

K čemu jsou dobré databázové návrhové vzory?

Helena Palovská,
Vysoká škola ekonomická v Praze
W. Churchilla 4
Praha 3
palovska@vse.cz

Abstrakt: Je prezentována myšlenka databázových návrhových vzorů, důvody jejího vzniku a přínosy významných autorů. Je podán přehled problematik řešených v publikovaných vzorech, je uvedeno jejich rozdělení na zárodečné modely a abstraktní vzory. Jsou sděleny zkušenosti z výuky databázových návrhových vzorů na vysoké škole. Na závěr jsou podány argumenty pro využití databázových návrhových vzorů v praxi.

Klíčová slova: datový model, návrh databáze, návrhové vzory, rapidní vývoj aplikací

Abstract: The idea of data model patterns, the cause of its emergence and contributions of important authors are presented. A survey of problems solved in published patterns is given, and their division into seed models and abstract patterns. University pedagogical experience in data model patterns learning is conveyed. Arguments for data model patterns usage in practice are put.

Keywords: data model, database design, design pattern, rapid application development

1. Úvod

Metodiky tvorby počítačových informačních systémů prošly za dobu existence tohoto oboru značným vývojem. Od neřízeného vývoje jednotlivých agend přes ANSI/SPARC model (Jardine, D. A., 1977) a následující různé více či méně formalizované a sofistikované systémy postupů (Voříšek, J. (ed.) 2008) až po agilní přístupy (Buchalceová, A., 2008). Od 70. let 20. století se ukazovalo vhodným integrovat jednotlivé agendy jedné organizace do informačního systému, s centrální správou dat.

Další zkušenost přinášela poznání, že požadavky na změnu ve funkcionalitě nějaké části takového systému často vedly k náročnému přebudovávání datové základny a problémům se zajištěním nadále funkcionality částí systému. Pro snížení nákladů a podstatných problémů s tím spojených se ukázalo vhodným usilovat o *stabilitu tzv. návrhu datové základny*, tj. modelu struktur pro ukládání dat a jejich užívání. Této stability se nejnázne dosahovalo, když model datových struktur byl nějakým způsobem *odrazem oblastí zájmu v realitě* (Gregor, J. & Chvalovský, V. 1988).

Problematika návrhu datové základny se stala samostatným oborem, se specifiky odlišujícími ho od ostatních aktivit při tvorbě a údržbě informačního systému. Podobně jako se vývoj aplikací stal jakýmsi uměním zahrnujícím mnoho sdílených dobrých zkušeností a doporučení k úspěšnému postupu, tak i při návrhu datových struktur vznikaly a vznikají znalosti, jež je možno sdílet. Taková doporučení zahrnují:

- jak postupovat při zjišťování informačních požadavků a business pravidel,
- jak kontrolovat správnost návrhu databáze formálními postupy,

- jak prezentovat model uživatelům pro ověření vecné správnosti v oblasti zájmu,
- jak a co dokumentovat,
- či jak zpřístupnit informace o stávající datové základně dalším tvůrcům.

Mezi těmito znalostmi, standardně předávanými studentům informatiky na vysokých školách či na školeních pro odborníky z praxe, jsou často implicitně předávány znalosti ještě jiné. Jde o příklady návrhu pro nějakou oblast zájmu.

Někteří autoři příklady návrhů databáze předávají v podobě nosných příkladů, na kterých ilustrují svůj výklad, jiní věnují ve svých publikacích zvláštní část několika „případům“.

Nemnoho autorů se samostatně věnuje bádání, *jak řešit návrh datové základny pro určité vybrané oblasti a jevy života*. Mezi takové autory, vycházející z dlouholeté praxe datového modelování, lze zařadit Richarda Veryarda (1992), Davida C. Haye (1996), Lena Silverstona (2001a), (2001b), (Silverston, L. & Agnew, P. 2009), Michaela Blahu (2010), ve slovenštině trojici spoluautorů Ľubora Šešeru, Alše Mičovského a Juraje Červeného (2001). Své produkty nazývají *vzory (patterns)*, podobně jako byly vzory nazvány známé konstrukty čtveřice autorů „Gang of Four“: Ericha Gammey, Richarda Helma, Ralpha Johnsona a Johna Vlissidise (1994), či jejich propagátora a následovníka Martina Fowlera (1996). V Česku se této problematice věnují Buchalceová (2004) a Pecinovský (2004).

Dnes jsou rozlišovány analytické vzory, návrhové vzory (design patterns), architektonické vzory, metodologické vzory... Pokud se jedná o vzory návrhu databáze, jsou některými z autorů nazývány *vzory modelu databáze (data model patterns)*. Protože slovem model se přitom myslí návrh datových struktur, nadále v tomto článku je budu nazývat *databázovými návrhovými vzory*.

2. Myšlenka návrhových vzorů

Myšlenku návrhových vzorů, jakožto zdokumentovaných řešení návrhových problémů v určitém oboru, přinesl architekt Christopher Alexander (Alexander, Ch. & Ishikawa, S. & Silverstein, M. & Jacobson, M. & Fiksdahl-King, I. & Angel, S. 1977). Tato myšlenka byla převzata v dalších oborech, včetně informatiky. Kromě sdílených znalostí k řešení daných problémů se za další přínos návrhových vzorů považuje *sdílený jazyk (pojmenování vzorů)*, jež mohou týmy splupracující na projektu využívat. Tak je tomu bezpochyby u známých návrhových vzorů Gang of Four (http://en.wikipedia.org/wiki/Design_Patterns#Patterns_by_Type). Od návrhového vzoru se požaduje specifikace problémů, které vzor řeší, jak je řeší, a kdy je vzor použitelný.

Z databázových návrhových vzorů prezentovaných jednotlivými autory zmíněnými v úvodu nevzniklo sdílené pojmenování vzorů. Každý autor pojmenovává situace aplikovatelnosti svého vzoru po svém, konkrétně v názvech příslušných kapitol. Nicméně tyto názvy jsou vysvětlující natolik, že lze podle nich potřebný vzor vyhledat. V příslušné kapitole jsou pak objasněny obtíže, s nimiž se při řešení daného typu problému setkáváme. Řešení jsou předložena ve více variantách, s nastíněním, kdy je jaká varianta vhodná.

Liší se mírou abstrakce, ze které jsou jednotlivými autory oblasti problémů nahlíženy. Z tohoto pohledu Michael Blaha (2010, str. 4) rozděluje konstrukce na *abstraktní vzory (patterns)* a *zárodečné modely (seed models)*: zárodečným modelem nazývá rozsáhlejší univerzální řešení použitelná pro určitou aplikační oblast, zatímco skutečným vzorem nazývá abstraktní model menšího rozsahu, a aplikačně nezávislý. Sám Blaha nabízí spíše abstraktnější vzory použitelné napříč aplikačními oblastmi.

3. „Co je obtížné modelovat“ dle Richarda Veryarda

V knize z roku 1992 se Richard Veryard na stranách 203-289 věnuje některým záležitostem, jež byly a jsou často modelovány. Připomíná jejich různé stránky, jež zdánlivě jednoduchou problematiku objasňují jako složitější. Spíše než návrhové vzory zde nalezneme *důkladnou analýzu*.

Zabývá se následujícími tématy:

- počítatelné a identifikovatelné věci
- nepočítatelné substance či neidentifikovatelné předměty
- čas a prostor
- aktivity, scénáře, taktiky, plány
- sledování a řízení událostí
- řízení organizace
- obchodní vztahy a smlouvy
- řízení a uchování znalostí

Výklad je proložen modely v crow-foot notaci. Modely jsou omezeného rozsahu, k jejich pochopení slouží Veryardův detailní rozbor.

4. „Zvyklosti v myšlení“ dle Davida C. Haya

V knize z roku 1996 věnuje David C. Hay devíti běžným oblastem podnikové reality samostatné kapitoly. Tématem provází vždy postupně, od jednoduchého řešení – a jemu odpovídajícího modelu – k řešení zahrnujícímu i varianty, jež v různých případech mohou nastat. Podle rozlišení Michaela Blahy se jedná o *zárodečné modely*.

Široký záběr vychází z autorovy předchozí praxe, řešená témata jsou:

- firma – lidé, pozice, partneři, organizační struktura a lokality
- věci – produkty, vybavení, materiál
- procedury a aktivity – postupy a jejich kroky, úkoly, role v úkolech, práce na úkolech, užití majetku při úkolech, časové harmonogramy, zakázky, projekty
- smlouvy – typy smluv, smluvní strany, specifikace požadavků, dodávky
- účtování – náklady a přínosy, vztahy majetků a účtů, rozpočty, transakce
- laboratoř – vzorky, testy, pozorování
- plánování materiálu
- nepřetržitá výroba – topologie zařízení, toky materiálů, testování a měření

- dokumenty - umístění kopií, struktura dokumentů, autoři a další role osob k dokumentům, témata dokumentů, klíčová slova dokumentů, verze dokumentů, formáty dokumentů

V knize Haye (1996) je velké množství modelů v čisté Barker-ER notaci, k jejichž pochopení slouží vždy předchozí rozbor a výklad.

5. „Univerzální vzory datových modelů“ Lena Silversona a Paula Agnewa

V prvních dvou knihách (2001a), (2001b) Len Silverson přináší rozsáhlou sbírku modelů. První díl nabízí řešení problematik vyskytujících se v každém podniku, druhý díl nabízí řešení pro specifické obory podnikání. V porovnání s knihou Haye je každý model probírán obšírněji, a ve druhém dílu se zaměřuje na jiné oblasti.

Třetí díl, z roku 2009, jež vydal spolu s Paulem Agnewem, se podstatně odlišuje. Dle rozlišení Blahy se jedná o *univerzální abstraktní vzory*, Blaha (2010, str. XV) tyto konkrétní typy vzorů nazývá *archetypy*. Každému je věnována vlastní obsáhlá kapitola opakujícího se formátu. Každý vzor má název, vysvětlení účelu, vysvětlení vzoru samotného i příklad užití, doporučení, kdy má být vzor užit, a jeho slabiny. Každá kapitola postupně přináší generičtější verze modelů. Tím se může kniha jevit zbytečně objemnou, ale čtenář může vyhledat tu verzi, která je mu srozumitelná a pro něj přínosná.

Jednotlivá témata:

- deklarace rolí - například partner, dodavatel, sponzor, zákazník, zaměstnanec
- role v kontextu - například sponzor v daném projektu
- hierarchie, agregace, rekurzivní vztahy - například pracovní vztahy
- kategorizace, klasifikace, typizace - i v případech, kdy je více více schémat třídění, jež navíc nemusí být hierarchické
- stavy - přechody mezi stavy, systémy pravidel přechodu
- kontakty - účely a metody kontaktů, například kontakty na zákazníky
- obchodní pravidla - faktory a subjekty pravidel, kategorizace systémů pravidel

Modely jsou znázorněny v crow-foot notaci s cizími klíči.

V další kapitole ukazují autoři aplikaci svých vzorů do několika komplexních řešení, a v závěrečné kapitole ukazují, jak zavést návrhové vzory do povědomí a nástrojů řešitelského týmu.

6. Zkušenosti s výukou databázových návrhových vzorů

Na Vysoké škole ekonomické v Praze se již od roku 2010 vyučuje předmět Databázové návrhové vzory (<http://isis.vse.cz/katalog/syllabus.pl?predmet=100115>). Obsahově se zaměřuje na *archetypy podnikové informatiky*. Po úvodním vysvětlení účelu předmětu a představení plánu témat jsou studentům hned na začátku poskytnuty „příběhy“ vymezující rozsah těchto témat. Studenti jsou vyzváni, aby danou problematiku vždy dopředu promysleli a pokusili se o její řešení. Na příslušném

semináři je potom formou brainstormingu vytvářen návrhový vzor. Studenti také odevzdávají své domácí přípravy, k nimž jim vyučující poskytuje zpětnou vazbu. Většina studentů *dosahuje viditelného pokroku* v jejich domácích přípravách. Kromě toho v průběhu semestru pracují na své závěrečné práci, již je konceptuální návrh databáze pro vybranou problematiku; jediným kritériem hodnocení této práce je věcná správnost řešení. Studenti, kteří v průběhu aktivně spolupracují při re-konstrukci jednotlivých návrhových vzorů, *předvádějí na závěr výborné práce*.

7. Výhled využití databázových návrhových vzorů

Podle Blahy (2010, str. 3) jsou výhody užití vzorů následující:

- *Obohacení jazyka modelování* – nemusíme se vyjadřovat pouze v primitivních pojmech, ale i v běžných kompozitech
- *Zlepšení dokumentace* – dokumentace, která používá jazyka vzorů, je přehlednější a srozumitelnější, a rozhodnutí jsou jasnější
- *Méně obtížné modelování* – vzory nabízejí řešení
- *Rychlejší modelování* – používá se již vytvořené
- *Kvalitnější modely* – vzory jsou ověřené
- *Znovupoužití* - jednou vynaložené úsilí na osvojení si modelu se opakovaně zúčtuje

Proti využití vzorů pak stojí dle Blahy (tamtéž, str. 4) následující:

- Databázový návrh aplikace nesestavíte výlučně jen z návrhových vzorů
- Vyhledání relevantního vzoru může být obtížné
- Vzory jsou složité abstraktní konstrukce, jejich osvojení vyžaduje úsilí
- Nekonzistence mezi vzory různých autorů
- Obor databázových návrhových vzorů je nerozvinutý

Zkušenosti z Vysoké školy ekonomické však ukazují, že úsilí věnované studiu byt' jen některých databázových návrhových vzorů zlepšuje schopnosti modeláře.

V poslední době se setkáváme s rapidně vytvářenými a též rapidně upravovanými a obohacovanými aplikacemi. V tomto prostředí se často přistupuje k nasazení bezschémového databázového řešení, takže, jako tomu bývalo v počátcích podnikové informatiky, jsou datové struktury zachycovány pouze v aplikačním kódu. Pro spolupráci na projektech a udržitelný vývoj aplikací to je problém. Použití bezschémového databázového systému neznamená, že by nemělo být používáno a udržováno schéma používaných informací. Ovšem informační model se musí přizpůsobovat tak rychlému vývoji, jak rychle se vyvíjejí celé aplikace (Ambler, S. 2014). A zde přichází ke slovu stará zkušenost, že udržitelnosti informačního modelu nejlépe dosáhneme, pokud odráží realitu v oblasti zájmu. Zárodečné modely dle terminologie Blahy zjednodušují poznání reality v běžných oblastech, abstraktní vzory zjednodušují vývoj modelů v nových oblastech.

Literatura

- Alexander, Ch. & Ishikawa, S. & Silverstein, M. & Jacobson, M. & Fiksdahl-King, I. & Angel, S., 1977: *A Pattern Language: Towns, Buildings, Construction*, 1st ed., Oxford University Press
- Ambler, S. *Agile Model Driven Development (AMDD): The Key to Scaling Agile Software Development*, Dostupné na: <http://www.agilemodeling.com/essays/amdd.htm> [staženo 20.12.2014]
- Blaha, M., 2010: *Patterns of Data Modeling*, 1st ed., CRC Press
- Buchalceková, A., 2004: Methodology Patterns. In: *CITSA 2004*, červen 21-25, 2004, Orlando
- Buchalceková, A., 2008: *Metodiky budování informačních systémů*, Oeconomica, Praha
- Halpin, T., 2008: *Information Modeling and Relational Databases*, Morgan Kaufmann
- Hay, D.C., 1996: *Data model patterns : conventions of thought*, 1st ed., Dorset House
- Gregor, J. & Chvalovský, V. 1988: *Projektování datové základny*, Praha: SNL
- Fowler, M., 1996: *Analysis Patterns: Reusable Object Models*. Addison-Wesley
- Gamma, E. & Helm, R. & Johnson, R. & Vlissides, J., 1994: *Design Patterns: Elements of Reusable Object-Oriented Software*, 1st ed., Addison-Wesley
- Jardine, D. A., 1977: *The ANSI/SPARC DBMS Model*, North-Holland Pub
- Pecinovský, R., 2007. *Návrhové vzory*, 1. vyd., Computer Press
- Silverston, L., 2001a: *The Data Model Resource Book: Volume 1: A Library of Universal Data Models for All Enterprises*, rev. ed., Wiley
- Silverston, L., 2001b: *The Data Model Resource Book: Volume 2: A Library of Data Models for Specific Industries*, rev. ed., Wiley
- Silverston, L. & Agnew, P., 2009: *The Data Model Resource Book: Volume 3: Universal Patterns for Data Modeling*, rev. ed., Wiley
- Šešera, L. & Mičovský, A. & Červeň, J., 2001: *Datové modelování v příkladech*, Grada
- Veryard, R., 1992: *Information modelling: practical guidance*, 1st ed., Prentice Hall
- Veryard, R., 1984: *Pragmatic data analysis*, 1st ed., Blackwell Scientific Publications
- Voříšek, J. (ed.), 2008: *Principy a modely řízení podnikové informatiky*, Oeconomica, Praha

JEL Classification: M10, C88