

УДК 314.1; 004.6

О. Р. Чертов

Національний технічний університет України «КПІ»
проспект Перемоги, 37, 03056 Київ, Україна

Архітектура систем обробки демографічної інформації

Розглянуто сучасні підходи до розробки корпоративних архітектур. Побудовано модель Захмана функціонально-інформаційної архітектури статистичної галузі України та запропоновано її розширення шляхом виділення рівня системної архітектури.

Ключові слова: корпоративна архітектура, модель Захмана, системна архітектура, демографічна інформація.

Вступ

Архітектурні питання під час розробки складної інформаційної системи відіграють настільки ж важливу роль, що й під час спорудження неординарної будівлі. Правильне архітектурне рішення значно знижує ризики всього проекту з розробки й супроводження системи та збільшує віддачу від вкладених коштів, оскільки дозволяє як ефективно використовувати існуючу інфраструктуру, так і оптимально спланувати її подальший розвиток. Ця обставина стала однією з основних причин того, що найвідоміша на даний момент методологія розробки програмного забезпечення Rational Unified Process була побудована саме як архітектурно-центрований процес [1]. Ще в 2002 р. аналітична компанія META Group (з 2005 р. входить до складу компанії Gartner, Inc.) зазначала, що планування архітектури та слідування прийнятим галузевим стандартам може до 30 % зменшити витрати на інформаційні технології (ІТ) [2]. Очевидно, що з того часу значення архітектури, як мінімум, не зменшилося.

Але наразі в галузі ІТ існує багато різних визначень і трактувань самого поняття «архітектура», при розв'язанні архітектурних питань застосовуються різноманітні стандарти, методології, підходи, які в ідейному плані достатньо близькі одне до одного, але суттєво відрізняються за формою свого подання чи опису. В той же час є принципово різні рівні архітектурного представлення системи. Все це, загалом, ускладнює процес формування оптимальної архітектури для комплексних інформаційних систем, що автоматизують діяльність великих корпорацій чи окремих державних інституцій.

На 2012 р. у статистичній галузі України заплановано два масштабних впровадження:

© О. Р. Чертов

— автоматизованої системи обробки даних Всеукраїнського перепису населення, який повинен відбутися того ж року;

— Інтегрованої системи обробки статистичних даних України, створення котрої розпочато в рамках проекту розвитку системи державної статистики України для моніторингу соціально-економічних перетворень при підтримці Міжнародного банку реконструкції та розвитку.

У зв'язку з цим особливо актуальними стають питання вироблення єдиних методологічних підходів до побудови архітектури зазначених та інших аналогічних систем, в основі котрих лежить обробка демографічної інформації.

Запит на детальні показники статистики населення України стимульований, насамперед, владними структурами для розрахунку видатків з державного бюджету на соціальні потреби населення країни і визначення міжбюджетних трансфертів у галузях охорони здоров'я, соціального захисту населення, культури, міського суспільного транспорту тощо, а також для обґрунтування законодавчих і нормативних актів державних програм соціально-економічного розвитку.

Метою публікації є аналіз сучасних методологій (стандартів) представлення архітектури в області ІТ і формування з єдиних позицій архітектурного опису систем обробки демографічної інформації.

Архітектура. Корпоративна архітектура

Архітектура забезпечує механізм розуміння та управління складними явищами [3]. Серед великої кількості означень архітектури наведемо три з них, які були закріплені в ІТ-стандартах:

- організаційна структура системи чи компоненту [4];
- набір правил для визначення структури системи і взаємозв'язків між її частинами [5];
- фундаментальне розуміння системи, втілене через елементи системи, їхні зв'язки одне з одним і з оточуючим середовищем, й принципи, що визначають розробку системи та її розвиток [6].

Спочатку в області ІТ термін «архітектура» застосовувався лише до апаратного забезпечення та технологічної інфраструктури, згодом — взагалі до інформаційних систем. Лише з часом стало зрозуміло, що потрібно системно підходити не тільки до побудови ІТ, але й всього підприємства. В результаті чого виник термін «архітектура підприємства» (enterprise architecture). Дана термінологія передбачає досить широке трактування поняття «підприємство»:

- одна організація чи декілька, які виконують загальну місію та керуються спільними цілями і задачами для надання певного виходу, наприклад, продукту чи послуги [7];
- група організацій, що керуються спільними цілями та задачами для надання продуктів і послуг [8].

Отже, термін «enterprise architecture» може застосовуватися і до архітектури державних організацій та відомств, тому, на наш погляд, більш адекватним є його переклад як «корпоративна архітектура». Коли мова буде йти виключно про корпоративну архітектуру конкретного відомства чи відповідної галузі, наприклад,

статистичної, то будемо використовувати еквівалентний термін «відомча функціонально-інформаційна архітектура».

Серед багатьох визначень корпоративної архітектури візьмемо за основу визначення, запропоноване «Всесвітньою організацією корпоративної архітектури» (Global Enterprise Architecture Organization, GEAO) — одним із провідних співтовариств у цій галузі: «Корпоративна архітектура пояснює спосіб, за допомогою якого загальне бачення діяльності організації відображається в його структурі та динаміці. На різних рівнях абстракції вона дає єдиний набір моделей, принципів, керівництв і політик, котрі використовуються для трансформації, синхронізації та розвитку систем у масштабі і контексті діяльності всієї організації в цілому» [9].

Загалом, під час опису корпоративної архітектури керуються наступними принципами [10]:

— рівень деталізації архітектури вибирається таким чином, щоб мінімізувати об'єм інформації про компоненту, вилучивши все, що не стосується взаємодії з іншими компонентами архітектури;

— визначення архітектури не повинно містити опис самих компонент.

Як правило, в корпоративній архітектурі виділяють чотири основні типи (рівні) [10–12], які зведено до табл. 1.

Таблиця 1. Основні рівні корпоративної архітектури

№	Назва рівня	Пояснення
1.	Архітектура діяльності (business architecture)	визначає бізнес-процеси та організаційні структури організації
2.	Архітектура інформації (information architecture)	описує логічну та фізичну організацію даних, ресурси з управління даними, санкціонування доступу до них
3.	Архітектура застосувань (application architecture)	надає для прикладних систем план та інтерфейси їх взаємодії між собою, з зовнішніми системами, джерелами чи користувачами даних, визначає зв'язок прикладних систем із бізнес-процесами організації, що підтримуються ними
4.	Технологічна архітектура (інфраструктура, technical architecture)	забезпечує роботу прикладних систем на рівні, описаному в операційних вимогах (надійність, масштабованість, пропускну здатність тощо), включає апаратні засоби, системне програмне забезпечення, мережі, комунікаційні протоколи і т.д.

Основні архітектурні методики (architecture frameworks).

Модель Захмана для функціонально-інформаційної архітектури статистичної галузі

У стандарті ISO 15704:2000 виділяються два типи корпоративних архітектур [7]:

— архітектури систем (інколи називаються архітектурами типу 1, котрі стосуються дизайну певної системи, наприклад, відомчої системи або комп'ютерної системи управління як частини загальної інтегрованої системи підприємства чи організації;

— архітектури або плани (проекти) рівня всього підприємства (інколи називаються архітектурами типу 2, такі загальні архітектури стосуються розробки та впровадження, наприклад, електронного уряду [13].

Наразі існує декілька десятків методик (frameworks) опису та побудови корпоративних архітектур типу 2. Аналітична компанія Gartner прогнозує [14], що в 2015 р. 95 % всіх організацій, які використовують у своїй діяльності методики корпоративної архітектури, будуть одночасно застосовувати по декілька таких методик. Ця ситуація викликана тим, що діяльність різноманітних приватних компаній та державних інституцій настільки різна, що їхня загальна архітектурна модель є малопродуктивною, а внесення певної специфіки (конкретики) може бути визначено з різних точок зору.

У роботі [15] виділяються як основні, що покривають близько 90 % всіх застосувань, наступні чотири методики:

— модель Захмана (Framework for Information Systems Architecture) — методика опису архітектури, насамперед, інформаційних систем [16];

— FEAF (Federal Enterprise Architecture Framework) — федеральна архітектура державних організацій США [12];

— TOGAF (The Open Group Architecture Framework) — методика опису архітектури, розроблена Open Group [11];

— методика опису архітектури, розроблена аналітичною компанією Gartner, Inc. [17].

З погляду забезпечення повноти опису відомчих функціонально-інформаційних архітектур заслуговує уваги також методика DoDAF (Department of Defense Architecture Framework) опису архітектури Міністерства оборони США [3].

Виділення зазначених методик також обумовлено їхньою «різнотиповістю»: модель Захмана — це таксономія (систематика), TOGAF — це опис архітектури як процесу, FEAF та DoDAF — це закінчені методології, а методика Gartner фактично є набором найкращих практик.

У всіх виділених методиках опис корпоративної архітектури розділяється на певні рівні абстракції та певні предметні області (домени). Проілюструємо їх на прикладі побудови моделі Захмана для функціонально-інформаційної архітектури статистичної галузі. Тим більше, що для зазначених методик існують відображення (mapping) на модель Захмана (наприклад, див. відповідне відображення методики TOGAF [18]).

Модель Захмана представляє собою таблицю (матрицю), розміром 6 на 6 (табл. 2). Колонкам цієї таблиці однозначно співставляються домени, які визначають цілі та обмеження системи (чому?), дані, котрі використовуються чи створюються (що?), процеси та функції (як?), місце дії чи виконання цих процесів (де?), учасники — люди, підрозділи, організації (хто?) та події (коли?).

Кожну предметну область можна розглядати та аналізувати на різних рівнях абстракції: від точки зору керівництва до уявлень безпосередніх виконавців та отримувачів відповідних даних і послуг. Таким чином, рухаючись по моделі Захмана вздовж однієї колонки зверху вниз, поступово та послідовно отримуємо різні ступені деталізації опису.

Таблиця 2. Модель Захмана для функціонально-інформаційної архітектури статистичної галузі України

Домени Рівні абстракції	ЧОМУ (мотивація)	ЩО (дані)	ЯК (функції)	ДЕ (дислокація)	ХТО (люди)	КОЛИ (час)
Контекст (сфера дії), рівень керівництва	законодавчо визначені цілі	статистичні обстеження та показники	перелік послуг і процесів	географічний розподіл органів статистики, місця проведення досліджень	ієрархічна (трирівнева) організаційна структура, споживачі інформації	програма статистичних досліджень
Концептуальна модель, рівень методологів	моделі показників звітності	метамоделі, моделі інформації, класифікатори	модель статистичних процесів	схема логістики (інформаційних потоків і потоків керування)	моделі потоків робіт (workflow), перелік ролей користувачів	календар (та періодичність) спостереження
Логічна модель, рівень архітекторів	суб'єкти та моделі діяльності	логічні моделі даних	архітектура застосувань	модель розподіленої архітектури	модель рольових відносин	діаграма подій, їхня синхронізація
Технологічна модель, рівень проєктувальників	посадові інструкції	фізичні моделі даних	специфікації використання	інфраструктура	інтерфейси користувача	специфікація подій
Деталі реалізації, рівень виконавців та розробників	керівництва користувача, інструкції із застосування, регламент робіт	опис структур даних	програмний код, програми, що виконуються	мережна архітектура	автоматизовані робочі місця	визначення часових прив'язувань
Практика застосування	реальні показники точності, своєчасності, надійності інформації, що надається	первинні та агреговані дані, мікродані	регламенти, що реалізуються послуги, що надаються	фізичне розташування обладнання, канали надання послуг	постачальники та споживачі інформації і послуг	розклад надання (чи доступності) інформації

Розглянута методика виступає потужним інструментом застосування системного підходу до планування, розробки та використання різноманітних інформаційних систем.

ційних систем. Створювана за даною методикою корпоративна архітектура має високу адаптивність і забезпечує ефективне керування змінами під час перебудови системи відповідно до нових (чи тих, що змінилися) вимог.

Системна архітектура як архітектура систем роботи з даними

При застосуванні загальних методик до опису архітектур першого, а не другого типу (в розумінні стандарту ISO 15704:2000) бажано виділити додаткові рівні абстракції і/або домени, що є суттєвими для відповідної проблемної області.

Для функціонально-інформаційної архітектури статистичної галузі ми пропонуємо додатково виокремити *системну архітектуру* як архітектуру систем роботи з даними. В табл. 1 цей архітектурний рівень повинен розміщуватися над архітектурами інформації та застосувань, а в табл. 2 — між рівнями методологів та архітекторів.

З точки зору системної архітектури сучасні системи обробки демографічних даних можуть бути трьох видів [19]:

— інформаційно-облікові (OLTP) [20], котрі забезпечують базову функціональність, таку як введення даних і матеріалів відповідних статистичних спостережень та опитувань, їхнє структуроване (як правило, за допомогою СКБД — системи керування базами даних) зберігання та облік, контроль в первинному та зведеному вигляді, розповсюдження результатів через різноманітні регламентні вихідні таблиці;

— інформаційно-аналітичні (OLAP) [21], за допомогою котрих користувачі можуть швидко будувати нерегламентні таблиці й проводити інші аналітичні дослідження статистичних даних, шукаючи передбачувані закономірності їхнього розподілу;

— системи інтелектуального аналізу даних, котрі беруть на себе найбільш громіздку та рутинну аналітичну операцію з пошуку прихованих закономірностей, що, можливо, існують у даних, що аналізуються.

На наш погляд, не є правомірним зроблене в роботі [10] віднесення до архітектури інформації механізмів (прикладних систем) доступу до даних. Типологія таких прикладних систем зв'язана не тільки і не стільки з даними, бо, наприклад, роботою з первинними даними фактично займаються всі три виділені нами види систем. Важливу роль також відіграють архітектура застосувань та, власне, ті категорії задач, що розв'язуються системами кожного виду.

Висновки та перспективи розвитку досліджень

Функціонально-інформаційна архітектура є тим стрижнем, з якого повинна починатись і навколо якого повинна вестись розробка будь-якої відомчої інформаційної системи. В статті побудована модель Захмана корпоративної архітектури статистичної галузі України та здійснене її розширення шляхом виділення рівня системної архітектури.

Подальша деталізація отриманої моделі, за необхідності, може бути зроблена, наприклад, у відповідності з міжнародним стандартом на загальну модель статистичних бізнес-процесів [22], котрий включає в себе опис 9 етапів, кожен з яких складається з різної кількості підпроцесів (від 3-х до 8-ми). Також представляєть-

ся перспективним і важливим (з точки зору значення для проблемної області) виділення та дослідження домену забезпечення конфіденційності інформації.

Розроблена модель буде використана під час створення Інтегрованої системи обробки статистичних даних України та автоматизованої системи обробки даних Всеукраїнського перепису населення 2012 р.

1. *Kroll P.* The Spirit of the RUP, The Rational Edge, December 2001 [Електронний ресурс] / P. Kroll. — Режим доступу: <http://www.ibm.com/developerworks/rational/library/content/RationalEdge/dec01/TheSpiritoftheRUPDec01.pdf>
2. *Enterprise Architecture Desk Reference*, 2002 Edition / META Group Inc., Executive insights Series, 2002.
3. *DoD Architecture Framework V2.0. Volume 2: Architectural Data and Models – Architect’s Guide* [Електронний ресурс]. — 2009. — 262 p. — Режим доступу: http://www.defenselink.mil/cionii/docs/DoDAF_V20_Volume_2.pdf
4. IEEE Std 610.12-1990, IEEE Standard Glossary of Software Engineering Terminology. — IEEE Standards Board, 1990. — 83 p.
5. ГОСТ Р ИСОМЭК 10746-2-2000. Информационная технология. Взаимосвязь открытых систем. Управление данными и открытая распределенная обработка. Часть 2. Базовая модель. — М.: ИПК Издательство стандартов, 2001. — 23 с.
6. ISO/IEC 42010:2007. Systems and Software Engineering — Recommended Practice for Architectural Description of Software-Intensive Systems. — ISO, 2008. — 23 p.
7. ISO 15704:2000. Industrial Automation Systems — Requirements for Enterprise-Reference Architectures and Methodologies. — ISO, 2000. — 43 p.
8. ISO 14258:1998. Industrial Automation Systems — Concepts and Rules for Enterprise Models. — ISO, 1998. — 15 p.
9. *Willett K.D.* Information Assurance Architecture / K.D. Willett. — Boca Raton: CRC Press, 2008. — 588 p.
10. *Данилин А.* Архитектура и стратегия. «Инь» и «Янь» информационных технологий предприятия / А. Данилин, А. Слюсаренко. — М.: Интернет-Ун-т Информ. технологий, 2005. — 504 с.
11. *Togaf Version 9 Enterprise Edition: A Pocket Guide* / [A. Josey, R. Harrison, P. Homan et al]. — Wilco: Van Haren Publishing, 2009. — 153 p.
12. *FEA Consolidated Reference Model Document Version 2.3* [Електронний ресурс]. — Режим доступу: http://www.whitehouse.gov/omb/assets/fea_docs/FEA_CRM_v23_Final_Oct_2007_Revised.pdf
13. *Шевцов О.М.* Архітектура електронних комунікацій / О.М. Шевцов // Інформаційні технології і засоби навчання: електронне наукове фахове видання [Електронний ресурс] / Ін-т інформ. технологій і засобів навчання АПН України, Ун-т менеджменту освіти АПН України. — 2010. — № 2(16). — Режим доступу: <http://www.ime.edu-ua.net/em16/emg.html>
14. *Gartner Predicts 95 Per Cent of Organisations Will Support Multiple Approaches to Enterprise Architecture by 2015*, Gartner, Inc., 2010 [Електронний ресурс]. — Режим доступу: <http://www.gartner.com/it/page.jsp?id=1358913>
15. *Sessions R.* A Comparison of the Top Four Enterprise Architecture Methodologies, 2007 / R. Sessions. — 44 p. [Електронний ресурс]. — Режим доступу: <http://www.objectwatch.com/whitepapers/4EAComparison.pdf>

16. *O'Rourke C.* Enterprise Architecture Using the Zachman Framework / C. O'Rourke, N. Fishman, W. Selkow. — Boston: Thomson Course Technology, 2003. — 752 p.
17. *Gartner* Enterprise Architecture Framework: Evolution 2005 / G.A. James, R.A. Handler, A. Lapkin, N. Gall. — Gartner, 2005. — 6 p.
18. *TOGAF 8.1.1* Online. ADM and the Zachman Framework [Електронний ресурс]. — Режим доступу: <http://www.theopengroup.org/architecture/togaf8-doc/arch/chap39.html>
19. *Чертов О.Р.* Учетные и аналитически информационные системы в области статистики населения / О.Р. Чертов // Наукові праці. — Миколаїв: Вид-во ЧДУ ім. Петра Могили, 2010. — Т. 134. — Вип. 121. Комп'ютерні технології. — С. 225–229.
20. *Weikum G.* Transactional Information Systems: Theory, Algorithms, and the Practice of Concurrency Control and Recovery / G. Weikum, G. Vossen. — San Diego: Morgan Kaufmann, 2001. — 852 p.
21. *Thomsen E.* OLAP Solutions: Building Multidimensional Information systems. — 2-nd ed. / E. Thomsen. — N.Y.: John Wiley & Sons, 2002. — 661 p.
22. *Generic Statistical Business Process Model, Version 4.0* (April 2009) / Steven Vale (ed.). — METIS, 2009. — 28 p.

Надійшла до редакції 05.08.2010