

УДК 004.89

## **АВТОМАТИЗИРОВАННАЯ РАЗРАБОТКА ОНТОЛОГИЧЕСКОЙ МОДЕЛИ ПРЕДМЕТНОЙ ОБЛАСТИ ДЛЯ ПОИСКА ОБРАЗОВАТЕЛЬНЫХ РЕСУРСОВ С ИСПОЛЬЗОВАНИЕМ АНАЛИЗА ТЕКСТОВ РАБОЧИХ ПРОГРАММ**

**Ужва А. Ю.**

*ГОУ ВПО Волгоградский государственный технический университет, 400131, Россия, Волгоград, пр. Ленина, 28, e-mail: alexey@uzhva.ru*

---

**Количество ежегодно публикуемых образовательных ресурсов увеличивается стремительными темпами. В связи с этим актуальной задачей становится автоматизация поиска образовательных ресурсов с учетом индивидуальных особенностей и текущих уровней знаний обучаемых. Анализ современного состояния вопроса показал, что наиболее перспективным методом адаптивного поиска является сочетание рассуждений на онтологии и рассуждений по прецедентам. В статье рассматривается процесс формирования онтологической модели, обеспечивающей поддержку алгоритма рассуждения по прецедентам для поиска электронных образовательных ресурсов, при помощи анализа текстов рабочих программ. Для этого разработана модель образовательного стандарта, рабочей программы. Представлены структура разработанной онтологической модели предметной области, архитектура автоматизированной системы, результаты тестирования предложенного метода, показавшие снижение трудоемкости формирования данной онтологии.**

---

**Ключевые слова:** онтологическая модель представления знаний, рассуждения по прецедентам, образовательные ресурсы, e-learning.

## **AUTOMATIC DEVELOPMENT OF ONTOLOGY MODEL FOR CASE-BASED REASONING IN SEARCH OF EDUCATIONAL RESOURCES USING ANALYZYS OF EDUCATION PROGRAMMS**

**Uzhva A. Y.**

*Volgograd State Technical University, 400131, Russia, Volgograd, Lenina av., 28, e-mail: alexey@uzhva.ru*

---

**Number of yearly published educational resources is rapidly growing. Therefore automatization of adaptive educational resources search become actual nowadays. Adaptive search should be performed with respect to learner profile and his current knowledge state. Analysis of existing models and methods shown that the most relevant method is combination of case-based reasoning with ontology based reasoning. Process of automatic ontology model creation considered. Resulting ontology can be used in adaptive search algorithm. Article includes description for structure of ontology model with support of case-based reasoning, architecture of software used for automatic model creation. The described model supports adaptive search algorithm with respect to learner profile.**

---

**Keywords:** ontology, case-based reasoning, educational resources, e-learning.

**Введение.** Количество ежегодно публикуемых образовательных ресурсов увеличивается стремительными темпами. Так, в 2011 году издано более 4 млн образовательных ресурсов, в то время как в 2002-м году их было издано менее 250 тысяч. Более 4,5 млн студентов в 2012 году проходило хотя бы 1 образовательный курс онлайн, по прогнозам в 2014 году данное количество достигнет 18 миллионов. Показано, что электронное образование способно повысить эффективность обучения, понизив при этом стоимость образовательного процесса. В связи с этим актуальной задачей становится автоматизация поиска образовательных ресурсов с учетом индивидуальных особенностей и текущих уровней знаний обучаемых.

Анализ подходов на основе нейронных сетей, решающих деревьев, онтологии, рассуждений по прецедентам, а также других подходов, показал, что наиболее перспективным является подход, сочетающий в себе рассуждения на онтологии и рассуждения по прецедентам.

Для реализации данного подхода должна быть создана онтологическая модель представления знаний. Неавтоматизированное создание данной модели является трудоемким процессом. Для устранения данного недостатка целесообразно автоматизировать первичное наполнение онтологии предметной области, используя имеющиеся неструктурированные данные. В качестве источника знаний о предметной области целесообразно использовать электронные версии рабочих программ, входящих в комплекс учебно-методических материалов для дисциплины. Для того чтобы было возможно осуществить автоматизированное построение онтологической модели на основе данных документов, структура модели должна соответствовать структуре используемых документов.

Помимо этого онтологическая модель представления знаний должна поддерживать интеграцию формализованных знаний в цикл рассуждений по прецедентам.

**Требования к составу и структуре онтологической модели предметной области.** Для реализации поставленной задачи было принято решение построить онтологическую модель предметной области, включающей в себя следующие структурные компоненты:

1. Модель образовательного стандарта, содержащая знания об образовательных дисциплинах, формируемых ими компетенциях, а также входящих в рабочую программу дисциплины изучаемых навыков (субкомпетенций).
2. Модель поля знаний, содержащая знания о зависимостях компетенций и субкомпетенций друг от друга.
3. Модель обучаемого, содержащая знания об обучаемом, включая его персональные характеристики и предпочтения, а также модель текущего поля знаний студента.
4. Модель образовательного ресурса, содержащая знания об образовательных ресурсах, включая их библиографические данные, набор охватываемых ресурсом тем, объединения ресурсов в образовательные коллекции.
5. Модель прецедента, обеспечивающую поддержку выбранных методов и алгоритмов построения персонифицированных образовательных коллекций.

Данные модели должны управляться мета-онтологической моделью при помощи связей включения, отношения, зависимости. Мета-онтологическая модель должна обеспечивать поддержку поиска образовательных ресурсов в соответствии с целями обучения, персональными предпочтениями и характеристиками пользователя, с использованием алгоритма рассуждений по прецедентам.

**Разработка модели.** Разработанная мета-онтологическая модель предметной области  $M$  состоит из следующих компонент:

$$M = \langle MO, I, RL, RR \rangle,$$

где  $M$  – мета-онтологическая модель предметной области;

$MO$  – множество управляемых онтологических моделей;

$I = \emptyset$  – множество экземпляров мета-онтологии;

$RL = \emptyset$  – множество правил вывода мета-онтологии;

$RR$  – множество отношений мета-онтологии.

$$MO = \langle OC, OL, OK, OColl, OStd \rangle,$$

где  $OC$  – онтологическая модель прецедента;

$OL$  – онтологическая модель обучаемого;

$OK$  – онтологическая модель поля знаний;

$OColl$  – онтологическая модель образовательной коллекции;

$OStd$  – онтологическая модель образовательного стандарта.

$$RL = \{includes, teaches, forms, has\},$$

где *includes* – отношение “включает”;

*teaches* – отношение “обучает”;

*forms* – отношение “формирует”;

*has* – отношение “имеет”.

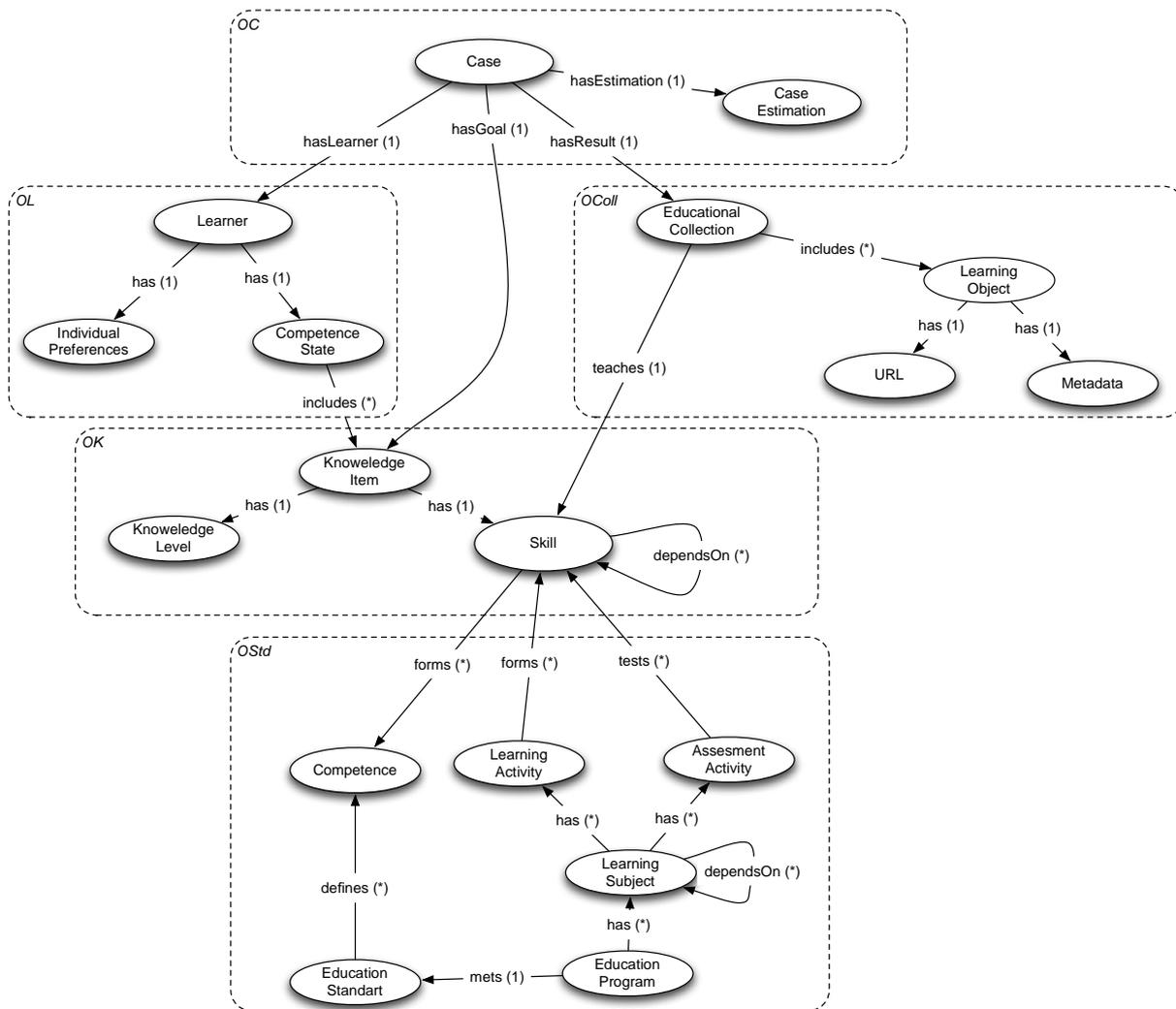


Рисунок 1. Графическое представление модели предметной области

На рисунке Рисунок 1 представлены графически основные концепты и отношения онтологии предметной области.

Онтологическая модель прецедента *OC* обеспечивает поддержку интеграции онтологических моделей представления знаний и алгоритма рассуждений по прецедентам.

Онтологическая модель образовательной коллекции *OColl* обеспечивает поддержку аннотирования электронных образовательных ресурсов, объединения их в образовательные коллекции для изучения концептов предметной области и задания соответствия образовательного ресурса и концепта предметной области.

Онтологическая модель обучаемого *OL* обеспечивает поддержку описания свойств и характеристик обучаемого, таких как его идентификационные данные (ФИО, группа, курс), текущее поле знаний, персональные характеристики и предпочтения.

Онтологическая модель образовательного стандарта  $OStd$  включает в себя описание образовательных курсов и их элементов, а также взаимосвязей между ними, определяющих принадлежность образовательных единиц к образовательному курсу, а также зависимость между образовательными единицами и междисциплинарные связи. Данная онтологическая модель отражает компетентностную модель образования, введенную в 3-м поколении федерального государственного образовательного стандарта.

$$OStd = \langle OC_{std}, IOStd, RLOStd, RROStd \rangle,$$

где  $OC_{std}$  – множество концептов онтологии образовательных курсов;

$IOStd$  – множество экземпляров;

$RLOStd$  – множество отношений;

$RROStd = \emptyset$  – множество правил вывода.

Структурой онтологии образовательного стандарта определены следующие концепты:

$$OC_{std} = \{Competence, Learning Activity, Assesment Activity, Learning Subject, Education Program, Education Standart\},$$

где *Competence* – компетенция, определяемая образовательным стандартом;

*Learning Activity* – образовательная деятельность, приводящая к освоению компетенции;

*Assesment Activity* – методика проверка результата (например, тест, экзамен, зачет);

*Learning Subject* – образовательная дисциплина;

*Education Program* – образовательная программа;

*Education Standart* – образовательный стандарт.

Онтологическая модель поля знаний  $OK$  обеспечивает поддержку описания элементов поля знаний предметной области. Ранжированное поле знаний обучаемого может быть представлено в виде множества:

$$RKF = \langle KI_1, KI_2, \dots, KI_n \rangle,$$

где  $RKF$  — ранжированное поле знаний обучаемого;

$KI_i = \langle C_i, L_i \rangle$  – элемент поля знания;

$C_i \in C$  – компетенция обучаемого;

где  $L_i$  – уровень владения компетенцией  $C_i \in C$ .

Где  $C$  – множество компетенций предметной области.

Экземпляры онтологических моделей  $IOStd$  и  $IOK$  изначально представляют собой пустые множества. Задача наполнения онтологии представляет собой заполнение данных множеств при помощи анализа текстов рабочих программ.

## **Автоматизированное формирование онтологической модели предметной области**

Рабочая программа учебной дисциплины содержит в текстовом формате знания о:

- названии дисциплины, семестре изучения;
- цели и задачах изучения дисциплины;
- формируемых компетенциях согласно образовательному стандарту;
- дисциплинах, на изучении которых базируется данный курс;
- списке субкомпетенций, формируемых данной дисциплиной;
- объеме (в академических часах) формирования субкомпетенции;
- отношениях компетенций и субкомпетенций.

Согласно представленной выше модели предметной области, могут быть извлечены и сформированы экземпляры для концептов *Competence, Learning Activity, Assesment Activity, Learning Subject, Education Program, Education Standart*.

На рисунке Рисунок 2 представлена архитектура системы извлечения знаний рабочих программ, позволяющая выполнить данную задачу.

Система состоит из двух основных модулей – модуля анализа рабочих программ и модуля управления зависимостями компетенций.

Модуль анализа рабочих программ состоит из двух подсистем – извлечения текстовых данных и анализа текстовых данных. Назначением данного модуля является получение максимального количества знаний из файлов рабочих программ. Данный модуль работает автоматически без управления пользователем.

Подсистема извлечения текстовых знаний оперирует с файлами рабочих программ в формате Microsoft Word. Данный выбор обусловлен фактическим способом хранения электронных учебно-методических комплексов дисциплин в ВУЗе. Выходными данными подсистемы является массив структурированных данных в формате JSON. Формат JSON выбран как универсальное средство обмена данными между гетерогенными инструментами и подсистемами.

Подсистема анализа текстовых данных оперирует со структурированными данными в формате JSON, полученными от подсистемы извлечения текстовых знаний и онтологической базой знаний. Данная подсистема формализует принимаемые на вход данные, используя структуру онтологии образовательного стандарта, и формирует знания о структуре предметной области в виде концептов и отношений онтологий образовательного стандарта *OS* и поля знаний *OK*.

Модуль управления зависимостями компетенций состоит из подсистемы пользовательского интерфейса и подсистемы установления зависимостей. Назначением данного модуля является обеспечение возможности доопределения зависимостей субкомпетенций друг от друга при помощи данных, получаемых от пользователя. Выходными данными данной системы является набор отношений *dependsOn*, формализованных при помощи структуры онтологии поля знаний *OK*.

Подсистема пользовательского интерфейса совместно с подсистемой установления зависимостей оперируют вводимыми пользователем данными для установления зависимостей субкомпетенций, которые не могут быть установлены автоматически в виду ограниченного количества данных, содержащихся в рабочих программах.

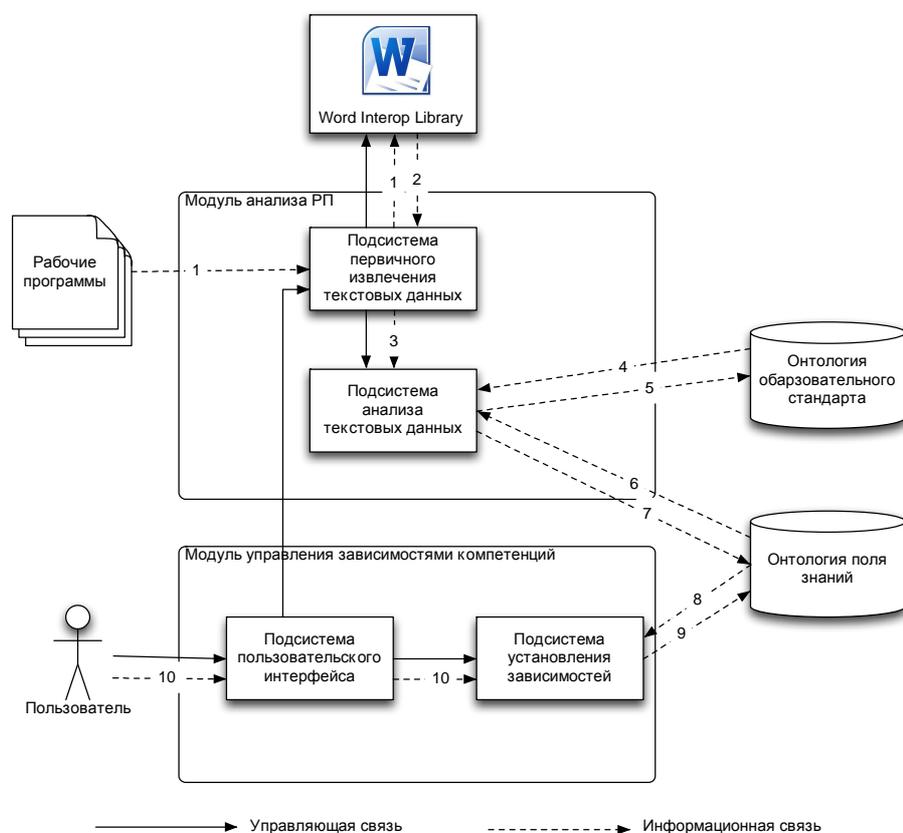


Рисунок 2. Архитектура системы формирования онтологии предметной области  
 Цифрами на схеме обозначены: 1 – файлы рабочих программ; 2 – структурированные данные рабочих программ; 3 – извлеченные данные в формате JSON; 4 – знания о структуре онтологии образовательного стандарта; 5 – знания о концептах образовательного стандарта; 6 – знания о структуре поля знаний; 7 – элементы поля знаний; 8 – знания о структуре и элементах поля знаний; 9 – знания о зависимостях элементов полей знаний; 10 – выбранные пользователями зависимые концепты поля знаний.

**Выводы.** Разработанный подход к автоматизированному формированию онтологической модели представления знаний о предметной области «электронные образовательные ресурсы» позволяет снизить трудоемкость формирования модели.

На данной онтологической модели формализованы знания о 26 предметах образовательной программы направления 231000 «Программная инженерия», в рамках которых автоматизированно заполнено 216 субкомпетенций. Детально проработаны первая часть части курса «Алгоритмические языки программирования» и вводной части курса «Интеллектуальные системы». На формирование онтологической модели было затрачено на 60 % меньше времени по сравнению с ручным заполнением.

### Список литературы

1. Интеграция онтологических моделей представления знаний и рассуждений по прецедентам для генерации персонифицированных образовательных коллекций / Дворянкин А. М., Ужва А. Ю., Жукова И. Г., Кульцова М. Б. // Инновационные технологии в обучении и производстве: сб. науч. тр. всеросс. конф. – Камышин, 2009. – С. 59.
2. Ужва А. Ю. Модели и методы поиска образовательных ресурсов на основе рассуждений по прецедентам и онтологии / А. Ю. Ужва, М. Б. Кульцова // XV региональная конференция молодых исследователей Волгоградской области (Волгоград, 9–12 ноября 2010 г.): тез. докл. / ВолгГТУ [и др.]. – Волгоград, 2011. – С. 189-191.
3. Ужва А. Ю. Онтологическая модель предметной области, обеспечивающая поддержку рассуждений по прецедентам для поиска образовательных ресурсов / А. Ю. Ужва, И. Г. Жукова, М. Б. Кульцова // Изв. ВолгГТУ. Серия "Актуальные проблемы управления, вычислительной техники и информатики в технических системах". Вып. 9: межвуз. сб. науч. ст. / ВолгГТУ. – Волгоград, 2010. – № 11. – С. 110-113.
4. Aamodt A. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches / A. Aamodt, E. Plaza // Communications. – 1994. – № 7. – С. 39-59.
5. Gaeta M. Ontology Extraction for Knowledge Reuse: The e-Learning Perspective / M. Gaeta // IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans. – 2011. – Т. 41. – № 4. – С. 798–809.
6. Heckmann, Gumo—the general user model ontology / D.Heckmann, T.Schwartz, B.Brandherm // User modeling. – 2005. – С. 149–149.
7. Henze N. Personalization services for e-learning in the semantic web / N. Henze // Design of Adaptive Feedback in a Web Educational System. – 2005. – С. 55.

8. Sheeba T. Semantic Web to E-Learning Content / T. Sheeba, S. Begum, M. Bernard // International Journal. – 2012. – Т. 2. – № 10. – С. 58–66.
9. Sicilia M.-Á. Modeling instructional-design theories with ontologies: Using methods to check, generate and search learning designs / M.-Á. Sicilia // Computers in Human Behavior. – 2011. – Т. 27. – № 4. – С. 1389–1398.
10. Zhou L. Ontology learning: state of the art and open issues / L. Zhou // Information Technology and Management. – 2007. – Т. 8, № 3. – С. 241 – 252.

**Рецензенты:**

Лукьянов Виктор Сергеевич, доктор технических наук, профессор, кафедра ЭВМиС, ВолгГТУ, г. Волгоград.

Камаев Валерий Анатольевич, доктор технических наук, профессор, зав. кафедрой САПР и ПК, ВолгГТУ, г. Волгоград.