

**Модель обобщенного репозитория метаданных корпоративной
информационной среды вуза**
**Model of common metadata repository for university's information
environment**

К.И. Шахгельдян

В работе рассматривается модель обобщенного репозитория метаданных, предназначенного для реализации схем интеграции данных, управления бизнес-процессов, необходимых для развития современной, управляемой информационной среды вуза.

The model of common metadata repository is the main subject of the article. The repository is intended to implement the schemes of data and business-processes integrations which are needed for modern, controlled university's information environment.

Введение

Информатизацию вуза можно рассматривать как процесс перехода к такому информационному университетскому обществу, в котором не только предоставлен доступ к качественной информации, но и информационные технологии (ИТ) являются обязательным средством выполнения своих должностных обязанностей сотрудниками вуза и средством обучения студентов. При этом должностные обязанности выполняются с большей эффективностью, а обучение студентов с большим качеством, чем без использования ИТ.

Корпоративная информационная среда (КИС) вуза является основным инструментом процесса информатизации. КИС вуза – это своеобразная проекция процедур и объектов деятельности вуза на плоскость ИТ. КИС состоит из информационной инфраструктуры, данных, приложений, обрабатывающих эти данные, и пользователей, оперирующих этими приложениями.

КИС, согласно общей теории сложности [1], в полной мере относится к сложным системам. Одна из проблем сложных информационных систем – это проблема сопровождения, которая в КИС связана с обеспечением качественной информации, с поддержкой модификации моделей данных и информационных систем и с поддержкой часто меняющихся бизнес-процессов (БП) вуза. Плохо управляемая среда в определенный момент становится не двигателем и поддержкой прогресса, а его стопором. Поэтому сопровождение системы - ее способность обеспечивать доступ к качественной информации, поддерживать быструю смену контингента пользователей, своевременную автоматизацию изменений в существующих БП и вновь появившихся процессов выходит на первый план вопросов развития КИС.

Основной концепцией построения КИС вуза является идея интеграции. Аспекты интеграции, которые необходимо рассматривать при построении КИС – это интеграция данных, приложений и БП.

В настоящее время одним из подходов к решению проблем сопровождения сложных информационных систем является подход, основанный на описании проблемной области в репозитории метаданных [2-4]. В частности репозиторий используется для обеспечения качества информации [2], для поддержки интеграции данных [3] или для управления БП [4].

В этой работе мы рассмотрим модель обобщенного репозитория метаданных, который предназначен одновременно для решения задач обеспечения качества информации, для управления БП и для обеспечения интеграции данных.

Подход к КИС, основанный на понятиях

Для того, чтобы КИС быстро и адекватно реагировала на изменяющиеся БП деятельности вуза, чтобы была действительно необходимым инструментом сотрудников и студентов, КИС должна быть управляема, настраиваема бизнес-аналитиками, т.е. специалистами предметниками, а не только программистами и администраторами. Для этого необходимо, чтобы КИС оперировала терминами предметных областей деятельности вуза.

Для этого КИС можно представить как *совокупность автоматизированных БП, оперирующих понятиями предметной области*. К понятиям предметной области относятся, например, *Сотрудник, Организация, Подразделение, Адрес, Студент, Дисциплина, Информационный ресурс, Телефон, Аудитория, Здание, Пластиковая карта, Договор, Тест, Расписание* и т.п. Эти понятия используются специалистами-предметниками – учебным управлением, деканатами и кафедрами, управлением планирования и бухучета, управлением персоналом, службой безопасности и т.д. Понятия ИТ-области - *Сервер, База данных, Пользователь, Проект, Роль, Фильтр, Атрибут, Таблица, Метод, Параметр, Компонент* и т.п. - так же являются частью КИС. Эти понятия используются ИТ-специалистами.

Понятие включает атрибуты, другие понятия и методы. Атрибуты и агрегированные понятия определяют внутреннее содержание понятия. Одно и то же понятие может быть агрегированным в несколько разных понятий. Между понятиями могут существовать отношения типа агрегации, когда понятие включает другие понятия, и отношения ассоциации, когда понятия ассоциируются друг с другом.

Понятие в большинстве случаев имеет проекцию на область базы данных, т.е. оно связано с некоторой таблицей или представлением. Одно понятие обычно имеет одну базовую таблицу или представление. Но полное описание понятия может включать несколько таблиц. Атрибуты

отображаются на переменные примитивных типов и обычно хранятся в базовой таблице понятия, являясь одним из полей этой таблицы (или представления). Агрегированные понятия хранятся в таблице (или представлении), не совпадающей с базовой.

Описание всех понятий КИС хранится в обобщенном репозитории метаданных (ОРМД). Понятия отличаются друг от друга по имени, в рамках КИС (или точнее в рамках ОРМД) имя понятия должно быть уникальным.

Связи между понятиями могут осуществляться на логическом и физическом уровне. Если связь между понятиями определена на уровне СУБД (в реляционных базах данных используется для этой цели концепция внешних ключей), то это физическая связь, если связь между понятиями определяется на уровне логики программ, работающих с понятиями, то такая связь называется логической. Так как таблицы, связанные с понятиями, могут располагаться на разных серверах баз данных, то невозможно использовать средства СУБД, чтобы описать эти связи, и в этом случае связь всегда будет логической.

Определены связи между понятиями 3-х типов:

1. связь один к одному подразумевает, что понятие *A* связано только с одним экземпляром понятия *B*;
2. связь один ко многим подразумевает, что понятие *A* может быть связан с несколькими экземплярами понятия *B*;
3. наследование подразумевает, что понятие *B* содержит все те же атрибуты, что и понятие *A*, но множество экземпляров понятия *B* является подмножеством экземпляров понятия *A*.

Связи между понятиями описываются указаниями на левый и правый источник связи (*A* и *B*), источник данных, в котором осуществляется связь, атрибуты связи, по которым сопоставляется понятия *A* и *B*, условия связи. При наличии дополнительных атрибутов у связи описывается понятие, которое содержит атрибуты связи между понятиями. Условия связи определяются по условиям, накладываемым на атрибуты. Такие условия можно использовать, когда в одном источнике связывается понятие *A* с понятиями *B* или *C* в зависимости от некоторого атрибута связи.

В ОРМД должны быть представлены как физические, так и логические связи. При этом физические связи получаются автоматически из метаописания СУБД и только дополняются администратором, а логические - вручную прописываются администратором ОРМД.

В ОРМД описаны так же связи между источниками данных, серверными компонентами и проектами [5]. Эти описания позволяют оценивать связи между пользователями, понятиями предметной области, источниками данных, серверными компонентами и проектами.

Логические отношения между понятиями

Логические отношения между понятиями (связи между понятиями на логическом уровне) могут быть контекстно-зависимые и контекстно-независимые. Контекстно-независимые отношения являются статическими и не изменяются во время работы с экземплярами понятий. Это означает, что если понятие A связано с понятием B , то A и B логически связаны все время своего существования.

Контекстно-зависимые отношения позволяют определить связи между понятиями, зависящие от атрибутов понятия. Т.е. в зависимости от значения одного или нескольких атрибутов понятия A , он будет связан с одним из понятий некоторого множества понятий $B = \{B_i\}_{i=1}^M$. Контекстно-независимые отношения – это отношения между понятиями, а контекстно-зависимые – это в большей степени отношения между экземплярами понятий.

В связи с контекстно-зависимыми отношениями необходимо определить абстрактное понятие. Понятие B , агрегированное в понятие A , называется *абстрактным*, если оно не связано ни с какой таблицей или представлением, и всегда заменяется на другие понятия, при работе с экземпляром понятия A .

В общем случае понятие A , содержит атрибут C и абстрактное понятие B , которое может быть заменен на один из элементов множества $D = \{D_i\}_{i=1}^N$, в зависимости от атрибута C .

$$A \rightarrow B = \{D_i, A \rightarrow C \in S_i\}_{i=1}^N$$

Здесь S_i - это множество допустимых значений атрибута C . Понятие D_i является частью некоторого понятия P . В этом случае понятия A и P связаны контекстно-зависимой связью.

Понятия и управление бизнес-процессами

Для управления БП необходимо разработать схему управления БП на основании понятий и атрибутов. ОРМД содержит не только описания понятий, которые используются в управлении БП, но и описания БП и схем его функционирования.

Целью такого подхода является создание инструмента, который бы позволил бизнес-аналитикам, специалистам предметникам самими формировать БП, определять порядок и условия следования действий в БП, а так же задавать бизнес-правила. Такой подход позволяет своевременно реагировать на изменения БП в системах их управления, а так же позволяет разработчикам КИС создавать новые информационные системы.

Введем понятие *элементарный БП*. Элементарные БП реализуют простую логику, которая реализована в программах, процедурах, методах

серверных компонент или как-то иначе. Элементарные БП описаны в ОРМД.

Пользователи КИС могут формировать составные БП, состоящие из элементарных БП. В общем случае составные БП состоят из других составных БП и элементарных БП. Понятие «состоять из» означает, что внутри составного БП элементарные и другие составные БП могут выполняться последовательно или параллельно.

Последовательность действий может управляться блоком условий. Они могут комбинироваться по «И», «ИЛИ» или «Не». В условиях используются понятия КИС и их атрибуты.

Доступ к элементарным БП внутри составного БП определены для ролей пользователей проектов [5]. Это означает, что пользователь, который имеет роль может или должен выполнить действие, определенное элементарным БП.

Когда некоторый БП вуза изменяется должны быть изменены и составные БП. В некоторых случаях требуется разработка новых элементарных БП. Новые элементарные БП описывают в ОРМД. В определенных случаях изменения БП приводит к необходимости вводить новые понятия, тогда от администратора КИС требуется описание этих новых понятий в ОРМД с привязкой к новым таблицам и представлениям.

Реализация элементарных БП выполняется как разработчиками, так и в отдельных простых случаях бизнес-аналитиками. Использование БП, создание составных БП, определение условий выполняется бизнес-аналитиком.

Понятия, интеграция данных и качество информации

Для поддержания интеграции данных используются описания контекстно-зависимых и контекстно-независимых связей между понятиями в ОРМД. ОРМД позволяет использовать в проектах понятия, созданные в другом проекте, поддерживать общую систему справочников для всех проектов КИС, а так же общую систему отчетности КИС. При составлении отчетов пользователи используют определения понятий из ОРМД, что позволяет КИС формировать произвольные, в том числе агрегированные отчеты из данных различных проектов без привлечения программистов.

Качество информации предполагает поддержание в КИС полной, непротиворечивой, актуальной и достоверной информации. Для этого в КИС вводятся критерии качества информации. Эти критерии определяются через понятия и атрибуты ОРМД, где они и хранятся. В том числе в ОРМД хранятся ссылки на все процедуры и сервисы, обеспечивающие поддержания качественной информации.

Реализация ОРМД

Во Владивостокском государственном университете экономики и сервиса разработана КИС, важным элементом которой в последний год

стал ОРМД, построенный на базе модели, рассмотренной в работе. С помощью ОРМД реализованы системы планирования учебным процессом и система управления бизнес-процессами, а так же системы поддержания качества информации.

В качестве примера контекстно-зависимых отношений в КИС ВГУЭС можно рассмотреть описание пользователей КИС, которыми являются – студенты, сотрудники, внешние пользователи, бывшие студенты, бывшие сотрудники, проекты и серверные компоненты. В зависимости от категории пользователя сведения о нем хранятся в различных базах данных на различных серверах.

Если в КИС необходимо ввести новый тип пользователей, например, работодатель, то для этого описывается новое понятие *Работодатель*, определяется источник данных и методы извлечения. Все остальные информационные системы КИС не меняются.

В качестве примера формирования блока условий составного БП можно рассмотреть процесс формирования дисциплин. В зависимости от атрибутов дисциплины { *форма обучения; уровень образования; технология образования* } пути следования в БП создания, согласования и утверждения дисциплины будут различными. В описании блока условий в ОРМД должны быть зафиксированы явно или косвенно все сочетания из

форма обучения × уровень образования × технология образования.

Для поддержания качества информации в КИС ВГУЭС используются многочисленные процедуры и сервисы, которые актуализируют данные, выполняются проверки на полноту, непротиворечивость. Эти процедуры используют описания понятий, их связей с источниками данных из ОРМД.

Поддержка обобщенной системы справочников КИС ВГУЭС так же использует ОРМД.

Список литературы

- [1] Урманцев Ю.А. Общая теория систем: состояние, приложения и перспективы развития//Система, Симметрия, Гармония. М., Мысль, 1988, с.38-124
- [2] Wang, R., Kon, H., Madnick, S. Data Quality Requirements Analysis and Modelling//9-th International Conference of Data Engineering, Vienna, Austria.- 2003
- [3] Гринев М., Кузнецов С., Фомичев А. Особенности СУБД Sedna. XML-СУБД Sedna: технические особенности и варианты использования//Открытые системы.-2004.-№8.
- [4] Черняк Л. ВМР – близкие перспективы и далекие горизонты//Открытые системы.- 2004.-№ 11–с. 23-29
- [5] Шахгельдян К.И., Крюков В.В., Гмарь Д.В. Система автоматического управления правами доступа к информационным ресурсам вуза//Информационные технологии №2 2006.-с.19-29.

Научный консультант: д.ф.-м.н., профессор, Клещев А.С. зав. отделом
ИАПУ ДВО РАН