

ТАКСОНОМИЯ НА ДАННИ ГРАДИВНИ БЛОКОВЕ

Васил Милев

Анотация: Разработката на архитектура на предприятието е процес на моделиране и документиране на всички аспекти на организацията. Разработена е таксономия на данни градивните блокове, която се базира на три основни класификационни признака: абстракция, обхват и зрялост. Тя може да послужи при специфициране и съхранение на информация за данни градивните блокове. Разработката може да се прилага като ръководен материал за ИТ отдели, които започват да реализират проекти в областта на синхронизацията на бизнеса и информационните технологии.

Ключови думи: архитектура на предприятия, архитектура на данните, моделиране на данни, таксономия, EIM, 3D Model, eTOM.

TAXONOMY FOR DATA BUILDING BLOCK

Vassil B. Milev

Summary: The Enterprise Architecture is the process of modeling and documentation of all aspects of the organization. The taxonomy of “data building block” is developed, which is based on three main classification criteria: abstraction, scope and maturity. It can be used for specifying and storing information for data building blocks. This work can be used as guidance material for IT departments, which will begin to implement projects in the field of synchronization of business and information technology.

Keywords: enterprise architecture, data architecture, data modeling, taxonomy, EIM, 3D Model , eTOM.

Доклада е публикуван в :

1 .ВЪВЕДЕНИЕ

Разработката на архитектура на предприятието [14,19] е процес на моделиране и документиране на *всички* аспекти на организацията, за да се гарантира, че услуги, процеси, приложения, информация, данни, технологии, места, хора, събития и срокове са съобразени с целите и задачите. Архитектурата е фундаментална организация на системи, реализирана чрез нейните компоненти, връзките помежду им, и с обкръжаващата среда, както и принципите, определящи нейното проектиране и развитие [13]. За да се осигури синхрон между развитието на бизнеса и информационните технологии (ИТ) [11] е необходимо създаване на таксономии [4], удовлетворяващи гъвкавото им развитие.

Обикновено при структуриране на данните се използват стотици или дори хиляди имена на архитектурните елементи от данни. По-голямата част от имената са омоними, синоними и псевдоними, създадени в отсъствието на три фактора:

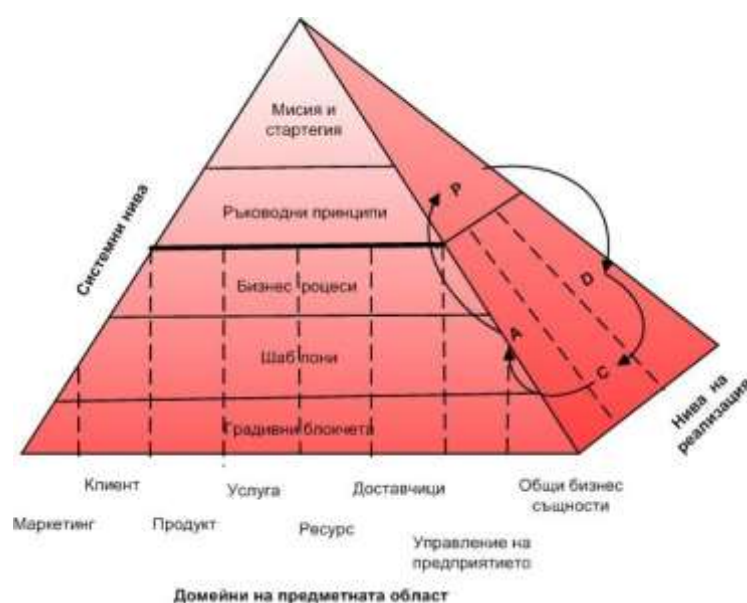
- стандарти за именуване;
- гледна точка на корпоративната архитектура на данни;
- планиране на създаване и използване на данните.

Цел на настоящият доклад е да представи система и средства за решаване на този проблем, т.е. чрез използване на малко на брой типови архитектурни елементи, наречени “данни градивни блокове“ (Data Building Blocks, DBB), които са систематизирани, описани и визуализирани, да се оптимизира дейността по моделиране на данни.

Тази цел се реализира чрез разработка на таксономия [4,17] на DBB, която класифицира по определени признаци основните архитектурни компоненти, използвани при моделиране на данните.

2. МЯСТО И РОЛЯ НА ДАННИ ГРАДИВНИТЕ БЛОКОВЕ И ТЯХНОТО ОПИСАНИЕ.

Мястото и ролята на DBB се изяснява чрез 3D МОДЕЛ на архитектура на данните [16] и разширения информационен модел (Extended Information Model, EIM) [7], представени на фиг.1 и фиг.2. Моделите описват конкретни архитектурни решения в областта на моделиране на данни [6,8], които се явяват типови и могат многократно да бъдат използвани в различни приложения.



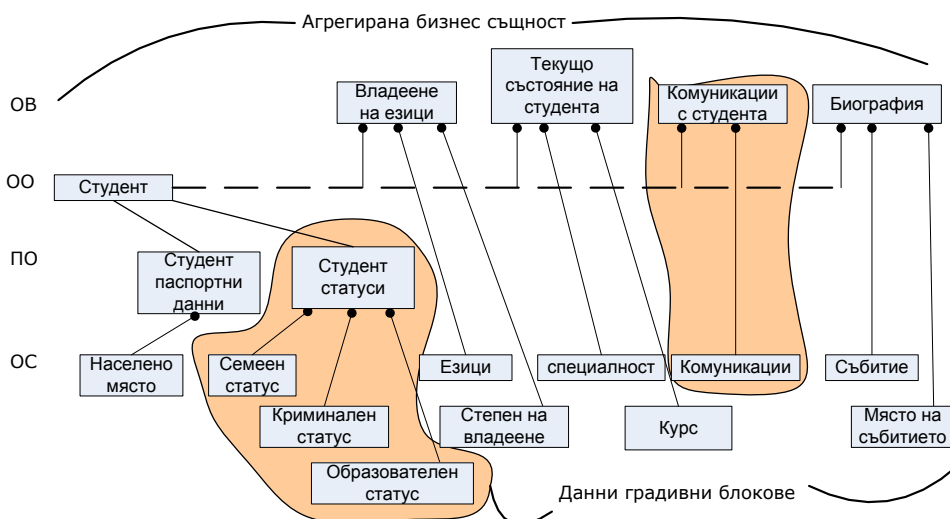
Фиг.1 . 3D МОДЕЛ на архитектура на данните [16].



Фиг.2 . РАЗШИРЕН ИНФОРМАЦИОНЕН МОДЕЛ [7] .

Представяват функционално обособени един или група обекти, които реализират конкретна функция в рамките на домейн от конкретна предметна област.

Примери за градивни блокове [9,10,15] са представени на фиг. 3: Комуникации със студента, Данни за студента, Статуси на студента в Агрегирана Бизнес Същност (Aggregated Business Entity, ABE), Студент в предметна област - Образование.



Фиг. 3 . АГРЕГИРАНА БИЗНЕС СЪЩНОСТ - Студент.

От изключително значение за разработка на качествени модели на данни е детайлното описание на DBB, което включва:

- извършва се текстово описание на същността;
- изброяват се ABE, в които градивния блок участва;
- изброяват се същностите които участват в градивния блок;
- описват се правила на използване и реализация.

Пример : Комуникации с персона

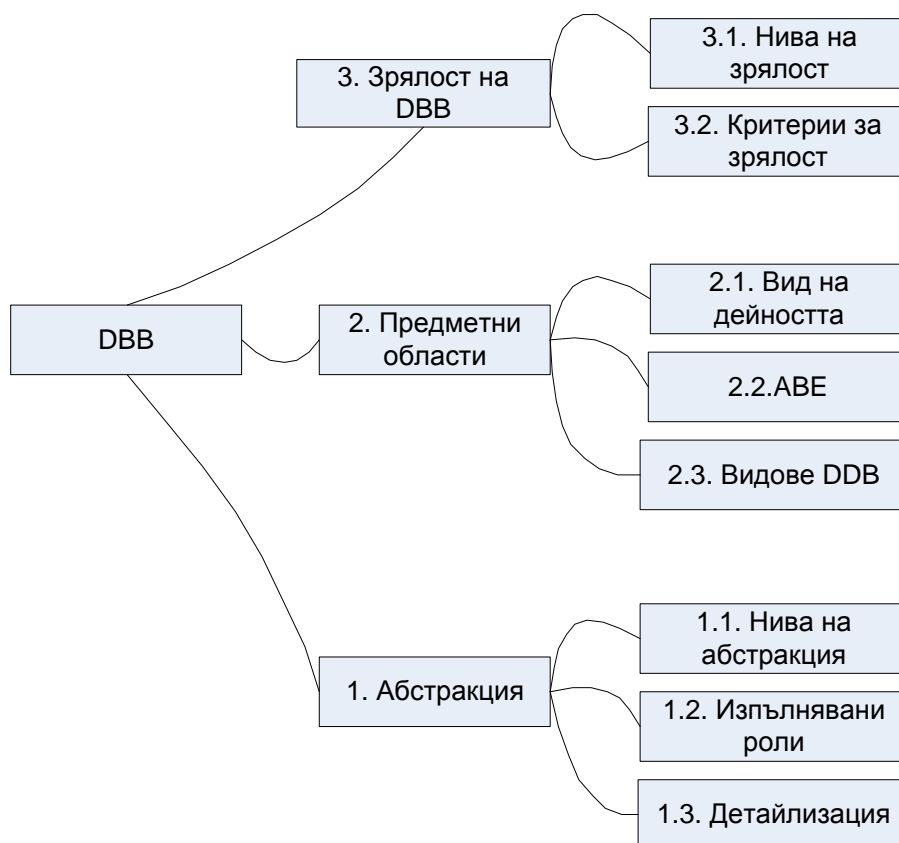
Таблица 1.

| Название | Комуникации с персона |
|-------------------|---|
| Описание | Подробно описание на градивния блок |
| АБС | Студент, преподавател, служител |
| Свързани същности | Персона, комуникации, комуникации с персоната |
| Правила | |

3. ТАКСОНОМИЯ НА ДАННИ ГРАДИВНИТЕ БЛОКОВЕ

Таксономията (от гръцки τάξις-строй, порядък и νόμος-закон) е наука за класификация, включваща правила, теория, методи и приложението им [4,17]. Представява колекция от контролирани условия и лексика, организирани в йерархична структура.

Бизнес таксономииите или така наречените класификации, вече се радват на доверие, като липсващо звено на проекти за управление на данни. Под таксономия в доклада ще се разбира набор от избрани термини, използвани за извличане на онлайн съдържание. Разработената примерна таксономия (фиг.4) ще се използва при специфициране и съхранение на информация за данни градивните блокове [9].



Фиг. 4. Таксономия на данни градивните блокове.

Основни компоненти на таксономията се явяват:

1. Абстракция

Указва предназначението на DBB, неговия разработчик и собственик, както и реализацията го модел. Включва следните нива :

1.1. Нива на абстракция - моделирането на данни се реализира в рамките на три модела - концептуален, логически и физически. Данни градивните елементи въпреки и част от общия модел на данни подлежат на същите обработки.

1.1.1. Концептуален модел - обикновено се създава в информационния етап стратегия за управление. Съдържа ключовите същности и отношения и представя графично, на високо ниво концепцията за дата градивния блок. Не зависи от използваната за реализация платформа и програмен език. Реализира се чрез използване на метода PER (Process-Entity-Relation) [6,8].

1.1.2. Логически модел - описва конкретните структурни елементи (таблици), техните атрибути и съответните типове данни. Определя приложимите бизнес правила.

1.1.3. Физически модел - крайното представяне на структурите на базата данни, които са получени от модела. Съдържа подробни спецификации за проектиране на бази данни.

1.2. Изпълнявани роли

1.2.1. Архитект - разработва системната архитектурата на предприятие, включваща архитектурата на данните, приложенията и инфраструктурата [13,19].

1.2.2. Проектант - проектира приложенията и съпътстващите ги бази данни, като създава логическия модел на данните.

1.2.3. Разработчик - разработва приложенията и физическият модел на базата данни.

1.2.4. Собственик - отговаря за съответни данни градивни блокове - съхранение, актуализации, нови приложения.

1.3. Детайлизация

1.3.1. Предприятие/организация - съвкупност от взаимодействащи елементи, които реализират конкретни бизнес процеси и преследват определени цели [13,14,19].

1.3.2. Проект - извършване на съвкупност от разнородни дейности за решаване на сложен нестандартен проблем при наложени ограничения относно време, разходи, качество и специфични изисквания към неговата организация.

1.3.3. Приложение - програма за реализиране на конкретна задача.

1.3.4. Модул на приложение - част или фрагмент от приложението, за реализиране на конкретната задача .

2. Предметни области

Основна задача на информационното моделиране е разработка на формално описание на дадена предметна област (обикновено засягаща едно предприятие), не зависещо от особеностите на реализация на модела и съответното програмно осигуряване. Предметната област касае сфера на дейност в реалния свят, за която е била поставена задачата да бъде прехвърлена в компютърния свят. Примери: телекомуникации, университети, медицина, търговия . Тя се разделя както следва :

2.1. Вид на дейността - за всяка организация са типични три направления в дейността, които се описват със съответните бизнес процеси. Явяват се основните домейни, които трябва да бъдат моделирани при изграждане на информационната инфраструктура и приложения в една организация. Много добре са развити в концепцията The New Generation Operations Systems and Software (NGOSS) на

TeleManagement Forum (TMF) [12] и по специално в модела на процеси The enhanced Telecom Operations Map (eTOM) [1].

На 1, 2 и 3 нива от модела eTOM [1] се извършва детайлизация на описаните по-горе области до подпроцеси (например, диагностика на проблеми или контрол на показателите за производителност на ресурсите):

2.1.1. Стратегия [6,7,8] - включва целия спектър въпроси свързани със стратегията на организацията, развитието на инфраструктурата и управление на жизнения цикъл на продуктите (услугите).

2.1.2. Операционни процеси [7,12] - явяват се най-важните процеси в една организация, които обезпечават представянето и осигуряването на услуги и продукти.

2.1.3. Управление на организацията - обхваща общи въпроси по осигуряване дейността на организацията, такива като управление на кадри, финанси и активи, знания, външни връзки и т.н.

2.2. Агрегирани бизнес същности [7] - представляват тясно свързани различни обекти, обединени в една група. Тя се състои от DBB, които изясняват и детайлизират основната бизнес същност-обекти, участващи в съответния процес. ABE се явяват основни за предприятието данни, които зависят от сферата, в която то действа. Пример за ABE в областта на образованието са : студент, преподавател, дисциплина, ресурси и т.н.

2.3. Видове градивни блокове

В модела SID (Shared information and data model) [2,12] за всяка ABE са определени следните данни градивни блокове.

2.3.1. Основни данни - описва основните данни на ABE, като определя същността на останалите обекти.

2.3.2. Стратегия и планиране - описва задачи, критерии и срокове за изпълнение на решения.

2.3.3. Спецификация - описва различни статуси на ABE.

2.3.4. Комуникация - отразява данни касаещи мобилни, стационарни и интернет комуникации.

2.3.5. Конфигурация - описва вътрешната структура на АВЕ - паспортни данни, фирмени данни, вход, изход и т.н.

2.3.6. Параметри за ефективност - описва параметри за оценка на производителността (ефективността) качеството на АВЕ.

2.3.7. Средства за тестване - описва средства и инструменти за оценка на АВЕ.

2.3.8. Проблеми - описва проблеми, които касаят АВЕ.

2.3.9. Събития във времето - описва събития и тяхната историческа свързаност в контекста на АВЕ.

В таблица 2 са дадени примери за данни градивни блокове на “Преподавател” и “Студент” за АВЕ.

Таблица 2 .

| Категория АВЕ DBB | Преподавател | Студент | Дисциплина |
|--------------------------|-------------------------------|-----------------------------------|--------------------------------|
| Същност | Преподавател | Студент | Дисциплина |
| Стратегия и план | Задачи и критерии | | Задачи и критерии |
| Статуси | Научна степен, научно звание, | Образователен, семеен, криминален | Изпит |
| Комуникация | Комуникации с преподавателя | Комуникации със студента | Комуникации с други дисциплини |
| Конфигурация | Паспортни данни | Паспортни данни | Данни за дисциплината |
| Параметри за ефективност | Параметри за ефективност | Параметри за ефективност | Параметри за ефективност |
| Средства за тестване | Атестация | Изпити | Обсъждане |
| Проблеми | Проблеми на | Проблеми на | Проблеми на |

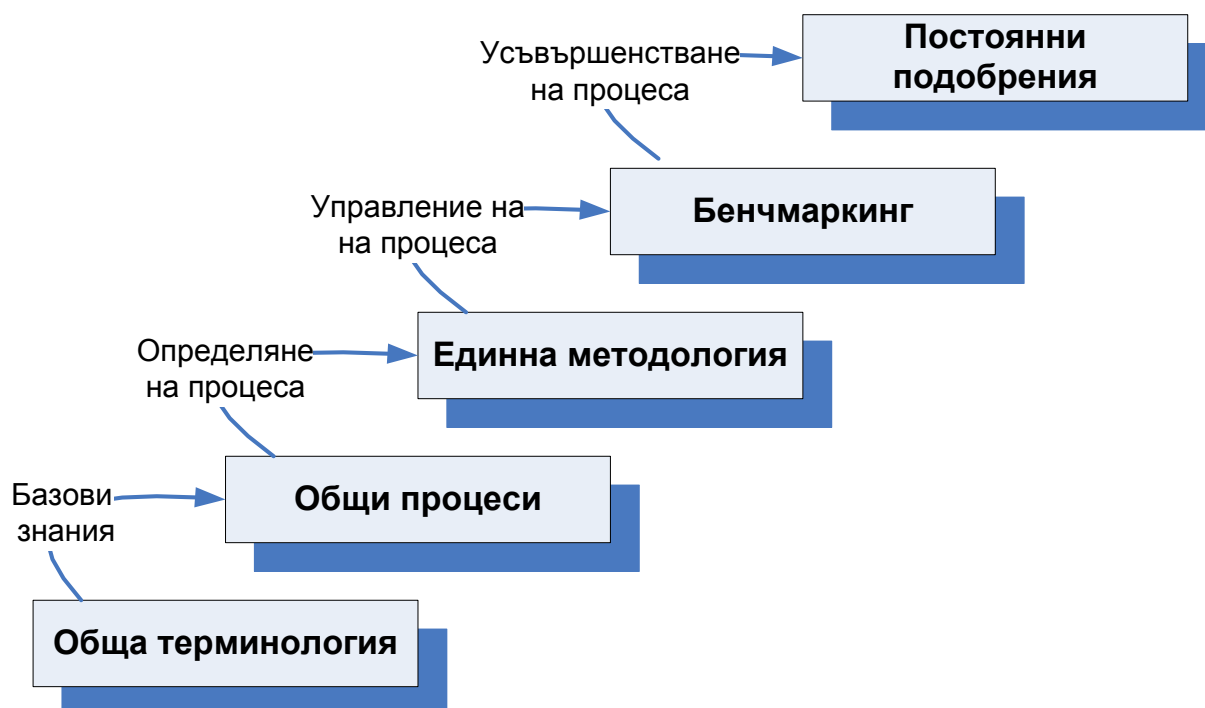
| | | | |
|--------------------------|----------------------------|-----------------------|-------------------------|
| | преподавателя | студента | дисциплината |
| История/биография | Биография на преподавателя | Биография на студента | История на дисциплината |

3. Зрялост на градивните блокове

Отчитане нивото на зрялост на архитектурните компоненти при моделиране на данните е в основата на получаване детайлна и ясна картина за състоянието на архитектурата на предприятието и в частност на данни градивни блокове. Моделите на зрялост [18,19,20] определят визията и пътната карта за усъвършенстване на процеса по моделиране на данни в една организация.

3.1. Нива на зрялост

Моделът на зрялост на Гаролд Керцнер [3], известен като Project Management Maturity Model (PMMM) се състои от 5 нива (фиг.5), всяко от които представя различна степен на зрялост на процесите по създаване на модели на данни в една организация.



Фиг. 5. Нива на зрялост на градивните блокове

3.1.1. Обща терминология - организацията осъзнава важността от управление на данните и си поставя цели по тяхното моделиране [6,7,9] и последващо пълноценно използване. Започват проучвания и дейности по изучаване и получаване на базови знания за моделиране на данни и в частност за данни градивните блокове.

3.1.2. Общи процеси - организация е натрупала знания в областта на моделиране на данни и осъзнава необходимостта от общи процеси по управление, съхранение и поддържане качеството на данните, гарантиращи стандартизация на дейността и многократно използване [13].

3.1.3. Единна методология - организация осъзнава важността от използване на единна методология, базирана на архитектурния подход [11,15,19], гарантираща синергетичен ефект, от едновременната разработка на архитектурите на бизнес процесите, данните, приложенията и инфраструктурата.

3.1.4. Бенчмаркинг - на това ниво организацията проучва най-добрите практики в областта на информационното моделиране [7], и внедрява тези от тях, които съответстват на нейните цели и задачи.

3.1.5. Постоянни подобрения - на това ниво организацията е създала и поддържа добро ниво на управление на данните си. Периодично се проверява качеството и пълноценното им използване.

За модела са валидни следните правила:

- всяко следващо ниво започва след завършване на предходното;
- нивата могат да се припокриват;
- за всяко ниво се разработват показатели, обхващащи петте перспективи на рамката.

3.2. Критерии за зрялост

Критериите за зрялост дават представа за характеристиките на които трябва да отговарят данни градивните блокове и за тяхното влияние върху възможностите за преминаване към следващи нива на зрялост. Те могат да получат следната класификация :

3.2.1. Коректност - определя дали DBB съответства на правилата и техниките за моделиране на данни [9,15] (т.е. дали е валиден модел на данните). Това включва диаграми конвенции, правила за именуване, правила за структуриране, правилата за състава и нормализиране.

3.2.2. Пълнотата - отнася се до това дали DBB съдържа цялата информация, необходима за подкрепа на изискваната функционалност на системата в АВЕ.

3.2.3. Почтеност - определя дали DBB, включва всички бизнес правила, които се прилагат към данните.

3.2.4. Простотата - означава, че DBB съдържа минималните възможни същности и отношения.

3.2.5. Гъвкавостта - определя се лекотата, с която DBB може да се справя с промени в бизнеса или в нормативната уредба.

3.2.6. Интеграция - определя се като съвместимост на DBB с останалите архитектурни компоненти на модела на данни на организацията.

3.2.7. Разбираемост - определя се като лекота, с която концепции и структури в данни градивния блок могат да бъдат разбрани.

3.2.8. Приложимост - определя се от това дали конкретния данни градивен блок, фигурира в хранилището на данни, както и може ли да бъдат изпълнен в рамките на срока, бюджетни ограничения и технология на проекта.

3.2.9. Стандартизираност - отразява доколко всички приложими стандарти са спазени, независимо от техния произход.

3.2.10. За многократна употреба - DBB трябва да осигурява възможност за многократна употреба, както самостоятелно, така и в състава на други по-мащабни модели.

3.2.11. Автоматизирана обработка - определя степента на формализация, която позволява да се използват програмни инструменти за въвеждане, съхранение, обработка, анализ, и визуализация на описанието.

ЗАКЛЮЧЕНИЕ

Представеният материал дава най-обща представа за същността, използваните принципи и характеристиките на една много обсъждана в ИТ средите тема, каквато е архитектурата на данните [9,15,18,19]. Направени са изводи за фундаменталната роля и мястото на архитектурата на данни в архитектурата на предприятието. Разглежда се мястото и ролята на важен според автора архитектурен елемент - данни градивни блокове. Представени са моделите, в които той участва, съдържанието на описанието му, както и примери за приложението му.

Разработена е таксономия на данни градивните блокове, която се базира на три основни класификационни признака: абстракция, обхват и зрялост. Тя може да послужи при специфициране и съхранение на информация за данни градивните блокове. Разработката може да се прилага като ръководен материал за ИТ отдели, които започват да реализират проекти в областта на синхронизацията на бизнеса и информационните технологии.

ЛИТЕРАТУРА

1. International Telecommunication Union , Enhanced Telecom Operations Map (eTOM) – The business process framework, printed in Switzerland, Geneva, <http://www.billingcollege.com/upload/M.3050.1.pdf>, 2005
2. International Telecommunication Union, Shared information and data model (SID, ITU-T M-Series Recommendations, Printed in Switzerland, Geneva, <http://www.itu.int/rec/T-REC-M.3190-200807-I>, 2009
3. Kerzner H. R, Using the Project Management Maturity Model: Strategic Planning for Project Management, John Wiley & Sons Inc, 2005
4. Lawrence A. T., Taxonomy for the Technology Domain, Information Science Publishing (an imprint of Idea Group Inc.), 2005
5. Lubyansky A., Using Data Model Patterns to Build High-Quality Data Models Analyst, Project Performance Corporation-Part of the AEA group, <http://www.ppc.com/assets/pdf/white-papers/Data-Model-Patterns.pdf>
6. Milev V, H. Tujarov, S. Kalchev, Data PER model , International Conference on Computer Systems and Technologies - CompSysTech'13, 28-29 June 2013- Ruse, Bulgaria ,,2013
7. Milev V, S. Kalchev, Extended Information Model (EIM), http://www.tuj.asenevtsi.com/Public/EIM_UPLOAD.pdf
8. Milev V, S. Kalchev, Method For Data Modeling With Per Model, International Conference on Information Technologies (InfoTech-2013), 15-16 September 2013 Bulgaria, 2013
9. Milev V., The “Building Blocks” Concept Using for Modeling Data, International Conference on Information Technologies (InfoTech-2011), 15-16 September 2011 Bulgaria, 2011
10. Milev V., Multiple Dimensions Of Enterprise Architectural Description, Научна конференция с международно участие “Предизвикателства пред висшето образование и научните изследвания в условията на криза”, Бургас 2010
11. Milev V., S. Kalchev, M. Stefanova, Architectural approach, based synchronization business and information technology, Научна конференция с международно участие “Mathtech2010”- Шумен 2010

12. TM Forum, Information Framework (SID), 2011,
<http://www.tmforum.org/InformationFramework/1684/home.html>13
13. Tujarov H., S. Kalchev, V. Milev, Component model of the enterprise architecture, Proceedings of the International Conference on Information Technologies (InfoTech-2010) 16-18 September 2010, Bulgaria
14. Tujarov H., S. Kalchev, V. Milev, Methods for creation of Enterprise Architecture, International Conference on Information Technologies (InfoTech-2011), 15-16 September 2011 Bulgaria, 2011
15. Tujarov H., S. Kalchev, V. Milev., Using The “Building Blocks“ For Modeling Data, Международна научна конференция “Информатиката в научното познание 2010”, Варна 2010
16. Милев В., “3D модел на архитектура на данните”. "Приложение на ИКТ в икономиката и образованието", 2-3 декември, 2011, УНСС, София.”16
17. Милев В., В. Монеv, Използване на таксономия при разработка на архитектура на данните, Международна научна конференция “Информационните технологии-стратегически приоритет в икономиката на знанието“, 14-15 октомври 2011г, Свищов, 2011
18. Милев В., С. Калчев, Архитектурата на данните като фундамент на архитектурата на предприятие Юбилейна международна научна конференция -50 години ВТУ “ Св. св. Кирил и методий”, Велико Търново,2013
19. Тужаров Х., С. Калчев, В. Милев, Архитектура на предприятие,
<http://tuj.asenevtsi.com/EA/Index.htm>, 2010
20. Тужаров Х., С. Калчев, В. Милев, Модел на зрялост на информационните технологии, международна научна .конференция “Информационни технологии в управлението на бизнеса”, Варна 2009