

Проектирование информационной системы с использованием онтологии предметной области в рамках обучающего Web-портала

Сфера образования становится все более открытой, сложной и дифференцированной: появляются новые образовательные технологии, возникают новые формы образования, все чаще встает вопрос о новых способах сертификации полученного образования. Обучение с помощью электронных информационных технологий – e-learning – прочно закрепилось в мире как относительно новая, удобная и современная форма получения образования с помощью открытой интернет-среды, основанная на использовании информационных технологий, обеспечивающих обмен учебной информацией на расстоянии и реализующих систему сопровождения и администрирования учебного процесса.

Предлагаемая информационная система базируется на объединении методов искусственного интеллекта и Интернет-технологий. Основой для информационной системы, работающей со знаниями, служит онтология предметной области. Различные функции системы могут использовать онтологию для своих нужд. Например, используя онтологию, подсистема работы с документами может в автоматическом режиме или при помощи пользователя составить семантическое описание нового документа при его добавлении в репозиторий. Подсистема поиска, получив запрос пользователя и обратившись к онтологии, выделит в нем знакомые понятия, на основе их взаимоотношений определит текущий контекст и составит семантическое описание запроса. Далее этот семантический образ будет сопоставлен с описаниями информационных ресурсов системы, в результате чего подходящие под запрос ресурсы будут представлены пользователю. Подсистема документооборота может использовать семантические описания различных процессов создания и обработки документации (*workflow*) для того, чтобы на конкретном этапе сотруднику, выполняющему определенные функции, предоставлять важную для него информацию, найденную также при помощи онтологии. Каждый информационный ресурс должен получить описание, выражающее его семантику в явной форме. Семантика ресурса может быть выражена в понятиях онтологии и записана при помощи какого-либо языка представления знаний (DAML+OIL, OWL). В дальнейшем на основе этих описаний будет выполняться семантический поиск информации в системе.

На сегодняшний день наиболее зрелой методологией проектирования онтологий является методология On-To-Knowledge. Согласно методологии On-To-Knowledge процесс создания онтологии состоит из последовательных этапов *анализа целесообразности, начального проектирования, детальной разработки онтологии, тестирования и оценки модели, сопровождения и развития онтологии*. На основании анализа этих этапов можно выделить основные задачи построения онтологии:

- 1) определение целей онтологии и требований к ней;
- 2) анализ предметной области:
 - a. описание процессов, протекающих в компании;
 - b. описание должностных обязанностей сотрудников компании;
 - c. разработка сценариев использования системы на базе онтологии;
 - d. определение вопросов компетенции системы;
- 3) составление инструкций по разработке онтологии;
- 4) составление списка источников знаний;
- 5) выбор стартовой онтологии (базовой таксономии) из доступных;
- 6) сбор информации:
 - a. анкетирование и интервьюирование экспертов;
 - b. сбор документов, относящихся к предметной области;
- 7) обработка информации, выделение понятий и отношений:
 - a. лингвистический и статистический анализ документов;
 - b. анализ, проводимый инженерами знаний;
- 8) расширение стартовой онтологии выделенными понятиями и отношениями;
- 9) формальное описание онтологии в терминах языка представления знаний;
- 10) тестирование онтологии:
 - a. проверка полученной модели на соответствие спецификации;
 - b. тестирование макета онтологии представителем заказчика;
 - c. составление списка критических замечаний и предложений;
- 11) выполнение задач, необходимых для устранения выявленных недостатков;
- 12) сопровождение онтологии – поддержание модели в актуальном состоянии.

На основании информации об этапах разработки онтологии и выделенных задачах проектирования можно предложить следующую функциональную модель системы для построения онтологий (рисунок 1).

Далее рассмотрим описание информационной системы, которая обеспечивает работу с онтологией и ее использование в рамках обучающего WEB-портала. Информационная система проектируется в виде многоуровневого Web-приложения. Это связано с тем, что круг пользователей системы широк и они нуждаются в удаленном доступе к информационной системе.

Ядром системы является сервер онтологии (рисунок 2). Сервер онтологии содержит модули, связанные с созданием и использованием онтологии:

- модуль, обрабатывающий поисковые запросы к онтологии;
- модуль, связанный с построением объектной модели онтологии;
- модуль, реализующий механизм логического вывода;

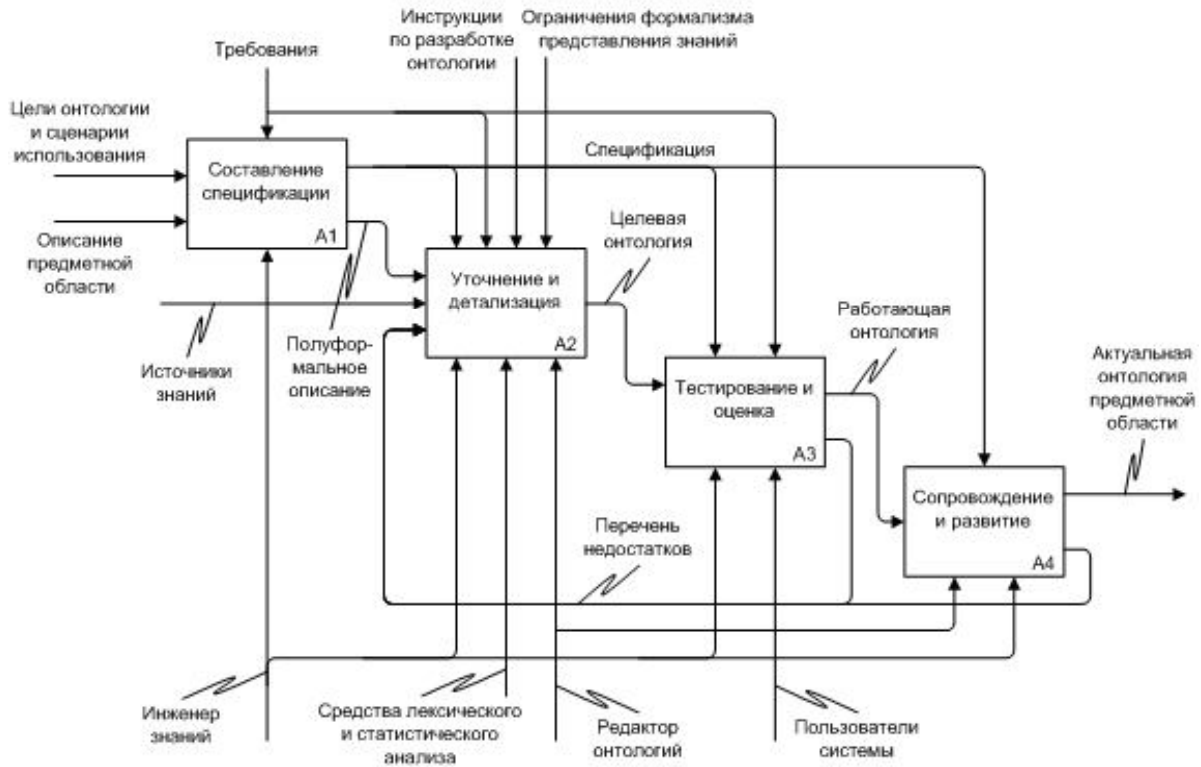


Рисунок 1 – Функциональная модель системы построения онтологии

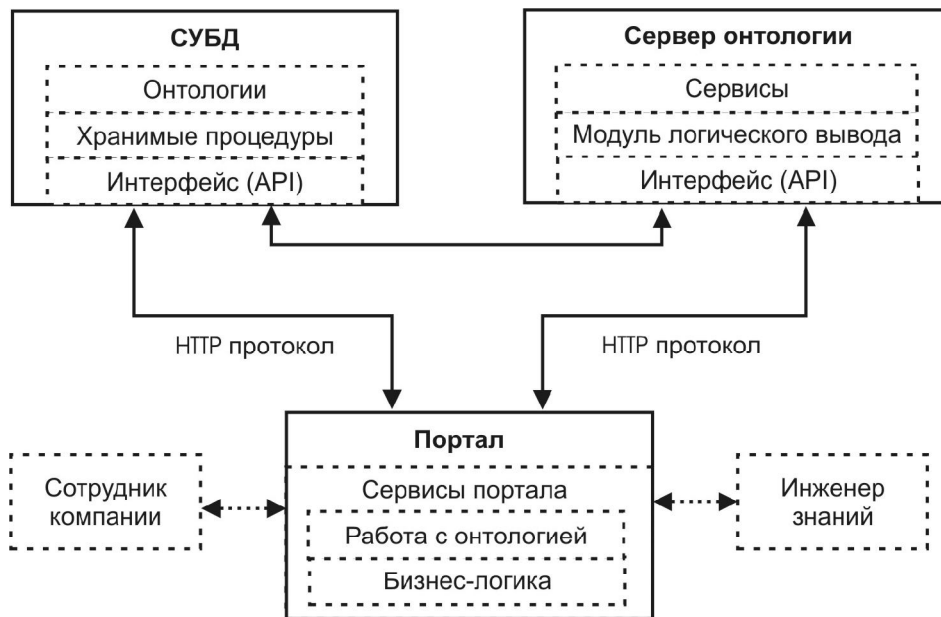


Рисунок 2 – Проект интеллектуальной подсистемы портала

- модуль, реализующий интерфейс доступа к хранилищу онтологий (СУБД).
- Эти модули будут использоваться сервисами портала, реализующими бизнес-логику, и сервисами, связанными с разработкой онтологии, к которым относятся:
 - сервис визуализации онтологии и навигации по ее структуре;
 - сервис, выполняющий аннотирование информационных ресурсов;
 - сервис поиска информационных ресурсов;
 - сервис лингвистико-статистического анализа тек-

ста, отвечающий за обработку текстов на естественном языке и выделение концепций онтологии;

- сервис редактора онтологии;
- другие сервисы.

С точки зрения создания онтологии важным элементом системы является хранилище знаний. Планируется хранить структуру онтологии в таблицах реляционной базы данных (РБД). Для этого должна быть разработана соответствующая схема РБД. Взаимодействие модулей системы с РБД будет осуществляться через специально написанный для этого интерфейс. Интерфейс должен реализовать такие функции, как:

- работа с онтологиями (создание, удаление, изменение служебной информации, связанной с конкретной онтологией);
- работа с понятиями (создание, удаление, модификация, выбор понятий по заданному критерию);
- работа с отношениями (добавление, удаление отношений между двумя понятиями, модификация существующих отношений, выбор всех понятий, связанных определенным отношением с заданным понятием);
- управление служебной информацией (авторы онтологий, пространства имен, типы данных, типы отношений и др. служебная информация системы).

Интерфейс не будет изменять непосредственно содержимое таблиц РБД. Функции интерфейса будут обращаться к хранимым процедурам РБД, что обеспечит определенный уровень независимости интерфейса от возможных изменений структуры базы данных.

В качестве редактора онтологии используется Protégé – расширяемая, платформенно-независимая среда для создания и редактирования онтологий и баз знаний. Protégé был разработан в Стэнфордском университете в лаборатории медицинской информатики. Этот редактор по многим критериям подходит для использования в распределенной системе на базе, хранимой в РБД онтологии.

Расширяемая функциональность. Protégé поддерживает разработку дополнительных модулей – плагинов, которые могут использовать объектную модель онтологии редактора для решения дополнительных задач. Выражаясь точнее, плагин представляет собой Java-класс, который использует методы базовых классов Protégé для расширения функциональных возможностей редактора. Protégé поддерживает три типа плагинов: плагины для хранения данных (*storage plug-ins*), плагины слотов (*slot widgets*) и плагины-вкладки (*tab plug-ins*). Плагин для хранения данных представляет собой модуль, который сохраняет или загружает модель знаний в/из определенного формата файла или базы данных. Плагин слота связан с редактированием значения определенного класса слотов (в Protégé ис-

пользуется терминология фреймовой модели представления знаний). Плагин-вкладка представляет собой панель редактора с набором управляющих элементов, совместно используемых для решения некоторой задачи, выходящей за рамки стандартных возможностей редактора.

Для редактирования онтологии будет использоваться OWL-плагин, позволяющий работать с формальным представлением онтологии на языке OWL. Язык OWL является наиболее перспективным языком представления знаний в сети WWW. Этот язык будет использоваться в системе в качестве формального представления онтологии вне хранилища и для взаимодействия ее компонентов.

Хранение онтологии в базе данных. Protégé позволяет использовать РБД в качестве хранилища онтологии. Стандартный способ записи онтологии в базу данных хранит всю модель знаний в одной таблице, что может быть не очень эффективно при обработке большого числа запросов. Благодаря расширяемой архитектуре редактора есть возможность создания плагина для хранения данных. Такой плагин будет записывать онтологию в базу данных с нужной нам схемой и затем считывать из нее онтологию для дальнейшего редактирования.

Удаленный многопользовательский доступ к онтологии. Protégé предусматривает возможность запуска на компьютере отдельного процесса-сервера и соединение с ним по протоколу HTTP удаленных клиентов. Сервер Protégé имеет список онтологий, доступных для удаленного редактирования и управляет правами доступа пользователей к редактированию этих онтологий. Управление правами доступа осуществляется редактированием служебной онтологии.

Таким образом, предложенная функциональная модель системы для построения онтологий обеспечивает эффективную работу с онтологией и ее использование в рамках обучающего WEB-портала для пользователей, нуждающихся в удаленном доступе к информационной системе.

СПИСОК ЛИТЕРАТУРЫ

1. W3C, «OWL Web Ontology Language Overview, W3C Recommendation 10 February 2004», <http://www.w3.org/TR/owl-features/>
2. W3C, «The Semantic Web», <http://www.w3.org/2001/sw>.
3. Studer R., Benjamins V.R. and Fensel D. Knowledge Engineering: Principles and methods, Data and Knowledge engineering (DKE), 25(1-2): 161-197, 1998.
4. Хорошевский В.Ф. Пространства знаний в сети Интернет и Semantic Web // Искусственный интеллект и принятие решений. 2008. № 1. С. 98-107.
5. Саданова Б.М., Яворский В.В. Преимущества объектно-ориентированного подхода при построении моделей образовательных ресурсов // Тр. ун-та; КарГТУ. Караганда, 2008. № 4. С. 41-44.