

О.А. Бубарева

Центр информационных систем Бийского технологического института (филиал
АлтГТУ)

ПРОГРАММНАЯ СИСТЕМА ИНТЕГРАЦИИ ДАННЫХ

Цель проекта

Цель проекта – создание программного инструмента интеграции данных информационных систем. В основе проекта лежит задача - установление взаимодействия информационных систем, ориентированных на различные пересекающиеся предметные области.

На сегодняшний день организации стремятся эффективно распоряжаться, управлять имеющимися у них данными. Для успешного управления данными в организациях создаются интегрированные автоматизированные информационные системы (ИАИС), обеспечивающие поддержку различных бизнес-процессов. Получение интегрированной информации зависит от эффективного взаимодействия входящих в структуру ИАИС информационных систем (ИС) с различными стандартами описания и представления данных. На текущий момент уровень интеграции таких систем весьма невысок. Бизнес-процессы часто автоматизируются с использованием программных решений собственных и сторонних разработчиков без учета их взаимосвязанности. При постоянных изменениях в этих бизнес-процессах разработчики ИАИС вынуждены заниматься корректировкой программ и моделей данных, что приводит к структурной и семантической неоднородности информационных элементов и, соответственно, необходимости повторной разработки программных конвертеров данных. Использование таких решений ведет к усложнению и, следовательно, снижению надежности ИАИС.

Общее описание проекта

Рассматриваемая программная система позволяет автоматически определять и устанавливать семантические связи двух онтологий с последующим установлением взаимосвязи схем интегрируемых ИС, а также выполнять алгоритм формирования запросов на добавление информации из одной системы в другую. На основе связи онтологических понятий связываются релевантные элементы объектных схем.

В структуру программной системы входят следующие компоненты (рис. 1):

- модуль создания и изменения онтологических описаний, отвечающий за создание и изменение понятий и иерархий в базовой онтологии;
- репозиторий метаданных, обеспечивающий хранение описания онтологий предметных областей;
- модуль сопоставления и интеграции онтологий, выполняющий слияние онтологий разных предметных областей;
- модуль разрешения конфликтов, возникающий при сопоставлении концептов онтологий;
- модуль формирования шаблона запросов на добавление информации. На основе связи онтологий создается SQL-запрос на добавление информации из интегрирующей ИС в главную;
- пользовательский интерфейс.

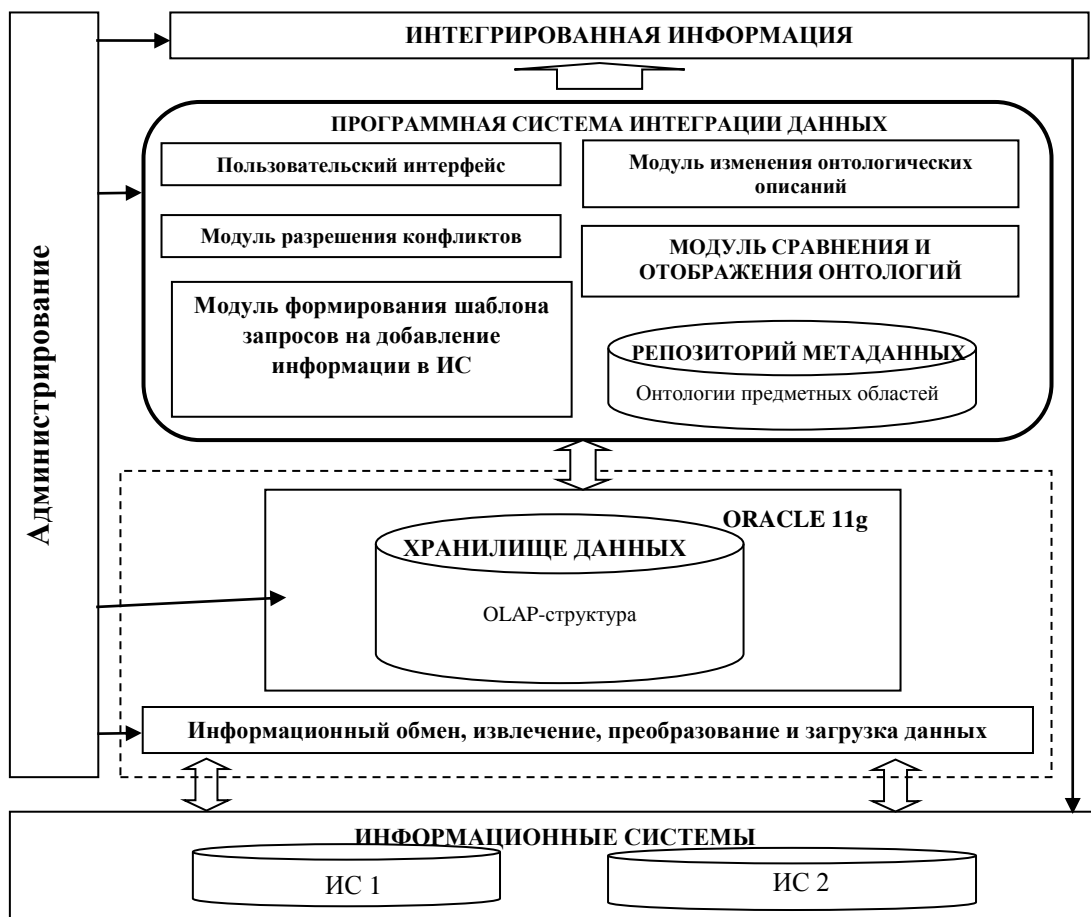


Рисунок 3.1 – Структура работы программы по интеграции ИС

Можно выделить следующие этапы работы программы, которые являются ключевыми.

1. Подключение к серверу БД интегрируемых ИС и извлечение онтологий этих ИС в формате OWL. Одна из выбранных онтологий указывается основной (главной), в которую необходимо произвести отображение концептов второй онтологии.

2. Устанавливаются семантические связи понятий онтологий для определения их сходства по определенным критериям.

На этом этапе пользователь системы имеет возможность дополнить и расширить онтологию. Но изначально все изменения, которые делает пользователь, сохраняются в его личной копии онтологии. Затем клиентская часть передает эту онтологию на сервер, где и происходит слияние. При возникновении нецелостности онтологии либо других конфликтов, ответственным за разрешение их является менеджер по работе со знаниями.

3. Интеграция онтологий неоднородных информационных систем.

4. Установление взаимосвязи ИС на структурном уровне, основываясь на согласованной семантике предметной области.

5. Интеграция данных. Формирование шаблона запросов на добавление информации из интегрируемой информационной системы в исходную.

Как следствие, обеспечивается совместная работа неоднородных ИС в контексте предметной области задачи на семантически значимом уровне.

Отличия от аналогов

Программная система интеграции данных неоднородных ИС с учетом сопоставления онтологий в целом лишена многих недостатков, присущих другим программам, в основе которых лежат чисто технические методы, и предоставляет возможность разработки интегрированных ИС, работающих с информацией на семантическом

уровне. Разработанная программа позволяет преодолеть семантическую разнородность информационных источников, то есть различие схем данных (онтологий) ИС. В отличие от существующих методов интеграции данных в нашем случае не требуется предварительной репликации или индексации сведений из информационных источников – вычисляемый системой ответ включает исключительно актуальные сведения, полученные непосредственно из источников данных.

Научная составляющая

В основе рассматриваемой программной системы лежит новая математическая модель интеграции данных информационных систем с неоднородными онтологическими спецификациями, учитывающая семантические связи, а также новый алгоритм сопоставления и интеграции неоднородных онтологических спецификаций ИС, особенностью которого является обеспечение согласованности онтологий за счет расширения набора вариантов взаимного позиционирования двух концептов с учетом их семантической близости.

Экономическая полезность

Реализация данного проекта позволяет снизить затраты времени и труда на создание интегрированных информационных систем. Программная система интеграции данных ИС позволяет преодолеть семантическую разнородность информационных систем, то есть различие их схем данных (онтологий). В частности, процесс интеграции трех ИС в рамках ИАИС Бийского технологического института занял около часа, что на порядок быстрее, чем в случае использования для этих целей ручных (неавтоматизированных) методов. Существенно выигрывает данная система в смысле эффективности и по сравнению с другими программными средствами, обеспечивая более высокое качество интегрированных данных и, соответственно, сокращение временных затрат на выполнение данного процесса.

Техническое обоснование проекта

Программный комплекс разработан в среде Oracle на языке PL\SQL. Объем программного кода составляет порядка 7 тысяч строк. Исполняемые модули и массив настроек программы занимают порядка 10 мегабайт.

Интерфейс программной системы представлен на рисунке 2 – 4.

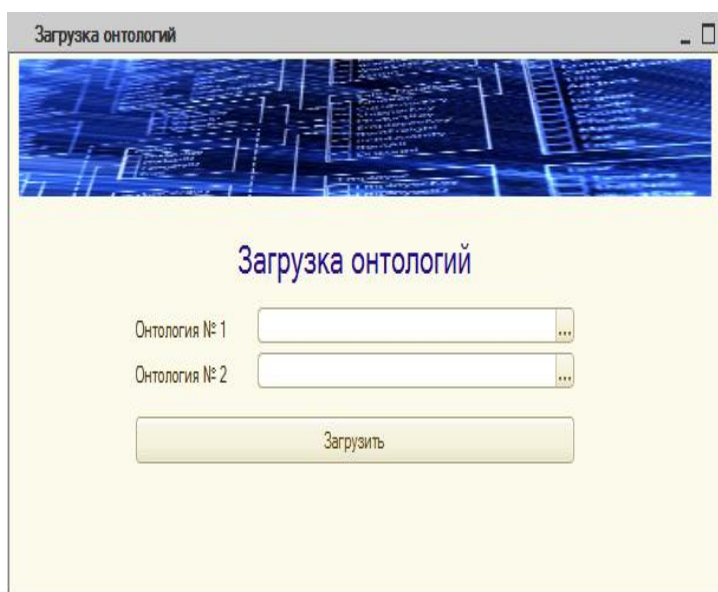


Рисунок 2 Главный экран программной системы

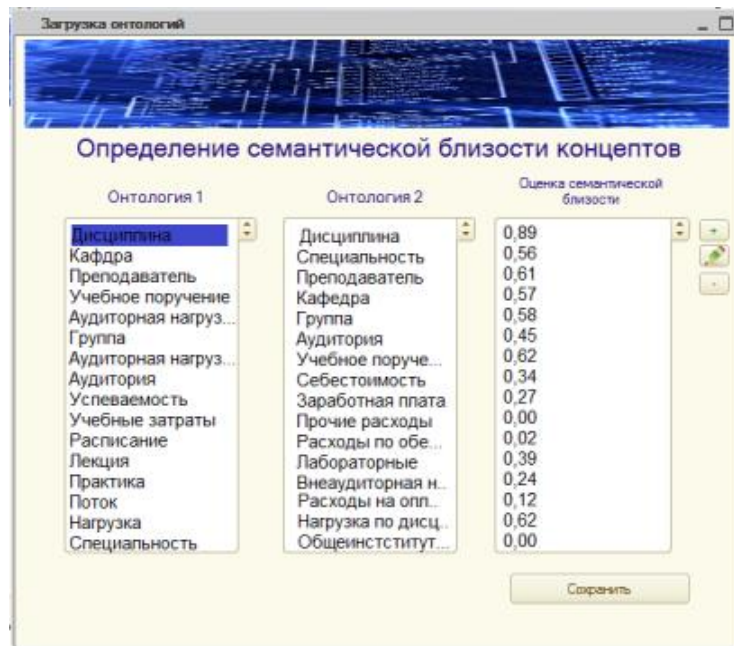


Рисунок 3 Определение семантической близости концептов

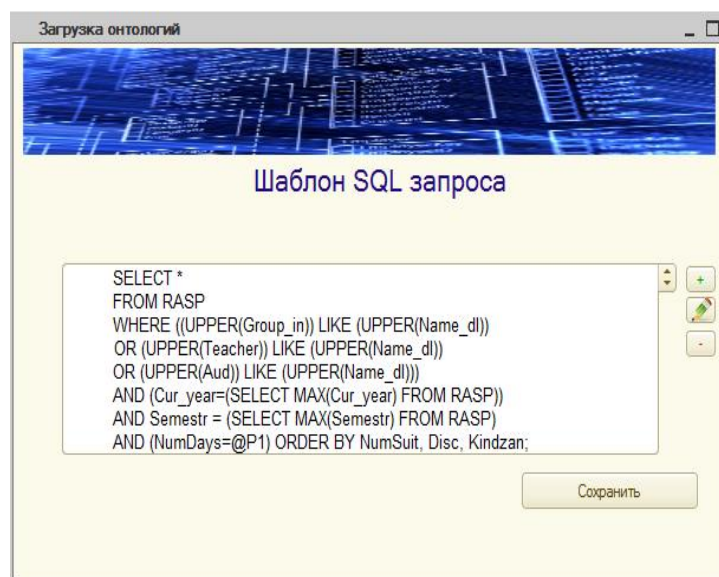


Рисунок 4 Формирование шаблона запроса

Заключение

Проект был апробирован и внедрен при создании ИАИС Бийского технологического института. Разработанная система может быть использована также при построении интегрированных информационных систем в других сферах деятельности.