

К ВОПРОСУ РАЗРАБОТКИ ИНФОРМАЦИОННО-СПРАВОЧНОГО ОБЕСПЕЧЕНИЯ ИНСТРУМЕНТАЛЬНОГО КОМПЛЕКСА ОНТОЛОГИЧЕСКОГО НАЗНАЧЕНИЯ

Палагин А.В., Петренко Н.Г., Величко В.Ю., Тихонов Ю.Л.
Семенков В.В., Митрофанова А.Е.

Аннотация: В работе описана программная модель модуля информационно-справочного обеспечения, функционирующего как в составе инструментального комплекса автоматизированного построения формальных онтологий, так и в качестве отдельного приложения. В последнем случае (с точки зрения пользователя) модуль функционирует как информационно-справочная система для работы с энциклопедиями, толковыми словарями и тезаурусами в произвольных предметных областях. Формализация разработки программной модели выполнена по UML-технологии.

Ключевые слова: онтологический подход, инструментальный комплекс автоматизированного построения формальных онтологий, информационно-справочная система.

ACM Classification Keywords: D.2 SOFTWARE ENGINEERING: D.2.10 Design.

Введение

Интеллектуальные информационные технологии стали едва ли не главным определяющим показателем развития во всех сферах научно-технической деятельности человеческого общества. Центральным их звеном является технология инженерии знаний, которая реализует процессы управления знаниями, и успехи в этом направлении во многом определяют интеллектуальный уровень и общую эффективность компьютерных систем. К сожалению, по ряду причин многие идеи направления искусственного интеллекта сегодня так и остаются не реализованными как полностью, так и частично. Одной из причин является недостаточная эффективность большинства современных компьютерных систем: последние в процессе управления знаниями должны оперировать не примитивными данными (в традиционном понимании), а знаниями, представленными в соответствующей формальной теории. Разработка и использование новых информационных технологий, таких как GRID-вычисления, многоагентные системы и др., со своей стороны, также требует знание-ориентированного подхода. При этом следует отметить необходимость междисциплинарных научных исследований и соответствующего инструментария их компьютерной поддержки. Со своей стороны методология системной интеграции должна включать, в том числе, формальную теорию (или теории) представления информации, которая обрабатывается на разных стадиях, методологию автоматизированного построения баз знаний предметных областей, архитектуры компьютерных систем новых поколений и многое другое.

Сказанное выше в равной степени относится и к “электронному образованию” вообще, и к созданию электронных курсов (ЭК) предметных дисциплин (ПдД). Онтолого-ориентированная автоматизированная система разработки (общезначимых) ЭК по различным ПдД должна включать в себя СУБД, являющуюся библиотекой справочной информации (БСИ). Помимо того, что основная функция такой СУБД – поддержка процесса проектирования онтологии заданной предметной области в инструментальном комплексе онтологического назначения (ИКОН) [1], БСИ является одним из основных блоков при наполнении ЭК. Таким образом, задача разработки БСИ является актуальной.

Постановка задачи

На современном этапе перехода к стадии информационного общества происходит содержательное наполнение парадигмы “электронного образования”, направленное на улучшение качества и доступности образования, расширение его форм и методов, интеграцию в общеевропейские образовательные программы [2].

В настоящее время осуществляется широкомасштабный процесс информатизации образования, однако педагогическая эффективность ЭК остается низкой, так как разработка технического задания (ТЗ) спецификации и требований к ЭК не имеет четкой регламентации. Внедрение ЭК в образовательный процесс не носит системного характера, не опирается на четкий аксиоматический и семантический базис общей логической теории и междисциплинарных связей, не учитывает возможности развития ЭК [3]. Только с использованием базовой системы категорий можно дать основу для построения онтологической модели действительности на том уровне, который отражён современной наукой [4], т.е. построить адекватный ЭК для данной ПдД. Онтология как представление концептуальной системы с помощью логической теории и как словарь, используемый логической теорией, является естественной составляющей системы автоматизированного построения ЭК.

Одним из самых трудоемких процессов при создании электронных курсов является построение библиотеки справочной информации ЭК, содержащей концептуальные знания ПдД.

ИКОН, предназначенный для автоматизированного построения онтологических баз знаний предметных областей (ПдО), предоставляет возможность автоматизации этого процесса и делает возможным построение онтолого-ориентированной автоматизированной системы разработки (общезначимых) ЭК по различным ПдД. На рис.1 представлена обобщенная архитектура ИКОН, при этом хранилища информации и управление их функционированием показано детально.

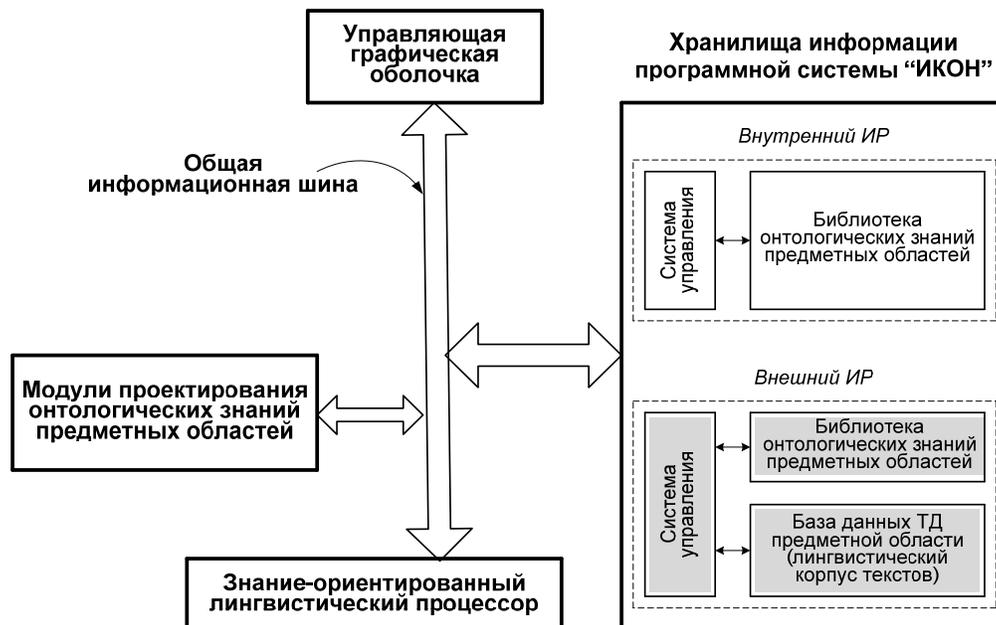


Рис. 1. Обобщенная архитектура ИКОН

ИКОН выполняет анализ и обработку больших объемов неструктурированных данных, в частности лингвистических корпусов текстов на украинском и/или русском языке, извлечение из них предметных

знаний с последующим их представлением в виде системно-онтологической структуры или онтологии предметной области.

Выделенный блок “*Внешний информационный ресурс*” создается на основе первоисточников. Создание “*Внешнего информационного ресурса*” - (электронные коллекции (ЭлК) энциклопедических словарей, толковых словарей и тезаурусов) предшествует процессу автоматизированного построения онтологии ПдО.

БСИ состоит из системы управления (СУ) и самой библиотеки энциклопедических знаний. СУ реализует функции, необходимые для обработки различных запросов пользователя и инженера по знаниям.

Модуль БСИ может функционировать в двух режимах: 1) как самостоятельное приложение; 2) в составе ИКОН. В первом режиме пользователь может самостоятельно наполнять систему новыми концептуальными знаниями, или добавлять к уже имеющимся в системе энциклопедии, толковые словари и тезаурусы по интересующим его ПдО. Во втором режиме основной функцией модуля БСИ является информационная поддержка работы модуля формирования множества понятий ИКОН.

Функциональная модель модуля БСИ

Онтология ПдО и, следовательно, онтология ПдД является открытой, общезначимой системно-онтологической структурой, доступной для адаптации к конкретному кругу задач. Как следствие, формальная онтология ПдД допускает многократное использование для различных наборов типовых задач (использование в учебных дисциплинах для разных специальностей, факультетов). Другими словами, один раз построенная формальная онтология ПдД пригодна для построения обширного набора ЭК. *Электронные коллекции энциклопедий, толковых словарей и тезаурусов* являются (по определению) общезначимыми ресурсами знаний в заданном домене прикладных областей.

Основные функции, выполняемые с помощью модуля БСИ, следующие:

- вывод описания терминов (понятий) на русском и/или украинском языке;
- связь с внешней БД;
- формирование связей между понятиями;
- поиск понятий;
- наполнения и организация словарей.

Для проектирования функциональной модели использована общепринятая стандартная методология и язык функционального моделирования UML (англ. Unified Modeling Language – унифицированный язык моделирования) – язык графического описания для объектного моделирования в области разработки программного обеспечения [6].

UML технологии стали основой для разработки и реализации во многих инструментальных средствах: в средствах визуального и имитационного моделирования, а также в CASE-средствах самого различного целевого назначения. Более того, заложенные в языке UML потенциальные возможности могут быть использованы не только для объектно-ориентированного моделирования систем, но и для представления знаний в интеллектуальных системах, которыми, по существу, являются перспективные сложные программно-технологические комплексы. Принимая во внимание все преимущества языка UML, на его основе была спроектирована функциональная модель модуля БСИ.

Функциональная модель ИНИТ представляет собой набор диаграмм трех видов:

- диаграмма вариантов использования (рис. 4);
- диаграмма активности (рис. 5);
- диаграмма классов (рис. 6).

Целями разработки диаграммы вариантов использования являются:

1. Определить общие границы и контекст моделируемой системы на начальных этапах проектирования модуля цифровой БСИ;
2. Сформулировать общие требования к функциональному поведению проектируемого модуля цифровой БСИ;
3. Разработать исходную концептуальную модель модуля цифровой БСИ для ее последующей детализации в форме логических и физических моделей.

Суть данной диаграммы состоит в следующем: проектируемая система представляется в виде множества сущностей или акторов (actor), взаимодействующих с системой с помощью так называемых вариантов использования. При этом актором или действующим лицом называется любая сущность, взаимодействующая с системой извне. Это может быть человек, программа или любая другая система, которая может служить источником воздействия на моделируемую систему. В свою очередь, вариант использования служит для описания сервисов, которые система предоставляет актору.

Для использования модуля БСИ в системе онтолого-ориентированной автоматизированной подготовки ЭК добавляется возможность экспорта в словарь ЭК.



Рис. 4 Диаграмма вариантов использования модуля БСИ

С помощью диаграммы активности (рис. 5) можно изучать поведение модуля цифровой БСИ с использованием моделей потока данных и потока управления. Диаграмма активности отображает некоторый алгоритм, описывающий жизненный цикл объекта, состояния которого могут меняться.

Диаграмма активности, в отличие от блок-схемы, имеет более широкую нотацию. Например, на ней можно указывать состояния объектов.

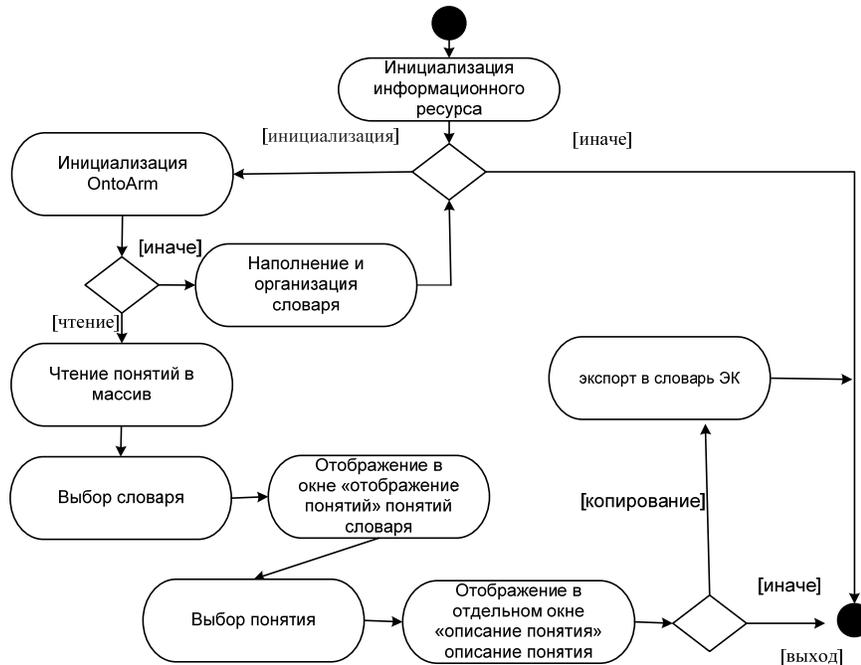


Рис. 5. UML-диаграмма активности программной модели модуля БСИ

Диаграмма классов описывает структуру объектов программной модели информационного ресурса: их индивидуальность, отношения с другими объектами, атрибуты, функции и процедуры. Модель классов создает контекст для диаграмм состояний и взаимодействия. На рис. 6 представлен фрагмент UML-диаграммы классов программной модели модуля БСИ (класс Dictionary - основной класс программы, реализует логику программы). Остальные классы: класс DictionaryItem, класс DictMap.

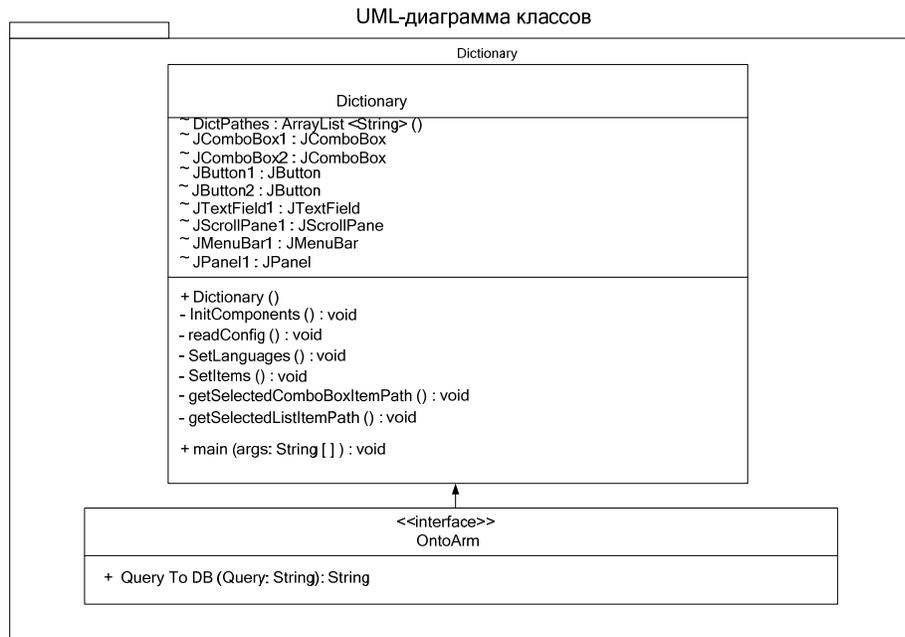


Рис. 6. Фрагмент UML-диаграммы классов программной модели модуля БСИ.

Описание работы программного модуля БСИ

- В начале работы модуля БСИ на экране появляется главное окно программы, в котором находится: строка меню (MenuBar) – содержит подменю Файл, Правка, Вид, Помощь, выпадающее меню «Выбор словаря», выпадающее меню снабжено всплывающей подсказкой «Выбор словаря» (ComboBox – строка выбора словаря), из которого можно выбрать необходимый тематический словарь, области для отображения списка ключевых терминов (List) – окно «общий список». Также есть кнопка Поиск и к ней строка набора искомого текста.
- После ввода термина в строку поиска и нажатия кнопки «Поиск» в окне «общий список» первой строкой будет искомое понятие.
- Для выбора тематического словаря, необходимо нажать на выпадающее меню «Выбор словаря», откроется весь список имеющихся словарей, после это необходимо выбрать нужный словарь – в результате в окне «Общий список» отобразится содержимое выбранного словаря в алфавитном порядке.
- Для получения описания термина нужно сделать двойной клик на нем, в результате откроется новое окно «Описание термина» в котором отобразится словарное описание термина на русском и украинском языках.
- Для экспорта понятия и его описания в файл словаря ЭК нужно нажать на кнопку экспорта, добавленную в окно модуля БСИ.

Для реализации хранения и обработки информации со словарей используется оригинальный метод организации доступа к информации и программное обеспечение - OntoArm [5].

На рис. 2 представлено главное окно программного модуля БСИ.

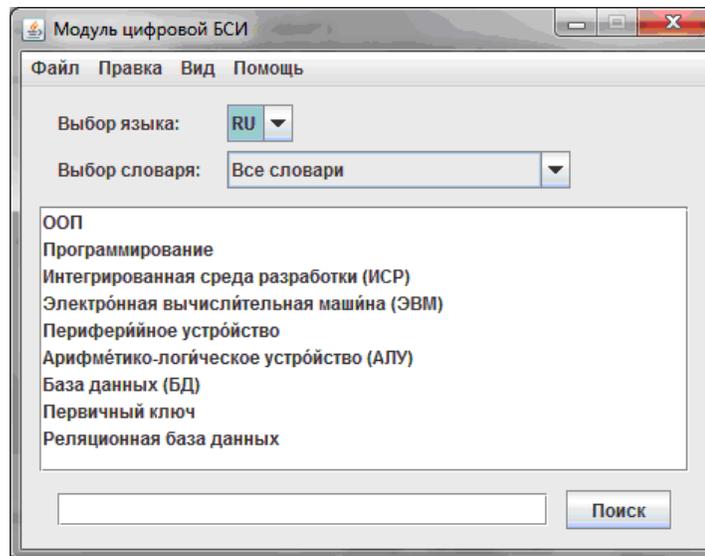


Рис.2. Главное окно модуля БСИ

Для системы автоматизированной подготовки ЭК в это окно добавляется кнопка «Экспорт».

Для выбора языка отображения имен словарей и содержащихся в них терминов используется выпадающий список «Выбор языка». Эта функция может быть использована при создании ЭК. Для выбора тематического словаря используется выпадающий список «Выбор словаря». Для вывода описания термина используется двойной щелчок кнопкой мыши по имени термина. Режим наполнения

выполнен в виде отдельной программы, которая позволяет вводить термины с клавиатуры или извлекать их из лингвистических корпусов оцифрованных текстов.

Выводы

В работе предложена программная модель цифровой библиотеки справочной информации и модуля управления библиотекой, входящих в оригинальный инструментальный комплекс онтологического назначения, дополненная средствами обеспечения функционирования БСИ как отдельного приложения. Определения и толкования библиотеки участвуют в формировании онтологии ПдД. Приведены UML-диаграммы, описывающие функциональную модель программного модуля БСИ ИКОН.

Электронные библиотеки являются общезначимыми ресурсами знаний в заданном домене близких прикладных областей и в этом смысле они инвариантны при адаптации и оптимизации онтологической системы к целевому применению, в частности, при адаптации к задаче автоматизации разработки ЭК по определенной предметной дисциплине.

Предлагаемая программная модель предназначена для поддержки процесса онтолого-ориентированного автоматизированного проектирования ЭК заданной предметной дисциплины в рамках архитектурно-структурной организации ИКОН.

Благодарности

Работа опубликована при финансовой поддержке проекта **ITHEA XXI** Института информационных теорий и приложений FOI ITHEA Болгария www.ithea.org и Ассоциации создателей и пользователей интеллектуальных систем ADUIS Украина www.aduis.com.ua.

Библиография

1. Палагин А.В. Онтологические методы и средства обработки предметных знаний / А.В. Палагин, С.Л. Крывый, Н.Г. Петренко. – [монография] – Луганск: изд-во ВНУ им. В. Даля, 2012. – 323 с.
2. Бухарестская декларация (общеевропейская конференция в Бухаресте по подготовке всемирной встречи на высшем уровне по вопросам информационного общества). – Доступно на www.itu.int/dms_pub/itu-s/md/03/wsispc2/doc/S03-WSISPC2-DOC-0005!!MSW-R.doc.
3. А.В. Палагин, Ю.Л. Тихонов, Н.Г. Петренко, В.Ю. Величко. Об онтологическом подходе в образовании //Сучасні тенденції розвитку інформаційних технологій в науці, освіті та економіці. (м. Луганськ, квітень 2011.)
4. Палагін О.В., Петренко М.Г. Модель категоріального рівня мовно-онтологічної картини світу // Математичні машини і системи. – 2006. – №3. – С. 91–104
5. К вопросу естественно-языковой адресации. Крассимира Иванова, Виталий Величко, Крассимир Марков // Problems of Computer Intellectualization. Kyiv, Ukraine - Sofia, Bulgaria, - 2012. (В печати)
6. С. Marshall, Enterprise Modelling with UML, ISBN 0-201- 43313-3, Addison-Wesley, Reading, MA, 2000.

Информация об авторах



Палагин Александр Васильевич – академик НАН Украины; заместитель директора Института кибернетики имени В.М. Глушкова НАН Украины; Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: palagin_a@ukr.net

Основные области научных исследований: системная интеграция трансдисциплинарных научных знаний, онтологический инжиниринг



Петренко Николай Григорьевич – к.т.н., старший научный сотрудник, Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: petrng@ukr.net

Основные области научных исследований: методология и инструментальные средства автоматизированного проектирования онтологий предметных областей, системная интеграция трансдисциплинарных научных знаний



Величко Виталий Юрьевич – к.т.н., старший научный сотрудник, Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: velychko@aduis.com.ua

Основные области научных исследований: индуктивный логический вывод, обработка естественно-языковых текстов.



Тихонов Юрий Леонтьевич – к.т.н., доц., доцент кафедры ИТС, Луганского национального университета им. Тараса Шевченко; e-mail: t2003i@mail.ru

Основные области научных исследований: разработка онтологизированных электронных курсов

Семенков Виталий Владимирович – ассистент кафедры ИТС, Луганского национального университета им. Тараса Шевченко; e-mail: semenkov@mail.ru

Основные области научных исследований: методы и алгоритмы построения СУБД справочного характера

Митрофанова Анна Евгеньевна – Ин-т кибернетики им. В.М. Глушкова НАН Украины, Киев-187 ГСП, 03680, просп. акад. Глушкова, 40; e-mail: amitrofanova@inbox.ru

Основные области научных исследований: методы и алгоритмы построения толковых словарей онтологического характера