

## МОДЕЛЬ ДАННЫХ РЕПОЗИТОРИЯ ПОНЯТИЙ КОРПОРАТИВНОЙ ИНФОРМАЦИОННОЙ СРЕДЫ ВУЗА

Игнатова Ю.А., Гмарь Д.В., Шахгельдян К.И. к.т.н., доцент

e-mail: [Yuliya.Ignatova@vvsu.ru](mailto:Yuliya.Ignatova@vvsu.ru)

ГОУ ВПО «Владивостокский государственный университет экономики и сервиса»

г. Владивосток, Приморский край

### Аннотация

Статья посвящена методам организации работы с корпоративными данными, программной реализацией одного из которых является репозиторий понятий информационного пространства вуза. В настоящей статье рассматривается модель данных репозитория понятий КИС ВГУЭС, в частности описывается структура и способы поддержания целостности данных, а так же основные принципы их извлечения и манипулирования в рамках этой модели.

## THE CONCEPTS REPOSITORY DATA MODEL OF UNIVERSITY CORPORATIVE INFORMATION ENVIRONMENT

Ignatova Y.A., Gmar D.V., Shakhgeldyan K.I.

### Annotation

The concepts repository of university corporative information environment as a data management method is the main subject of the article. The concepts repository data model, which describes the aspects of data structure, data integrity and data processing, is considered in this article.

Корпоративная информационная среда (КИС) Владивостокского государственного университета экономики и сервиса (ВГУЭС) помимо совокупности телекоммуникационного, сетевого оборудования и программного обеспечения, включает так же и данные, которые обрабатываются приложениями КИС и с которыми взаимодействуют пользователи КИС. В процессе организации работы с такими данными часто требуется оперировать некоторыми понятиями информационной среды, не создавая при этом отдельных программных средств управления этими понятиями. Для поддержания рассматриваемого процесса необходимо удобное и простое в использовании единое инструментальное программное средство, позволяющее эффективно организовать создание, хранение, представление и управление понятиями информационной среды. Для таких целей в КИС ВГУЭС используется репозиторий понятий. Подсистема хранения данных репозитория содержит полное описание понятий информационной среды, которые являются фундаментальными и служат основой для создания их экземпляров. Данная подсистема предназначена главным образом для хранения и поддержки экземпляров понятий. Система управления репозиторием регламентирует процесс порождения и согласования экземпляров понятий, включая создание их атрибутов, ограничений и отношений с другими понятиями информационного пространства. Каждый из экземпляров понятий подразумевает хранение условного постоянного набора данных, которые

вводятся однократно и в дальнейшем могут быть дополнены (расширены), с целью использования в других информационных системах КИС в качестве базовых.

Поскольку репозиторий должен обеспечивать работу с динамически увеличивающимся количеством экземпляров понятий информационного пространства, имеющих при этом переменную структуру (различный состав атрибутов различных типов), подсистема хранения данных репозитория должна отвечать не столько требованиям оптимальности, сколько универсальности. Так, использование классической методики проектирования баз данных [1], – предполагающей создание отдельной таблицы в базе для каждой описываемой сущности и далее, в процессе нормализации, выделение индивидуальных таблиц для хранения атрибутов сущностей, – потребовало бы постоянной реорганизации базы данных репозитория и, со временем, привело бы к неограниченному разрастанию количества ее объектов. В связи с этим, в качестве схемы проектирования базы данных репозитория был выбран один из вариантов реализации модели “Сущность-атрибут-значение” (EAV, «Entity-attribute-value» data model [2]) на основе реляционной СУБД (среда MS SQL Server 2000). Модель EAV представляет собой некую универсальную структуру описания данных, которая в самом общем смысле подразумевает хранение всех данных в виде так называемых таблиц-справочников: справочника сущностей, справочника атрибутов и справочника значений. Для различных типов данных (строковых, числовых и т.д.) выделяются отдельные таблицы-справочники значений, в простейшем случае используется одна таблица с применением наиболее общего формата данных. Справочник значений, помимо хранения всех данных, обеспечивает взаимосвязь с остальными таблицами-справочниками, поскольку содержит поля метаданных для ссылки на справочники сущностей и атрибутов, при этом каждый из атрибутов определенной сущности отображается на одну запись в справочнике значений. Таким образом, весь массив хранимых данных, образованный совокупностью значений таблиц-справочников, организуется фактически в один линейный список, что обеспечивает принципиальное единообразие хранения всех данных, а, следовательно, и существенную гибкость (настраиваемость) системы, разрабатываемой на основе данной модели.

В рамках рассматриваемого подхода, база данных репозитория понятий КИС ВГУЭС включает совокупность следующих реляционных таблиц (рис.1):

- таблица D\_DictList – используется в качестве справочника сущностей, ключевое поле IdObjectDepend однозначно идентифицирует каждый из экземпляров понятий информационного пространства. Поле IdDictType – определяет принадлежность экземпляра к классу внешних или внутренних понятий по отношению к локальному источнику данных репозитория; поле IdObjectGroup – ссылка на справочник групп экземпляров понятий;
- таблица D\_Dictionary – используется в качестве справочника групп экземпляров понятий информационного пространства, созданных с целью объединения экземпляров по тематическому признаку; ключевое поле Ids

обеспечивает взаимосвязь между понятиями на уровне их метаописания (тип связи понятие-родитель → понятие-потомок);

- таблица D\_Params – используется в качестве справочника атрибутов, ключевое поле IdParam уникально идентифицирует каждый из атрибутов экземпляров понятий, поле IdType – определяет тип данных атрибута, поле IdObject – используется для ссылки на справочник сущностей;
- таблица D\_Item – используется в качестве справочника групп значений атрибутов (записей). Поскольку доминирующей задачей репозитория является возможность использования экземпляров понятий информационного пространства для хранения наборов данных - т.е. каждому атрибуту определенного экземпляра соответствует более одного (множество) значения - появляется необходимость определять группы взаимосвязанных значений атрибутов в рамках отдельных записей. Для идентификации записей используется ключевое поле ItemId, поле IdObject – для ссылки на справочник сущностей;
- таблица D\_ParamDetails – используется в качестве справочника значений для хранения всего массива данных. Поле ParamText непосредственно используется для хранения данных строкового типа (длиной до 8000 символов), поле ParamMoney – для числовых данных, поле ParamDate – для данных типа дата/время, поле ParamValue – для организации связей с другими понятиями (устанавливается взаимосвязь между понятиями на уровне данных, тип атрибут родительского понятия→атрибут понятия-потомка), поле IdObject – ссылка на справочник сущностей, поле IdParam – ссылка на справочник атрибутов, поле InfoId – ссылка на справочник записей;
- таблица D\_LongData – используется для хранения текстовых блоков данных длиной до 2147483647 символов.

Логическая схема данных (рис.1), помимо описания собственно структуры, налагает на данные ряд определенных условий (ограничений), выполнение которых обеспечивает поддержание целостности данных в базе. В частности, ограничения целостности сущностей (первичные ключи и уникальные индексы), ограничения целостности атрибутов (значения по умолчанию, обязательность и необязательность значений Null\Not Null), а так же ограничения целостности типов данных, выполняются на декларативном уровне средствами СУБД. Правила описания понятий и организации их взаимосвязей различного типа и различного уровня (связи «понятие→понятие», «понятие→атрибут», «атрибут→атрибут» и т.д.), в т.ч. ограничения ссылочной целостности (целостность внешних ключей), принципы добавления, удаления и изменения данных, а так же дополнительные ограничения, не определенные логической схемой, выполняются на уровне прикладных программ. Логика выделяется в отдельный программный слой, реализуемый на сервере приложений, и, совместно с интерфейсом пользователя (web-интерфейсом), представляет систему управления репозиторием. Конечные наборы данных (данные самого низкого уровня, собственно пользовательские данные)

содержатся в экземплярах понятий, являющихся, в свою очередь, объектами системы управления репозиторием.

Логическая схема данных репозитория понятий КИС ВГУЭС, стандарт IDEF1X, ERWin [3]

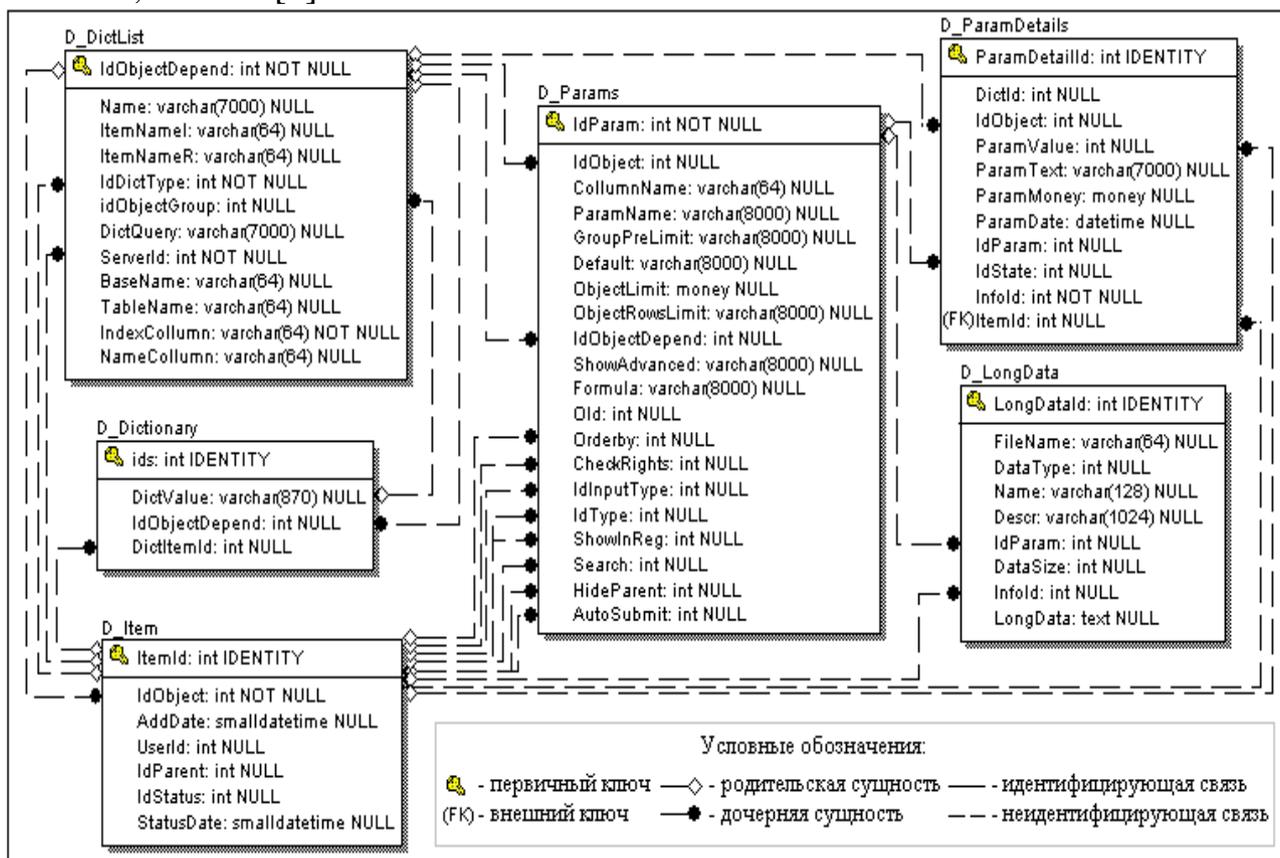


Рис.1

При построении интерфейса пользователя реализуется механизм по загрузке свойств и связей объектов системы из базы данных репозитория, согласно правилам описания соответствующих им экземпляров понятий, обеспечивая, таким образом, возможность работать с понятиями в привычном для пользователя «развернутом» табличном виде. При этом для каждого понятия в web-интерфейсе создается (отрисовывается) отдельная таблица и предоставляется набор элементов управления для манипулирования понятиями и их свойствами, а так же сохранению внесенных изменений в базе данных репозитория. Для построения системы управления репозиторием и реализации вышеуказанных его функций используется набор служебных объектов системы, оформленных в виде ее компонентов, которые могут работать с информацией о произвольных экземплярах понятий информационного пространства, что позволяет компоновать интерфейсные решения для работы с определенными экземплярами понятий. В виду однообразности модели хранения данных, служебные объекты системы управления (вплоть до справочников видов отрисовки элементов, типов полей, стилей отображения и т.п.) создаются в качестве понятий информационного пространства и хранятся в базе данных репозитория наряду с другими понятиями, таким образом, позволяя построить систему управления репозиторием на основе собственной же составной части,

т.е. рекурсивно. Такой подход позволяет не создавать в базе данных репозитория отдельных таблиц метаданных для описания структуры системы управления и организации ее web-интерфейса, а так же не разрабатывать отдельных программных средств по работе с этими таблицами.

Пользователи КИС работают с данными репозитория понятий через web-интерфейс системы управления, доступный в качестве одного из сервисов корпоративного информационного портала вуза, при этом прямой доступ к таблицам базы данных для них запрещен, и структура хранения данных скрыта. Пользователь, переходя к электронному каталогу объектов системы управления репозитория, и, обращаясь к определенному объекту, запрашивает статическую web-страницу в формате ASP. В ответ на запрос пользователя сервер приложений (web-сервер MS IIS) модифицирует соответствующую (заранее подготовленную) страницу-шаблон, используя параметры полученного запроса, и при помощи сценариев VBScript (используя свойства и методы встроенных объектов) производит подключение и выборку данных из базы репозитория. С помощью этих данных, генерируется видимое содержимое результирующей web-страницы путем вставки результата выполнения сценария в виде стандартных HTML-дескрипторов и текста. В результате пользователю в ответ на запрос посылается HTML-код, который был специально сгенерирован для него.

#### Библиографический список:

1. Основы проектирования реляционных баз данных, Кириллов В.В., СПб.: ИТМО, 1997, url: <http://khpi-iip.mipk.kharkiv.edu/library/dbms/kir1/>
2. The EAV/CR Model of Data Representation, Nadkarni PM, J Am: Med Inform Assoc, 2000, url: [http://ycmi.med.yale.edu/nadkarni/eav\\_CR\\_contents.htm](http://ycmi.med.yale.edu/nadkarni/eav_CR_contents.htm)
3. BPwin и ERwin. CASE-средства разработки информационных систем, Маклаков С.В., М., Диалог-МИФИ, 2001
4. -