organizační a informační pyramidu organizace; tvar pyramidy demonstruje objem a návaznost dat a potřebu určitého typu informací napříč organizačními úrovněmi.

Na každé organizační úrovni by si manažeři měli klást otázky [20]– Má podnikový informační systém(y) všechny ty agendy, které bychom potřebovali pro svou práci? Je práce s informačním systémem uživatelsky přijatelná? Pracuje informační systém

v „naší logice“, tedy můžeme v něm postupovat v takovém sledu návazných činností, které vyhovují nám? Lze se na data spolehnout, tedy máme jistotu, že data jsou naprosto správná a že nejsou zastaralá? Dostáváme na výstup takové shluky dat (ve smyslu rozsahu dat či podnikové tematiky), které potřebujeme pro svoji činnost? Jsou informační systémy podporou při naplňování a dosahování podnikových cílů? Jsou data z informačních systémů podporou pro průběh a výkon podnikatelských procesů?



**Obr. 1: Organizační a informační pyramida**

*Zdroj:[18][19][20]; vlastní zpracování*

Podniková data mají plně sloužit k podpoře podnikatelských procesů a k podpoře manažerského rozhodování, tzn. podnikové informační systémy musí být vyvíjeny výhradně v souladu s těmito požadavky.

## 2 Formulace problematiky

Složitá problematika nakládání s podnikovými daty je dána situací, že podnik produkuje data pro každodenní podporu své podnikatelské činnosti, nicméně tato data nejsou vždy uživatelsky relevantní (např. nemají vhodnou návaznost či vhodnou míru detailizace či agregace aj.). Další problémovou oblastí je, že obrovské objemy vyprodukovaných podnikových dat nejsou dostatečně využívány pro podporu rozhodování a pro manažerskou predikci. V souvislosti s tendencí přesunu od kapitálových strategických zdrojů k informačním zdrojům vyvstává nutnost, rozvíjet a lépe využívat ostatní zdroje v podnikatelském subjektu; tzn. je zdůrazňován význam informací a informačních zdrojů pro podnik a nutnost rozvoje intelektuálního kapitálu [13]. Různé propracované přístupy či modely či rámce podnikové informatiky se snaží přispět k řešení této problematiky.

### 2.1 Přístupy podporující využitelnost a relevantnost podnikových dat

Přístup Business Intelligence zahrnuje sadu procesů, aplikací a technologií, jejichž cílem je účinně a účelově podporovat řídicí aktivity ve firmě. Aplikace Business Intelligence jsou postaveny na principech multidimenzionálních pohledů na podniková

data, podporují analytické a plánovací činnosti organizací, přičemž pokrývají analytické a plánovací funkce většiny oblastí podnikového řízení, tj. prodeje, nákupu, marketingu, finančního řízení, controllingu, majetku, řízení lidských zdrojů, výroby, IS/ICT aj. [14]. Další přístup, Competitive Intelligence, lze vyjádřit jako konkurenční zpravodajství, je definován jako cílevědomý postup pro získávání, analyzování a řízení externích informací, které mohou mít vliv na plánování, rozhodování a fungování organizace [2]; smyslem je nesoustředit se pouze na podnikové interní a archivované informace nebo informace volně dostupné na webu. Přístup Computer Intelligence nevyjadřuje zavedený termín ve smyslu Business Intelligence, ale navozuje téma kvality výpočetní platformy, resp. zapojení vědeckých disciplín do výpočetních postupů [1]. Jedná se o uplatnění SW nástrojů podporující oblast rozhodování a řízení podniku, které aplikují vědecké postupy výpočetní inteligence. Např. metody umělé inteligence jsou využívány na řešení takových problémů, kde cesty řešení pro všechna možná řešení nejsou předvídatelné a kde jsou k dispozici neurčité, neúplné a velmi rozdílné strukturované údaje a informace [16]. Řízení podnikové informatiky (řízení celku i řízení dílčích částí) bývá uplatňováno za podpory určitého rámce či modelu, mezi něž patří standardy ITIL a COBIT. ITIL je rámcem spíše pro IT management podniku, kterému nabízí návody, šablony, diagramy a další osvědčené postupy pro řízení IT služeb [3] [8] [22]. COBIT je rámcem pro IT Governance [7], jedná se o strategický rámec pro řízení IT prostředí s cílem zajistit soulad mezi řízením podnikové informatiky a mezi cíli a řízením podnikové činnosti. Uvedené přístupy či rámce řeší spíše management IT procesů ve vztahu k podpoře podnikatelských procesů a k naplňování podnikových cílů. To je pochopitelně významné, neboť kvalita informačního systému je dána mírou, kterou informační systém přispívá k výkonnosti a efektivnosti podnikových procesů, činností a jednotlivých uživatelů [12].

Uvedené přístupy a rámce se však zabývají zejména manipulací s daty, nicméně neřeší základní vymezení vstupního datového obsahu, s kterým až poté může být manipulováno efektivně či méně efektivně. Vymezení datového obsahu souvisí s vlastním vývojem informačního systému. Vývoj informačního systému je svým rozsahem, složitostí a speciálností rozsáhlým projektem. Jsou definovány různé metodiky či modely pro postup vývoje informačního systému jako např. model kaskádový, prototypový, spirálový aj. [6] [11], nicméně z hlediska pohledu managementu lze vymezit tři základní etapy pro informační systém – příprava, zavádění a provozování. Pro úspěšnost takového projektu je důležité věnovat velkou pozornost etapě přípravy, zahrnující fáze jako plánování strategie, analýzu, prvotní návrh aj. Chyby z těchto fází se projeví až později a jejich odstranění je obvykle obtížné a nákladné [22].

Základní vymezení datového obsahu je tedy primárním požadavkem. Na formulaci tohoto datového obsahu se však musí podílet budoucí / stávající správci dat a uživatelé dat, tedy pracovníci managementu. Jenom oni přesně vědí a mohou definovat, jaké činnosti v jaké návaznosti probíhají, tedy jaká data k těmto činnostem potřebují, a to konkrétně jaká data, v jakých návaznostech, v jaké formě aj. Tedy jenom oni mohou dodat relevantní požadavky pro vlastní vývoj datového modelu.

### 3 Řešení problému

#### 3.1 Vývoj datového modelu

Datový model je v rámci podnikové praxe téměř výhradně realizován relačním modelem dat. Základním konstruktem je relace. Schéma pro relaci [4] [17] pojmenovanou R lze vyjádřit

$$R (A1:D1, A2:D2, \dots, An:Dn) \quad (1)$$

kde  $D_i = \text{dom}(A_i)$ , pro  $i \in \langle 1, n \rangle$ ; A je atribut; D je doména atributu.

Dom je zobrazení definované na množině jmen atributů, kdy je k atributu přiřazena jeho doména; doména je specifikovaná množina hodnot, kterou daný atribut nabývá.

Relační model dat dále zahrnuje integritní omezení, jež se vyjadřují jako podmínky, které mají být na datech v databázi splněny. Relační schéma databáze je dvojice (R, I), kde R je množina schémat relací a I je množina integritních omezení. Integritní omezení lze rozlišovat jako lokální omezení týkající se jednotlivých schémat relací (tj.  $I_i$  pro  $i \in \langle 1, K \rangle$ ), a jako globální omezení (I) udávající vazby mezi daty z různých relací. Pak lze psát [17]:

$$R = \{ (R1:I1), \dots, (RK:IK); I \} \quad (2)$$

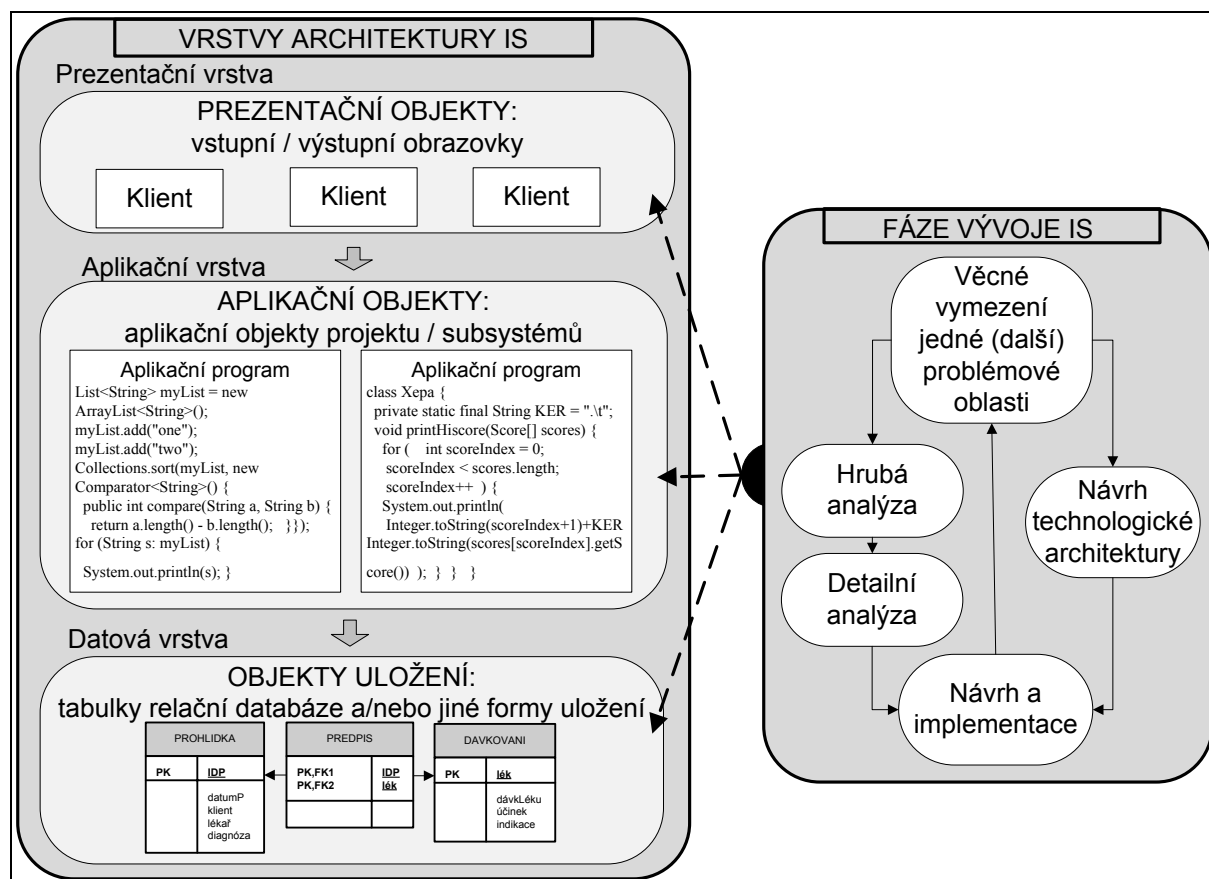
Návrh relačního modelu dat je vytvořen na základě transformace analytických modelů v rámci procesu datového modelování. Datové modelování má přístupy – strukturovaný přístup a objektově orientovaný přístup [4] [5] [15] [20]. Strukturovaný přístup je orientovaný přímo na potřeby tvorby databáze, resp. je cílený k návrhu a implementaci informačního systému na technologii databázového systému; analýza tímto přístupem je realizována s využitím diagramu typu Entity Relationship Diagram. Objektově orientovaný přístup se opírá o standard UML (Unified Modeling Language), který pomáhá specifikovat, vizualizovat a dokumentovat modely návrhu aplikací a architektur, ale také datových struktur a podnikových procesů. Jedná se o obecný modelovací jazyk, který obsahuje specifikaci základní sady diagramů včetně možnosti jejich rozšíření. Analýza tímto přístupem je realizována s využitím diagramů – diagram případu použití (Use Case), sekvenční diagram (Sequence Diagram), diagram tříd (Class Diagram) aj. Tyto přístupy datového modelování jsou uplatněny v rámci některé z metodik, které použije vývojový tým (ať už vlastní projektový / IT tým nebo tým v rámci outsourcingu). Součástí metodik je identifikace integritních omezení ve všech fázích vývoje datového modelu.

Použití uvedených přístupů a nástrojů však vyžaduje adekvátní odbornost a zručnost v problematice datového modelování i při využití vhodných IT case nástrojů. Tato odbornost je samozřejmá u členů projektového týmu, nicméně nelze tuto odbornost delegovat na stávající / budoucí tvůrce dat a uživatele dat. Jenže právě oni musí zásadním způsobem vymezit datový obsah, přičemž toto vymezení se pak stává základní součástí integritních omezení.

### 3.2 Vymezení datového obsahu pro identifikaci integritních omezení

Vymezení datového obsahu je podkladem pro definici integritních omezení, která se stanou součástí vyvíjeného relačního modelu dat. Vymezení datového obsahu je tedy určitým požadavkovým modelováním, které je pak zpracováno či transformováno do integritních omezení pro datové modelování. K určení nástrojů pro vymezení datového obsahu je třeba vycházet charakteristik:

- Identifikace datového obsahu je tvořena definicí / popisem všech datových souvislostí ve vztahu k vykonávaným podnikatelským aktivitám.
- Tyto datové souvislosti znají pouze aktéři, kteří související podnikatelské aktivity buď přímo sami vykonávají nebo jsou odpovědní za výstupy těchto aktivit; aktéři jsou tedy tvůrci dat a/nebo uživatelé dat.
- Aktéři vymezením datového obsahu zásadním způsobem ovlivní kvalitu datového modelu, tzn. kvalitu těch fází vývoje informačního systému, které již zpracovávají projektoví / IT odborníci.
- Na aktéry nemohou být delegovány požadavky na znalosti a zručnosti spojené s konkrétními vývojovými metodami a technikami.
- Aktéři definují vymezení datového obsahu formou modelů, a to grafických a textových.



Obr. 2: Vrstvy architektury informačního systému a fáze vývoje systému

Zdroj: [9][20]; vlastní zpracování

Vývoj datového modelování systému probíhá se zohledněním jednotlivých vrstev architektury systému (viz obrázek 2). Architektura informačního systému je vyjádřena jako součinnost tří vrstev – datová, aplikační a prezentační. Datová vrstva je tvořena uloženými daty nejčastěji v propojených tabulkách relačního databázového systému založeného na relačním modelu dat. Aplikační vrstva je tvořena aplikacemi, které pracují s uloženými daty datové vrstvy. Prezentační vrstva pak zde znamená formu komunikace uživatele se systémem prostřednictvím rozhraní. Fáze vývoje životního cyklu informačního systému jsou zde koncipovány přístupem přírůstkovým a iterativním, kdy i uvnitř jednotlivých etap jsou modely (posléze i programový kód) vytvářeny iterativním přístupem; tzn. systém je vyvíjen po přírůstcích. Šipky naznačují, že vývoj systému jako celek (resp. každá z jeho fází) má vliv na funkčnost a správnost všech vrstev architektury systému. Všechny fáze vývoje systému vyžadují komunikaci a spolupráci mezi IT dodavateli (analytici, návrháři, programátoři) a managementem (uživatelé dat).

Vymezení datového obsahu by mělo zahrnovat oblasti:

- Identifikace datových funkcionalit, které mají být k dispozici pro podporu podnikatelských aktivit.
- Charakteristika těchto datových funkcionalit.
- Identifikace datových objektů v rámci žádaných datových funkcionalit.
- Charakteristika vlastností těchto datových objektů.
- Identifikace vazeb mezi těmito datovými objekty.
- Charakteristika těchto vazeb.

Výchozí charakteristiky pro metody a techniky k vymezení datového obsahu:

- Základním požadavkem je jednoduchost syntaxe, srozumitelnost pro aktéra a rychlá aplikovatelnost.
- Nástroje mohou být použity a kombinovány z různých metodik datového modelování.
- Použití techniky a nástrojů musí poskytnout výstup, který bude použitelným vstupem do datového modelování na úrovni analýzy i návrhu informačního systému.

Nástroje pro vymezení datového obsahu:

- Identifikace datových funkcionalit, které mají být k dispozici pro podporu podnikatelských aktivit:
  - Základní vymezení žádaných datových funkcionalit by mělo vycházet z existujících podnikových procesních map.
  - Diagram případů užití (use case) ze standardu UML je srozumitelným a rychle uchopitelným nástrojem pro vyjádření identifikovaných datových funkcionalit.
- Charakteristika těchto datových funkcionalit:
  - Scénář případu užití ze standardu UML lze užít ve formě textové tabulky. Aktér vyjádří ve scénáři nutné vstupní podmínky, žádaný výstup a zejména sled návazností, v jakém má být pracováno s danou datovou funkcionalitou.
- Identifikace datových objektů v rámci žádaných datových funkcionalit:

- Východiskem je scénář případu užití (vytvořený u předchozího bodu) ze standardu UML, který obsahuje sled činností a výčet objektů. Aktér v textu scénáře nalézá datové objekty.
- Charakteristika vlastností těchto datových objektů:
  - Aktér pro datový objekt určí: datový typ, doménu, identifikovatelnost (identifikátor), podmínku platnosti, nutnost vyplnění, výchozí hodnotu.
- Identifikace vazeb mezi těmito datovými objekty:
  - Diagram entitně-relační ze strukturovaného přístupu je srozumitelným nástrojem pro vyjádření vazeb mezi datovými objekty. Forma vyjádření může být grafická a/nebo textová.
- Charakteristika těchto vazeb:
  - Vyjádření kardinality a parciality v souladu se strukturovaným přístupem. Aktér určí, zda identifikovaná vazba vždy nastat musí či nemusí (parcialita) a určí četnost vztahu (kardinalita).

Aktéři takto definují své požadavky na požadovaný datový obsah. Vyjádřené požadavky se stávají integritními omezeními pro vývojový tým, který s nimi dále pracuje jak na úrovni analýzy, tak na úrovni návrhu v rámci vývoje podnikového informačního systému, resp. v rámci vývoje jeho datového modelu. Takto vyvíjené požadavky mohou aktéři dále charakterizovat a detailizovat, tak jak vývojový tým postupuje od hrubé analýzy k detailní analýze, od analýzy k návrhu, resp. tak jak v rámci iterativního postupu dochází k propracovanosti datového modelu.

#### 4 Závěr

Řízení podnikových procesů, resp. řízení organizace jako celku musí využívat podporu informačního prostředí. To je tvořeno podnikovými informačními systémy, které poskytují určitou formu práce s podnikovými daty – data jsou jednak vytvářena v souladu s podnikatelskou činností firmy (podnikové procesy), dále jsou data využívána pro manažerské procesy. S daty pracují tvůrci dat a uživatelé dat, tzn. pracovníci managementu. Informační systémy poskytují přístupy k datům v různé formě, v různé návaznosti, v různé stupni detailizace či agregace aj. Kvalitu formy / způsobu / obsahu poskytovaných dat pak mohou ovlivnit pouze jejich tvůrci a uživatelé. Pouze tito aktéři znají, jaké vykonávají aktivity v rámci podnikových procesů a jaká data k těmto aktivitám potřebují, tzn. v jakém sledu, v jakém tvaru, jakém detailu, s jakými kontrolami aj. Proto pouze aktéři mohou definovat žádaný datový obsah. Přímé zapojení aktérů do vývoje datového modelu však není možné, protože není únosné na ně delegovat odbornost spojenou s vývojem datového modelu, resp. informačního systému. Tato odbornost je příslušná členům vývojového týmu (ať je řešena vlastním IT nebo outsourcingem) a představuje sadu specializovaných znalostí a zručností, jako využívání a znalost vhodných metod a technik, práci s IT modelovacími nástroji typu case aj. Nástroje aktérů musí být jednoduché a rychle aplikovatelné, nesmí být omezeny dodržováním konkrétních metodik, musí poskytnout takový výstup, který bude použitelným vstupem do datového modelování na úrovni analýzy i návrhu informačního systému. Vymezení datového obsahu by mělo zahrnovat oblasti – identifikace datových funkcionalit (které mají být k dispozici pro

podporu podnikatelských aktivit), charakteristika těchto datových funkcionalit, identifikace datových objektů v rámci žádaných datových funkcionalit, charakteristika vlastností těchto datových objektů, identifikace vazeb mezi těmito datovými objekty a charakteristika těchto vazeb. Vymezení datového obsahu ze strany manažerů je podkladem pro definici integritních omezení, která se stanou součástí datového modelu (relačního modelu dat); přičemž vlastní vývoj datového modelu zakončený implementací a provozováním informačního systému je již záležitostí specializovaného týmu. Manažeři svým vymezením datového obsahu přesně charakterizují své požadavky na data, která potřebují k výkonu svých podnikatelských aktivit; jedná se tedy o určité požadavkové modelování, které je pak zpracováno či transformováno do integritních omezení pro datové modelování. Jen tímto způsobem získají manažeři nástroje ve formě informačních systémů, které jim poskytují data relevantní pro podnikové / manažerské procesy.

### Použité zdroje

- [1] ALTER, S. *Information Systems*. Upper Saddle River: Prentice Hall, 2002. 286 s. ISBN 0-8493-1661-8.
- [2] BASL, J., BLAŽÍČEK, R. *Podnikové informační systémy: podnik v informační společnosti*. Praha: Grada, 2008. 283 s. ISBN 978-80-247-2279-5.
- [3] CARTLIDGE, A. a kol. (překlad HUDEC, J.). *Úvodní přehled ITIL V3*. Praha: Hewlett-Packard s.r.o., 2007. 56 s. ISBN 0-95551245-8-1.
- [4] DATE, C. J. *An Introduction to Database Systems*. Boston: Addison-Wesley, 2004. 983 s. ISBN 0-321-19784-4.
- [5] ERIKSSON, H., PENKER, M., LYONS, B., FADO, D. *UML 2 Toolkit*. Indianapolis: Wiley, 2004. 511 s. ISBN 0-471-46361-2.
- [6] GÁLA, L., POUR, J., TOMAN, P. *Podniková informatika*. Praha: Grada, 2006. 484 s. ISBN 80-1278-4.
- [7] IT Governance Institute. *COBIT 4.1*. USA: IT Governance Institute, 2007. 213 s. ISBN 1-933284-72-2.
- [8] *ITIL ®- IT Infrastructure Library* [online]. [cit. 2011-04-25]. Dostupný z WWW < <http://www.itsil-officialsite.com/AboutITIL/WhatisITIL.asp> >.
- [9] LBMS. Materiály firmy LBMS, s.r.o. *Objektová analýza a návrh v UML*. Praha. 2007.
- [10] MALAGA, R. *Information Systems Technology*. New Jersey: Prentice Hall, 2005. 386 s. ISBN 0-13-049750-9.
- [11] MERUNKA, V. *Datové modelování*. Praha: Alfa, 2006. 176 s. ISBN 80-86851-54-0.
- [12] MOLNÁR, Z. *Efektivnost informačních systémů*. Praha: Grada, 2000. 142 s. ISBN 80-7169-410-X.
- [13] MYŠKOVÁ, R. *Informace v rozhodování manažera*. Habilitation Theses. Zlín: Univerzita T. Bati ve Zlíně, 2007, 41 s., ISBN 978-80-7318-522-0.



- [14] NOVOTNÝ, O., POUR, J., SLÁNSKÝ, D. *Business Intelligence: Jak využít bohatství ve vašich datech*. Praha: Grada, 2005. 254 s. ISBN 80-247-1094-3.
- [15] *Object Management Group – OMG* [online]. [cit. 2011-03-12]. Dostupný z WWW <<http://www.omg.org/>>
- [16] OLEJ, V. *Modelovanie ekonomických procesov na báze výpočtovej inteligencie*. Hradec Králové: Miloš Vognar - M&V, 2003. 160 s. ISBN 80-903024-9-1.
- [17] POKORNÝ, J., HALAŠKA, I. *Databázové systémy*. Praha: Vydavatelství ČVUT, 2003. 148 s. ISBN 80-01-02789-9.
- [18] SODOMKA, P. *Informační systémy v podnikové praxi*. Brno: Computer Press, 2006. 351 s. ISBN 80-251-1200-4.
- [19] STAIR, R., REYNOLDS, G. *Principles of Information Systems*. Boston: Thomson Learning, 2006. 758 s. ISBN 0-619-21525-9.
- [20] ŠIMONOVÁ, S. *Modelování procesů a dat pro zvyšování kvality*. Pardubice: Univerzita Pardubice, edice Monografie. 193 s. ISBN 978-80-7395-205-1.
- [21] ŠIMONOVÁ, S., Svobodová, K. Klasifikace procesních změn v podmínkách diverzifikovaných kompetencí. In *Scientific Papers of the University of Pardubice Series D*. Pardubice: Univerzita Pardubice, 2010. Series D, Faculty of Economics and Administration, 16 (2010), ISSN 1211-555X, s. 297-306.
- [22] VRANA, I., RICHTA, K. *Zásady a postupy zavádění podnikových informačních systémů*. Praha: Grada, 2005. 188 s. ISBN 80-247-1103-6.

#### **Kontaktní adresa**

##### **Ing. Stanislava Šimonová, Ph.D.**

Univerzita Pardubice

Fakulta ekonomicko-správní, Ústav systémového inženýrství a informatiky

Studentská 84, 532 10 Pardubice, Česká republika

Email: Stanislava.Simonova@upce.cz

Tel.: 466 036 009

##### **Ing. Martin Kořínek, Ph.D.**

Univerzita Hradec Králové

Přírodovědecká fakulta, Katedra informatiky

Rokitanského 62, 500 03 Hradec Králové III, Česká republika

Email: Martin.Korinek@uhkk.cz

Tel.: 493 331 166

Doručeno redakci: 30. 04. 2011

Recenzováno: 15. 07. 2011

Schváleno k publikaci: 09. 08. 2011