

ПРИНЦИПЫ И МЕТОДЫ СЕМАНТИЧЕСКОГО СТРУКТУРИРОВАНИЯ ИНФОРМАЦИОННО-ОБРАЗОВАТЕЛЬНОГО ПРОСТРАНСТВА НА ОСНОВЕ РЕАЛИЗАЦИИ ОНТОЛОГИЧЕСКОГО ПОДХОДА

УДК 004.423.4:

Юрий Филиппович Тельнов,
д.э.н., профессор, заведующий кафедрой
Прикладной информатики в экономике,
Московский государственный универси-
тет экономики, статистики и информатики
(МЭСИ)
Тел. (495) 442-80-98
Эл. почта: YTelnov@mes.i.ru

В статье раскрываются принципы семантического структурирования информационно-образовательного пространства объектов знаний и научно-образовательных сервисов с использованием методов онтологического инжиниринга. Новизна предлагаемого подхода заключается в сопряжении онтологии контента и онтологии научно-образовательных сервисов, что позволяет осуществлять эффективную композицию сервисов и объектов знаний в соответствии с моделями профессиональных компетенций и потребностей обучающихся. В результате применения методов семантического структурирования информационно-образовательного пространства обеспечивается интеграция использования разнородного распределенного научно-образовательного контента учебными заведениями для проведения научных исследований, методических разработок и обучения.

Ключевые слова: информационно-образовательное пространство, научно-образовательный контент, объекты, сервисы, репозиторий, онтология.

THE PRINCIPLES AND METHODS OF INFORMATION AND EDUCATIONAL SPACE SEMANTIC STRUCTURING BASED ON ONTOLOGIC APPROACH REALIZATION

Yurij F. Telnov,
Doctorate of Economics, Professor, the
Head of the Department of Applied Informatics
in Economics, Moscow State University
of Economics, Statistics and Informatics
(MESI)
Tel. (495) 442-80-98
E-mail: YTelnov@mes.i.ru

This article reveals principles of semantic structuring of information and educational space of objects of knowledge and scientific and educational services with use of methods of ontologic engineering. Novelty of offered approach is interface of ontology of a content and ontology of scientific and educational services that allows to carry out effective composition of services and objects of knowledge according to models of professional competences and requirements being trained. As a result of application of methods of information and educational space semantic structuring integration of use of the diverse distributed scientific and educational content by educational institutions for carrying out scientific researches, methodical development and training is provided.

Keywords: information and educational space, scientific and educational content, objects, services, repository, ontology.

1. Принципы семантического структурирования информационно-образовательного пространства

Развитие интернет-технологий привело к созданию и применению сервисно-ориентированной архитектуры (СОА) программных систем, основанной на реализации прикладных функций в виде программных сервисов, к которым осуществляется доступ в интернет и/или интранет среде независимо от местоположения в глобальной вычислительной сети [1,2]. Применительно к проблеме создания единого информационно-образовательного пространства сервисная архитектура обеспечивает распределенный гибкий доступ к разнородным источникам знаний преподавателей, научных сотрудников, аспирантов и студентов сетевого сообщества учебных заведений для решения разнообразных научно-образовательных задач [3,4,5].

В настоящее время сервисная организация информационно-вычислительного обслуживания базируется на большом числе независимых, несвязанных и дублирующихся сервисов, на основе которых формируется структура динамических информационных процессов. Использование сервисно-ориентированной архитектуры обеспечивает организациям быструю адаптацию программного обеспечения к изменениям условий и требований рынка, полное соответствие информационных технологий конкретному виду деятельности, многократное повторное использование сервисов. При этом встает проблема выбора и адаптации сервисов для решения конкретных задач. Для таких процессов необходимы более гибкие механизмы, позволяющие проектировать и/или перепроектировать информационные процессы «на лету», т.е. непосредственно во время выполнения процесса.

Существующие сервисно-ориентированные архитектуры не учитывают зависимость выполняемых сервисами функций от особенностей модели проблемной области и решаемых задач, не решены задачи оптимального выбора сервисов из множества доступных и их динамической адаптации к конкретным потребностям решаемых задач, что может привести к принятию неэффективных решений по организации информационных процессов. Отсутствие теоретических основ для решения данной задачи обуславливает необходимость проведения исследования в области семантического описания сервисов и источников знаний и их динамического связывания на основе концептуального моделирования проблемной области и отражения концептуальных моделей в рамках онтологий.

Создание информационно-образовательного пространства на основе семантического представления источников знаний и сервисов как раз и направлено на повышение актуальности и эффективности коллективного использования в обучении и научных исследованиях распределенного в вычислительной сети контента, имеющего различия в стандартах, платформах, структуре и форме представления.

Целью данной работы является разработка методов и средств создания информационно-образовательного пространства на основе применения онтологического подхода, которые обеспечивают распределенное хранение и совместное использование с помощью сервисов разнородного научно-образовательного контента всеми участниками научно-образовательного процесса для проведения научных исследований, методических разработок и обучения.

В качестве основных принципов создания единого информационно-образовательного пространства с использованием сервисов следует выделить следующие:

- объединение различных источников информации по различным дисциплинам, направлениям подготовки и участникам научно-образовательного процесса в рамках единого ИОП;
- обеспечение постоянного развития системы за счет обновления теоретического знания и непрерывного накопления нового опыта;

- предоставление релевантной информации по конкретным образовательным программам каждому из участников научно-образовательного процесса в соответствии с требованиями реализации профессиональных компетенций;

- обеспечение распределенного доступа к контенту участникам научно-образовательного процесса, территориально распределенных в образовательной сети.

- организация коллективной среды взаимодействия участников образовательного процесса.

Интеграция разнородных источников знаний, междисциплинарный характер их использования, необходимость привлечения внешних источников знаний, обмен знаниями между пользователями в образовательном процессе предполагает проведение разработки архитектуры единого информационно-образовательного пространства, которое можно представить на трех взаимодействующих уровнях [6,7]:

Синтаксический (объектный) уровень – хранение аннотированных с помощью специально разработанной системы категорий (онтологий) источников знаний и сервисов в виде объектов репозитория и их индексирование.

Семантический (понятийный) уровень – определение концептуальной модели структуры знаний, общей для всех источников знаний и сервисов на основе системы онтологий.

Прагматический уровень (уровень приложений) – определение механизмов отбора, адаптации и исполнения сервисов для решения разнообразных научно-образовательных задач.

Рассмотрим перечисленные уровни организации информационно-образовательного пространства более детально.

2. Методы структурирования знаний на объектном уровне

Распределенный характер внешних и внутренних источников знаний, а также научно-образовательных сервисов предопределяет необходимость централизации управления этим контентом, которая осуществляется в виде цифровых репозитория объектов. При этом хранимые объекты в репозитории, как правило, содержат ссылки на реальные физические источники, которые находятся в распределенной сети Интернет/Интранет. Объекты

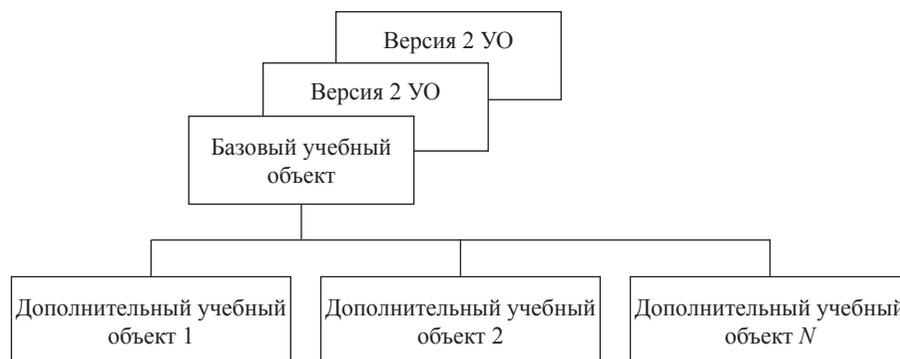


Рис. 1. Структура научно-образовательного объекта

могут иметь сложную иерархическую структуру агрегативного типа (отражение отношений «целое – часть») и версию природу хранения с учетом различных дат и авторов (рисунок 1).

Обычно учебно-научные объекты создаются из следующих источников знаний:

- Учебно-методические материалы;
- Научные проекты научно-педагогических работников, аспирантов и студентов;
- Лучшая практика и референсные модели, обобщающие опыт деятельности в различных сферах.
- База знаний экспертов, которые могут быть привлечены к учебно-методической деятельности и выполнению научно-исследовательских проектов.
- Нормативная и справочная информация.
- Базы знаний новостей и публикаций и др.

Наибольший интерес среди перечисленных источников знаний представляют базы знаний лучшей практики, поскольку они отражают опыт решения тех или иных научно-образовательных задач, который может непосредственно использоваться на практических занятиях путем адаптации к конкретным условиям рассматриваемой предметной области.

Как правило, базы знаний лучшей практики отражают:

- Определение методов решения задач.
- Метрики (показатели) эффективности применения методов.
- Описание успешных и неуспешных прецедентов внедрения методов решения задач.
- Справочные материалы: подборки презентаций, учебных курсов, статей.

В качестве основы структуризации источников знаний информационно-образовательного пространства для сервисов обычно используется

модель цепочек ценности (создания стоимостных объектов) М. Портера, в соответствии с которой происходит выделение основных и вспомогательных (поддерживающих) процессов.

3. Методы структурирования знаний на семантическом уровне

Концептуальный уровень ИОП обеспечивает семантическую интерпретацию запросов к научно-образовательному контенту, обеспечивает реализацию унифицированного интеллектуального доступа к множеству источников знаний. В основе концептуального уровня структуры знаний лежит таксономия используемых понятий или онтология, предназначенная для идентификации различных компонентов знания. Онтологию можно рассматривать, как систему рубрикации предметной области, с помощью которой интегрируются разнородные источники знаний [8,9]. С другой стороны, онтология часто рассматривается, как словарь-тезаурус, совместно используемый в СУЗ для упрощения коммуникации пользователей, формулирования и интерпретации их запросов. Классическим определением онтологии является определение Т.Грубера [10], трактующее онтологию как «явное описание концептуализации знаний», подразумевающее доступность концептуальной модели знаний пользователям или их компьютерным агентам в процессе взаимодействия с компьютерной системой управления знаниями.

Сложность отображения моделей проблемной области, представленных в различных методологиях, обусловлена сильной привязкой в существующих подходах содержания модели к формализму. В разрабатываемых в последнее время системах управления знаниями акцент делается как раз не на форму, а на сущность отображаемых

явлений или концептуализацию знаний о проблемной области в онтологиях. В связи с этим разработка подхода к моделированию проблемной области на основе онтологий, в которой стандартизуется метамодель мира, то есть такие понятия, как объект, субъект, функция, ресурс и их взаимодействия, представляется основой для решения поставленной задачи интеграции применения различных стандартов в моделировании проблемной области.

В результате создания онтологии достигаются следующие ключевые преимущества применения ИОП:

- Точный и эффективный доступ к источникам знаниям, релевантным контексту задачи (конкретной ситуации).
- Лучшее понимание и интерпретацию пользователем полученных знаний в данном контексте с возможностью дополнительных справочных обращений к корпоративной памяти.
- Обеспечение информационного моделирования, то есть итерационное уточнение информационных потребностей в процессе решения задачи.

Структура знаний, представляемая в онтологии, должна отображать постоянные парадигматические отношения понятий, независимые от контекста решения задачи, и правила формирования переменных синтагматических отношений понятий, возникающих в некотором контексте решения задачи.

Парадигматические отношения понятий. К таким отношениям относятся отношения синонимии, омонимии, полисемии, обобщения («Род-Вид»), агрегации («Целое-Часть»), ассоциации «Причина – Следствие», «Средство – Цель», «Аргумент – Функция» и ролевые ассоциации «Роль – Процесс» и т.д., которые превращают словарь в тезаурус.

Для понятий онтологии могут быть также заданы свойства (атрибуты), формулы вычислений и примеры экземпляров.

Онтология отражает *синтагматические отношения понятий* в виде семантических ограничений (аксиом, логических правил). С помощью правил, с одной стороны, можно правильно интерпретировать смысловые отношения в конкретном контексте решения задачи, то есть контролировать целостность определений пользователя, а с другой стороны, можно расширять базовый список понятий за счет использования зависимостей

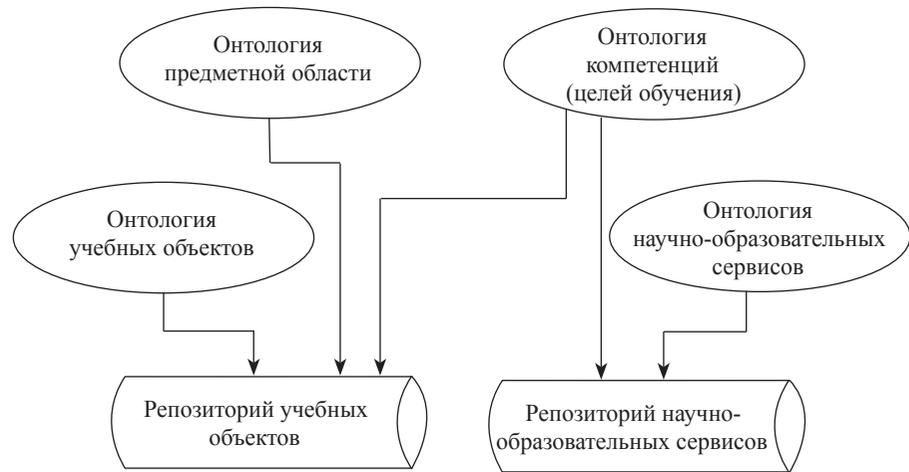


Рис. 2. Семантическая разметка научно-образовательного контента и сервисов с помощью онтологий

своих свойств, которые связывают их с другими понятиями, то есть осуществлять логический вывод. Различают как стандартные правила, например, симметричности, транзитивности, инверсности, так и специализированные правила.

Для информационно-образовательного пространства онтологическое знание организуется в виде четырех онтологий (рисунок 2):

- онтология предметной области, с помощью которой задается семантика профессиональной деятельности;
- онтология учебных объектов, с помощью которой обеспечивается управление учебным контентом в SCORM-спецификациях;
- онтология компетенций, с помощью которой выстраивается последовательность изучения учебных материалов.
- онтология сервисов, с помощью которой определяется классификация научно-образовательных задач.

С помощью модуля семантической разметки на основе онтологий осуществляется аннотирование и индексирование источников знаний в репозиториях учебных объектов и научно-образовательных сервисов (рисунок 2).

4. Методы отбора и адаптации сервисов для решения научно-образовательных задач

Научно-образовательные задачи, решаемые на основе информационно-образовательного пространства, относятся к классу интеллектуальных задач и отличаются слабой формализованностью, предполагающей нечеткость постановки целей и описания условий

решения задачи. Кроме того, уровень знаний и система критериев оценки решения у различных пользователей могут отличаться. Обычно типичное решение интеллектуальной задачи сводится к следующим шагам:

- Формулировка постановки задачи.
- Отбор релевантных задаче источников знаний.
- Понимание отобранного материала (обучение, консультация).
- Решение задачи (выработка, конфигурация решения).
- Проверка допустимости решения задачи (оценка решения задачи, консультация с экспертами, экспертными системами).
- Принятие решений и мониторинг его реализации.
- Запоминание результатов решения задачи в ИОП.

На каждом шаге решения интеллектуальной задачи может использоваться база знаний. При этом работа сводится к итерационной серии поисков в ИОП, обеспечивающей точность достижения цели каждого этапа.

На каждом этапе решения интеллектуальной задачи определяются требования по входу и выходу, используемым методам и средствам, а также используемым ресурсам (конкретным исполнителям, временным, материальным и стоимостным ограничениям) и часто внешне выглядит как заполнение некоторой экранной формы на основе семантического шаблона, в котором фиксируются:

- Общая цель активности.
- Контекстная информация, известная из состояния процесса или задаваемая вручную на данном шаге.

Экранная форма, как правило, заполняется с помощью списков ключевых слов, связанных с рубриками онтологии. Таким образом, экранная форма запроса отражает текущий локальный контекст поиска, который определяет особенности выполнения следующего шага решения задачи.

Повышение интеллектуальных способностей ИОП по поиску, навигации пользователя в информационном пространстве, организация взаимодействия пользователей на каждом шаге решения задачи обеспечивается в результате интерпретации онтологического знания, отбору релевантных для выполнения сервисов.

Для решения перечисленных задач в работе предлагается семантический шаблон описания сервиса, реализующего ту или иную активность (функцию), который представляется кортежем следующих атрибутов (ролевых признаков):

Semantic Template (VD, G, I, In, Act, Rec, T, M, Source, Service)

где *VD* – вид деятельности (например, проектирование),
G – цели (например, нормализация базы данных, сокращение затрат на обработку данных и др.),
I – вход (например, документы, показатели, функциональные зависимости),
O – выход (например, отношения в третьей нормальной форме),
In – инициатор (например, руководитель проекта),
Act – актер (например, проектировщик),
Rec – получатель (например, руководитель проекта),
T – период времени (составление технического проекта),
M – метод решения задачи (метод Д.Кодда),
Source – источник знаний (ссылки на описание метода нормализации, примеры лучшей практики и т.д.),
Service – ссылка на информационно-образовательный сервис, который предназначен для обработки источников знаний.

Алгоритм отбора и адаптации сервисов можно представить следующим образом:

1. Выбор семантического шаблона функции, начиная с определения вида деятельности.

2. Ручное заполнение ролевых признаков семантического шаблона значениями понятий из доменов предметной онтологии.

3. Семантический контроль заполнения шаблона и расширение по правилам проверки зависимостей понятий.

4. Вычисление семантического расстояния близости метода исполнения функции заданным исходным условиям. В случае неоднозначности получаемых решений выполняется моделирование вариантов.

5. Декомпозиция функции на подфункции по правилу агрегации/декомпозиции в соответствии с методом выполнения функции, выбранным на шаге 4.

6. Передача наследуемых значений ролевых признаков декомпозированным функциям.

Переход при необходимости продолжения процесса конфигурирования на пункт 1.

В процессе выбора значений их доменов может выполняться проверка корректности задания значений ролевых признаков и доопределение связанных ролевых признаков по правилам проверки зависимостей понятий.

Таким образом, сгенерированный на основе интерпретации семантического шаблона набор ссылок на имеющиеся в ИОП ресурсы позволяет автоматизировать выбор и адаптацию адекватных сервисов для решения разнообразных научно-образовательных задач.

Заключение

Предложенные в статье принципы и методы структурирования информационно-образовательного пространства (ИОП) на основе онтологического подхода направлено на повышение актуальности и эффективности коллективного использования в образовании и научных исследованиях распределенного в вычислительной сети контента, имеющего различия в стандартах, платформах, структуре и форме представления. В результате обеспечивается быстрая адаптация информационно-образовательных сервисов к быстро изменяющимся образовательным программам и потребностям научно-педагогических работников и обучающихся, многократное повторное использование сервисов. Тем самым, достигается:

- расширение возможностей для создания новых и актуализации существующих образовательных программ и курсов;

- гибкость настройки научно-образовательных процессов из сервисов в соответствии с изменяющимися условиями;

- сокращение затрат на владение и сопровождение аппаратно-программных комплексов благодаря сервисно-ориентированной архитектуре.

Литература

1. Thomas Erl. SOA Principles of Service Design – Prentice Hall Ptr, 2007.
2. Dirk Krafzig, Karl Banke, and Dirk Slama. Enterprise SOA. Prentice Hall, 2005
3. Тельнов Ю.Ф., Данилов А.В., Казаков В.А. Сервисно-ориентированная архитектура динамической интеллектуальной системы управления бизнес-процессами // Научно-практический журнал "Открытое образование", № 6, 2010
4. Тельнов Ю.Ф., Казаков В.А., Козлова О.А., Динамическая интеллектуальная система управления процессами в информационно-образовательном пространстве высших учебных заведений // Научно-практический журнал "Открытое образование", №1, 2013
5. Трембач В.М. Методы формирования, использования и анализа баз знаний // Экономика, статистика и информатика, Вестник УМО, № 4, 2010.
6. Toward a Technology for Organizational Memories / Abecker A., A.Bernardi, K.Hinkelman, O.Kuhn // IEEE Intelligent Systems. –1998. –No.3.
7. Тельнов Ю.Ф. Реинжиниринг бизнес-процессов: компонентная методология. – М.: Финансы и статистика, 2004 (2-е изд-е).
8. Гаврилова Т.А., Гулякина Н.А. Визуальные методы работы со знаниями: попытка обзора // Искусственный интеллект и принятие решений, 2008, №1.
9. Кудрявцев Д.А. Системы управления знаниями и применение онтологий. – СПб.: Изд-во Политехн. ун-та, 2010.
10. Gruber T.R. Translation Approach to Portable Ontologies // Knowledge Acquisition. – 1993. – No 5 (2).

References

1. Thomas Erl. SOA Principles of Service Design – Prentice Hall Ptr, 2007.
2. Dirk Krafzig, Karl Banke, and Dirk Slama. Enterprise SOA. Prentice Hall, 2005

3. Telnov Yu.F., Danilov A.V., Kazakov V.A. Service-oriented architecture of a dynamic intellectual control system of business processes // Nauchno-prakticheskij zhurnal "Otkrytoe obrazovanie", №6, 2010

4. Telnov Yu.F., Kazakov V.A., Kozlova O.A., Dynamic intellectual control system of processes in information and educational space of higher educational institutions // Nauchno-prakticheskij zhurnal "Otkrytoe obrazovanie", №1, 2013

5. Trembach V.M. Methods of formation, use and analysis of knowledge bases // Ekonomika, statistika i informatika, Vestneyk UMO, № 4, 2010.

6. Toward a Technology for Organizational Memories / Abecker A., A. Bernardi, K. Hinkelman, O. Kuhn // IEEE Intelligent Systems. – 1998. – No.3.

7. Telnov Yu.F. Reengineering of business processes: component methodology. – M.: Finance and statistics, 2004 (2nd prod.-e).

8. Gavrilova T.A., Gulyakina N.A. Visual methods of work with knowledge: review attempt // Iskusstvennyj intellekt i prinyatie reshenij, 2008, №1.

9. Kudryavcev D.A. Control systems of knowledge and application of ontologies. – SPb.: Izd-vo Politehn. un-ta, 2010.

10. Gruber T.R. Translation Approach to Portable Ontologies // Knowledge Acquisition. – 1993. – No 5 (2).