

УДК 004.652

Методы разработки баз данных для информационных систем

В. И. Есин

Харьковский национальный университет им. В.Н. Каразина, Украина

Рассматривается проблема необходимости создания эффективной технологии разработки элементов информационной системы. В качестве решения предлагается новая технология, основывающаяся на метаонтологиях модели «объект-событие», использовании универсальной модели данных, языка модели данных, специального программного инструментария разработчика базы данных. Раскрываются принципиальные отличительные особенности предлагаемой технологии и методы ее реализующие.

Ключевые слова: модель данных, модель данных «объект-событие», универсальная модель данных, база данных, информационная система.

Розглядається проблема необхідності створення ефективної технології розробки елементів інформаційної системи. Як рішення пропонується нова технологія, що ґрунтується на метаонтологіях моделі «об'єкт-подія», використанні універсальної моделі даних, мови моделі даних, спеціального програмного інструментарію розробника бази даних. Розкриваються принципові відмінні особливості запропонованої технології і методи що її реалізують.

Ключові слова: модель даних, модель даних «об'єкт-подія», універсальна модель даних, база даних, інформаційна система.

The problem of necessity of creation of effective technology of development of elements of information systems is described. New technology, based on meta-ontology's of model "event object", using the universal data model, data model language, special software development tools database is propose as a solution. The principles distinctive features of the proposed technology and methods its realizing is described.

Keywords: data model, "object-event" data model, universal data model, database, information systems.

1. Постановка задачи

Процесс создания современных информационных систем (ИС) предполагает разработку базы данных (БД) и программных приложений. Причем и базы данных, и приложения информационной системы могут работать либо в конкретной имеющейся среде (с заданной конфигурацией аппаратных средств, топологией сети, используемой архитектурой и т. д.), либо в среде специально создаваемой для конкретной ИС. При этом система будет считаться успешной, если на момент ее запуска и в течение всего времени ее эксплуатации можно обеспечить [1]: требуемую функциональность и степень адаптации к изменяющимся условиям функционирования; требуемую пропускную способность; требуемое время реакции системы; безотказную работу системы; простоту эксплуатации и поддержки системы; необходимую безопасность. Поэтому, проектируя базу данных для информационной системы и соответствующие приложения работы с ней, необходимо, во-первых, предусмотреть определенную гибкость реализуемой системы: последняя должна обладать способностью не просто давать ответы на запросы, которые пользователь задает сегодня, но также в ней должна быть заложена возможность

предвидения той информации, которую он захочет получить завтра. Во-вторых, обеспечить требуемые пропускную способность и время реакции системы. И, в-третьих, строя или учитывая среду функционирования ИС, следует заботиться о безотказности, безопасности работы системы, простоте ее эксплуатации и поддержке.

Сегодня при разработке информационных систем в целом, и баз данных в частности, используются различные известные методы, такие как [1,2]: каскадная модель; поэтапная модель с промежуточным контролем; спиральная модель; модель эволюционной разработки (метод приращений), каждый из которых имеет свои области применения, достоинства и недостатки. Однако какой бы метод проектирования не был выбран, придется хотя бы однажды пройти все его стадии, независимо от того делаете ли вы это последовательно или итеративно, состоит ли задача в разработке системы в целом или ее отдельной части, является ли ваша технология формальной или неформальной.

И в этом случае эффективным решением могла бы быть определенная технология, которая: а) имела бы комплексное и общее решение, позволяющее разрабатывать основные составные элементы ИС: базы данных и программные приложения с ней работающие, интероперабельными. То есть способствовала бы использованию уже существующих компонентов, с целью отказа от экстенсивного ручного программистского труда в пользу интенсивных методов сборки, а не разработки их каждый раз уникальным способом под конкретное задание; б) позволяла бы прозрачно и в сжатые сроки разрабатывать любого уровня сложности новые БД и программные приложения с ней работающие; в) позволяла бы проводить реинжиниринг существующих баз данных, построенных на различных принципах, платформах и содержащих значительный объем разнообразных данных; г) упрощала и удешевляла бы процесс дальнейшей эксплуатации баз данных и ИС в целом.

2. Технология разработки баз данных

В результате проведенного анализа развития информационных технологий, классических методов проектирования баз данных, требований, предъявляемых к корпоративным информационным системам и БД, а также используемых в различных предметных областях (Про) моделей данных была разработана такая технология. Она, как совокупность предлагаемых методов не привязана к конкретным программным решениям и платформам, хотя сегодняшняя ее реализация связана с определенными системами и платформами.

Рассмотрим ее методы.

2.1. Метод разработки баз данных информационной системы на основе использования схемы базы данных с универсальной моделью данных

При создании данного метода была использована классическая технологии проектирования баз данных, являющаяся мощным течением «со своей философией, способами восприятия реальности и способами существования в ней» [3]. А именно, за основу были взяты базовые ее этапы, которые и получили дальнейшую конкретизацию и уточнения с учетом особенностей и возможностей семантической модели данных «объект-событие» [4] и

универсальной модели данных (УМД) [5,6,7]. То есть, опираясь на: а) основные этапы классической технологии проектирования баз данных; б) формализмы (метаонтологии) модели «объект-событие» и ее форму представления данных ПрО в виде связанного иерархического дерева (способствующих взаимопониманию между специалистами различных областей знаний и квалификации в процессе проектирования баз данных); в) использование универсальной модели данных (позволяющей разработчику базы данных любой ИС прозрачно и быстро создать ее логическую модель, которую впоследствии автоматически перевести в компьютерную среду); г) использование языка модели данных (позволяющего программистам больше внимания уделять продумыванию интерфейса и функциональности приложений, а специалистам предметной области более тесно взаимодействовать с программистами, ставя им конкретные задачи и контролируя их действия на протяжении всего жизненного цикла эксплуатируемой БД),

а также на программный инструментарий разработчика базы данных информационной системы и библиотеки подпрограмм разработчика приложений, был разработан метод построения баз данных для любой ИС.

Характеристика основных этапов метода разработки баз данных

Основные этапы метода разработки базы данных информационной системы для любой ПрО, основывающегося на использовании схемы базы данных с универсальной моделью данных, а также порядок их выполнения представлены на рис. 3.

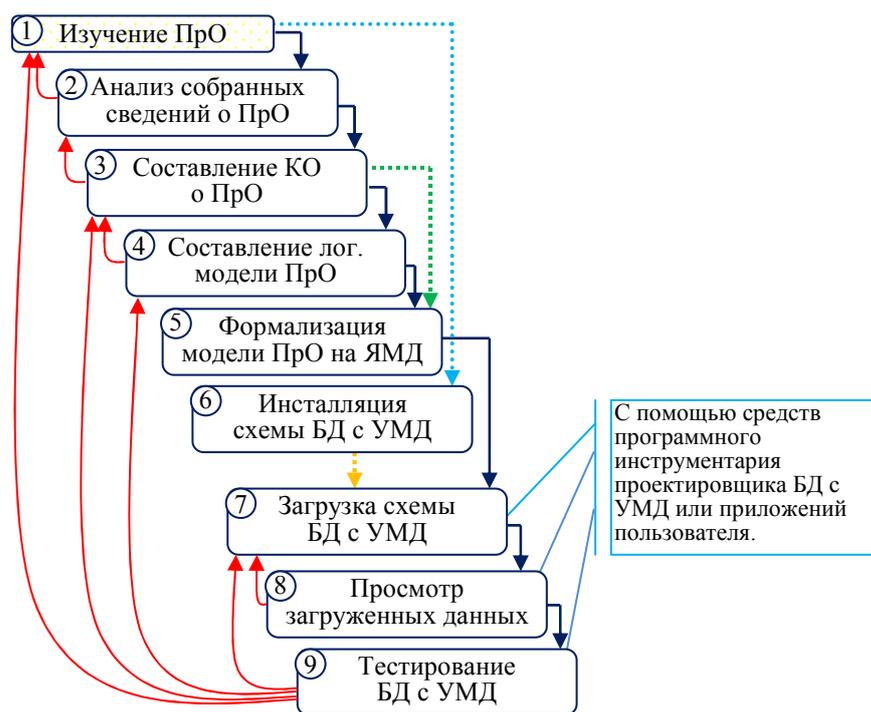


Рис. 3. Этапы метода разработки базы данных

Этап 1. Изучение (сбор сведений) предметной области и выделение ее основных значимых элементов. На естественном языке формулируется описание предметной области и ограничения, накладываемые на выбранные элементы ПрО (бизнес-правила).

Этап 2. Анализ собранных сведений о ПрО и их фиксация. Конкретизируются и формализуются множества значимых элементов ПрО путем отнесения их к одной из метаонтологий семантической модели «объект-событие» (раздел, класс объектов, класс событий, их экземпляры, характеристики и т. п.). Выясняется иерархия классов, экземпляров объектов ПрО, динамика событий с ними происходящими, их параметры и т. д. Определяются и формализуются ограничения целостности элементов ПрО.

Этап 3. Составление концептуального описания (КО) ПрО с помощью модели данных «объект-событие». В соответствии с правилами, изложенными в [8], составляются концептуальные описания метаданных и данных ПрО.

Зачастую этот и следующий этап недооцениваются проектировщиками БД (особенно начинающими) и воспринимаются ими как дополнительная, излишняя работа. Эта точка зрения абсолютно неверна. Во-первых, построение подробного и наглядного концептуального описания позволяет более полно оценить специфику моделируемой предметной области и избежать возможных ошибок на стадии проектирования схемы БД. Во-вторых, на этих этапах производится важная документация, которая может оказаться очень полезной не только при проектировании схемы БД, но и при эксплуатации, сопровождении и развитии уже сформированной БД.

Этап 4. Составление логической модели ПрО с помощью универсальной модели данных. В соответствии с правилами, изложенными в [9], составляется логическая модель ПрО.

Этап 5. Формализация модели предметной области с помощью ЯМД. На данном этапе с помощью языка модели данных (ЯМД) осуществляется формализованное представление модели ПрО, составленной на предыдущих двух этапах. Такое формализованное представление позволяет облегчить совместную и индивидуальную работу прикладных программистов, специалистов предметной области и конечных пользователей. А именно, ослабляются требования к знаниям прикладными программистами тонкостей модели БД и особенностей диалектов языка SQL запросов. При составлении приложений, имеющих доступ к данным схемы БД с УМД, прикладным программистам необходимо знать язык модели данных и структуру интерфейса доступа к серверным (хранимым) процедурам интерпретатора ЯМД. В результате чего они могут больше внимания и времени уделять продумыванию интерфейса и функциональности самих приложений. Специалисты предметной области, понимая простые операторы ЯМД, могут более тесно взаимодействовать с программистами при постановке им конкретных задач и контроле их действия на протяжении всего жизненного цикла эксплуатируемой ИС. Конечные пользователи также, понимая несложные операторы ЯМД, могут исполнять и составлять строки метаописания ЯМД самостоятельно, не прибегая к помощи программистов. Сами же строки метаописания, являющиеся формализованным представлением модели ПрО, могут использоваться как

определенная документация, контрольные примеры, так и исходные данные для тестирования БД с УМД, необходимые для эксплуатации и сопровождения базы данных.

Этап 6. Установка схемы БД с УМД. Данный этап может быть выполнен и в самом начале процесса разработки БД (для облегчения процесса установки схемы БД с УМД разработан специальный скрипт и инструкция по его использованию). То есть процесс разработки физической структуры БД любой ИС сводится не к разработке каждый раз новых объектов базы данных, а к установке стандартной схемы БД со структурой УМД, которая в дальнейшем наполняется элементами конкретной предметной области. Это позволяет разработчикам, больше сосредоточиться на ПрО, а, следовательно, лучше продумать состав необходимых данных в соответствии с реализуемыми бизнес-процессами.

Этап 7. Загрузка схемы БД с УМД реальными данными. Загрузка данными БД с УМД может осуществляться как с помощью специально разработанного (в рамках предлагаемой технологии) программного инструментария проектировщика БД, так и с помощью программных приложений, разработанных самостоятельно заказчиком или по его просьбе разработчиком индивидуально для конкретного пользователя.

Этап 8. Просмотр загруженных данных. Просмотр осуществляется с целью проверки их соответствия представлениям экспертов ПрО с помощью специальной программы из состава разработанного в рамках новой технологии программного инструментария разработчика БД с УМД, в которой данные представлены в виде связанного иерархического дерева. Анализируя такие представления, можно выявить несоответствия, отсутствующие или ошибочные связи между реальными данными ПрО и данными, занесенными в БД с УМД. После чего с помощью, опять же, предоставляемого программного инструментария легко провести перепроектирование, повторить загрузку реальных данных и получить новый вариант базы данных для анализа на адекватность предметной области. В результате нескольких итераций проектируется БД, адекватная представлениям о ПрО. При этом следует учесть, что итерационный характер носит не проектирование схемы БД и ее элементов, а заполнение сформированной уже схемы БД с ее элементами. Что позволяет существенно экономить временной и финансовый ресурс, соответственно выделяемый (затрачиваемый) на разработку базы данных.

Этап 9. Тестирование БД с УМД. Тестирование может проводиться как с помощью имеющегося программного инструментария разработчика БД с УМД, так и с помощью программного обеспечения (ПО), разработанного пользователем самостоятельно или сторонним разработчиком. Данный этап является очень важным не только в плане правильности функционирования последней, но и в плане адекватного моделирования ПрО в базе данных, обеспечения необходимой безопасности информации, хранящейся в БД с УМД, а также развития и улучшения ее концептуальной структуры в соответствии с новыми возможными изменениями в представлении предметной области.

Итерационное исполнение всех перечисленных этапов (в данном случае рассматриваемый механизм разработки представляет собой модель

эволюционной разработки [2]) гарантирует прозрачное и оперативное (в сжатые сроки) создание любого уровня сложности базы данных для любой предметной области, отвечающей функциональным потребностям заказчика. Кроме того, следуя предписаниям третьего, четвертого, да и пятого этапов предлагаемого метода, разработчик автоматически создает достаточно подробную и понятную всем участникам проекта (специалистам различного профиля и квалификации) документацию, описывающую различные уровни представления данных разрабатываемой БД. При этом следует отметить важный момент – разработчик специально не концентрируется и не тратит время на ее написание, она формируется автоматически (для ее размещения во вторичной памяти необходимо воспользоваться возможностями предоставляемого в рамках новой технологии инструментария). А роль ее значительна. Наличие документации полезно и на этапе разработки конкретной БД, и при эксплуатации, и при сопровождении, и при развитии уже функционирующей БД.

Использование предлагаемого метода позволяет развивать и улучшать концептуальную структуру БД рассматриваемой ПрО без изменения физической структуры схемы базы данных. Что может быть реализовано благодаря использованию либо специально разработанного программного инструментария проектировщика БД, либо программных приложений, разработанных самостоятельно заказчиком или по его просьбе разработчиком индивидуально для конкретного пользователя (см. реализацию седьмого и восьмого этапов метода).

2.2. Метод разработки баз данных, основывающийся на специальных типовых процессах реинжиниринга существующих баз данных в схему базы данных с универсальной моделью

Сегодня можно говорить, что эра, когда разработчики информационных систем приходили в организацию и начинали проекты информатизации «с нуля», прошла. Большинство организаций уже имеет некоторые информационные системы, которые со временем становятся бременем компании и начинают требовать реинжиниринга. Занятие реорганизацией информационных систем не является сегодня для разработчика БД чем-то необычным [10]. Это объясняется тем, что специалистам любой крупной и давно существующей компании, которая обладает несколькими базами данных, относящимися к разным видам деятельности, данные в которых могут иметь разные представления и быть даже несогласованными, становится очень трудно связывать и анализировать содержащуюся в них информацию. К тому же многие БД, составляющие основу таких ИС, как правило, построены на уже устаревших «платформах» (например, Dbase, FoxPro и т. д.). При этом данные, в них хранящиеся, имеют большую практическую ценность. Как результат, имея самые разнообразные изолированные друг от друга источники данных, найти необходимые для деятельности компании данные становится невозможным.

Поэтому чтобы упростить и ускорить доступ к такой информации, дать возможность проследивать связи между разными источниками данных и обеспечить использование для управления на всех уровнях, в том числе, и

унаследованных настольных баз данных, необходимо каким-то образом интегрировать эти данные.

На сегодняшний день в мире существует большое количество подходов, методов и технологических решений, напрямую или косвенно соотносимых с деятельностью по реинжинирингу ИС. Однако они не интегрированы на уровне методологий (процессов разработки). Как результат, можно наблюдать наличие огромного их количества, где основной акцент сделан на опять же разработку ИС «с нуля», и практическое отсутствие «стройных» методологий, целью создания которых являлось бы комплексное, целостное решение задач реинжиниринга ИС [11]. Поэтому ниже предлагается метод, с помощью которого можно решить эту проблему.

Характеристика основных этапов метода разработки баз данных

Предлагаемый метод разработки базы данных основывается на использовании схемы БД с УМД и специально определенных типовых процессах реинжиниринга существующих баз данных в схему базы данных с универсальной моделью. Основные этапы метода приведены на рис. 4.

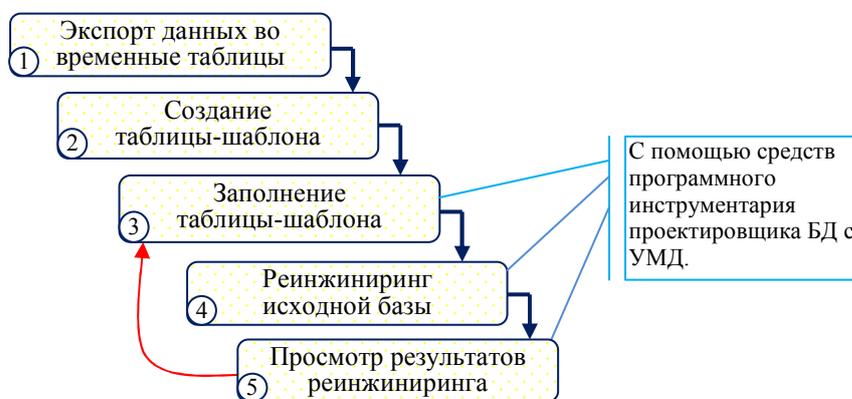


Рис. 4. Этапы метода разработки новой базы данных путем реинжиниринга существующей

Этап 1. Экспорт данных во временные таблицы. Данные базы данных, подвергающейся реинжинирингу, экспортируются в одну из СУБД: Oracle, PostgreSQL или Access. *Этап 2. Создание таблицы-шаблона.* В той СУБД, куда были экспортированы данные, создается, так называемая, таблица-шаблон, более подробно с которой можно ознакомиться в работе [12]. *Этап 3. Заполнение (формирование) таблицы-шаблона.* С использованием различных конструкций и операторов языка модели данных [13], заполняется таблица-шаблон. Некоторые особенности правил ее формирования приведены в работе [12]. *Этап 4. Реинжиниринг исходной базы данных.* На основании содержания таблицы-шаблона (данных ее строк и столбцов) и определенных строк (не всех сразу) определенных таблиц исходной БД, подвергающейся реинжинирингу, специально разработанное в рамках предлагаемой технологии приложение автоматически формирует корректные строки метаописания ЯМД, которые затем и исполняет. В результате таких действий данные, ранее хранящихся в исходной БД, помещаются в схему БД с УМД в виде определенных метаданных

и данных рассматриваемой ПрО. *Этап 5. Просмотр результатов реинжиниринга.* С помощью специально разработанной в рамках предлагаемой технологии программы просмотра проверяется правильность преобразования БД. В случае обнаружения несоответствий между описанием предметной области и занесенными в БД данными, последние с помощью различных приложений (предусмотренных технологией) удаляются. После чего вносятся требуемые изменения, помещаемые в таблицу-шаблон в виде соответствующих строк метаописаний ЯМД, и процесс преобразования (этапы 3...5) повторяется.

Такое повторение осуществляется до тех пор, пока не будет найдено адекватное отображение ПрО в схеме БД с УМД, способствующее, в том числе, и улучшению функциональности получаемой новой БД.

3. Выводы

1. Предлагаемый метод разработки баз данных информационных систем, основывающийся на использовании схемы базы данных с универсальной моделью, позволяет: 1) оперативно создавать базы данных отвечающие функциональным потребностям заказчика, для любой предметной области; 2) создавать достаточно подробную документацию, описывающую различные уровни представления данных разрабатываемой БД и понятную специалистам различного профиля и квалификации; 3) развивать и улучшать концептуальную структуру базы данных для конкретной информационной системы, рассматриваемой предметной области, без изменения физической структуры схемы базы данных.

2. Созданный в рамках новой технологии проектирования баз данных метод разработки базы данных, основывающийся на специальных типовых процессах реинжиниринга существующих баз данных в схему базы данных с универсальной моделью, в отличие от традиционных методов имеет комплексное, целостное решение задач реинжиниринга ИС на уровне методологии, основывающейся на универсальной модели данных. Благодаря разработанному методу стало возможным использование данных «устаревших» информационных систем в системе, основанной на новой технологии с последующим возможным упрощением комплексирования появляющихся новых программных систем на основе использования готовых компонентов со стандартными интерфейсами (сохранение интероперабельности) и получение принципиально новых дополнительных функциональных возможностей, какие ранее были недоступны.

Использование данного метода способствует расширению функциональности вновь создаваемой базы данных и улучшению других количественных и качественных ее характеристик, в том числе снижается стоимость сопровождения, вероятность возникновения значимых для заказчика рисков, уменьшаются сроки работ по ее сопровождению.

ЛИТЕРАТУРА

1. Хаф Л. Проектирование информационных систем [Электронный ресурс] / Л. Хаф. – Режим доступа : http://www.info-system.ru/is/theory/theory_design_is.html#15

2. Райордан Р. Основы реляционных баз данных Пер. с англ. / Р. Райордан – М. : Издательско-торговый дом «Русская Редакция», 2001. – 384 с.
3. Зиндер Е.З. Проектирование баз данных: новые требования, новые подходы [Электронный ресурс] / Е. З. Зиндер. – Режим доступа: <http://www.citforum.ru/database/kbd96/41.shtml>
4. Есин В. И. Семантическая модель данных «объект-событие» / В. И. Есин // Вісник Харківського національного університету. – Х. : Харьковский национальный университет им. В.Н. Каразина, 2010. – № 925. – С. 65-73 – (Серія : "Математичне моделювання. Інформаційні технології. Автоматизовані системи управління" ; вип. 14).
5. Есин В. И. Универсальная модель данных и ее математические основы / В. И. Есин // Системи обробки інформації. – Х. : Харківський університет Повітряних Сил, 2011. – № 2(92) – С.21–24.
6. Есин В. И. Универсальная модель данных и ее отличительные особенности / В. И. Есин // Вісник Харківського національного університету. – Х. : Харьковский национальный университет им. В.Н. Каразина, 2011. – № 960. – С. 141–147 – (Серія : Математичне моделювання. Інформаційні технології. Автоматизовані системи управління" ; вип. 16).
7. Есин В. И. Основные компоненты универсальной модели данных / В. И. Есин // Вісник Харківського національного університету. – Х. : Харьковский национальный университет им. В. Н. Каразина, 2012. – № 1015. – С. 131–141 – (Серія : "Математичне моделювання. Інформаційні технології. Автоматизовані системи управління" ; вип. 19).
8. Есин В. И. Формы представления данных в модели «объект-событие» / В. И. Есин // Системи управління, навігації та зв'язку. – К. : ДП «Центральний науково-дослідний інститут навігації і управління», 2012. – № 1(21), том 2 – С.198–201.
9. Есин В. И. Метод моделирования предметной области с помощью универсальной модели данных / В. И. Есин // Системи озброєння і військова техніка. – Х. : Харківський університет Повітряних Сил, 2011. – № 2(26) – С.128–131.
10. Харрингтон Д. Л. Проектирование реляционных баз данных / Д. Л. Харрингтон. – М. : Издательство "Лори", 2006. – 230 с.
11. Ахтырченко К. В. Методы и технологии реинжиниринга ИС / К. В. Ахтырченко, Т. П. Сорокваша // Труды Института системного программирования РАН. / Под редакцией чл.-корр. РАН В. П. Иванникова. – М. : ИСП РАН, 2003. – Т. 4. – С. 141–162.
12. Есин В. И. Реинжиниринг существующих баз данных / В. И. Есин // Системи обробки інформації. – Х. : Харківський університет Повітряних Сил, 2012. – № 3(101), т. 2 – С.188–191.
13. Есин В. И. Язык для универсальной модели данных / В. И. Есин, М. В. Есина // Системи обробки інформації. – Х. : Харківський університет Повітряних Сил, 2011. – № 5(95) – С.193–197.