

# ОНТОЛОГИИ В СИСТЕМАХ УПРАВЛЕНИЯ ЗНАНИЯМИ

**ЗАХАРОВА О. В.**

кандидат экономических наук

Харьков

**В** истории человеческой цивилизации, связанной с технологиями и системами накопления и распространения знаний, можно выделить три ключевые достижения. Первое – появление письменности; второе – открытие книгопечатания, и третье – становление Internet и Web-технологий. Развитие Internet-пространства кардинально меняет коммуникационные возможности людей, существенно затрагивает интеллектуальные сферы их деятельности, влияет на создание и использование электронных образовательных ресурсов, формирование и развитие новых форм бизнес-процессов.

В начале 60-х годов прошлого века возникла идея создания глобальной сети знаний и интерактивных компьютерных вычислений. Технологические решения приобрели реальную почву после изобретения Т. Бернерсом-Ли – World Wide Web. В конце 80-х годов он разработал язык разметки HTML, сетевой протокол передачи гипертекстовых данных HTTP и первый WWW-браузер. Сегодня – возглавляет Консорциум W3C, – рекомендательный орган, разрабатывающий стандарты и определяющий пути развития World Wide Web.

В современных Web-технологиях наряду с HTML широко используется разработанный в 1996 г. язык разметки XML и его разновидности, который стал основным форматом при обмене данными в автоматизированных системах обучения и управления знаниями. В развитии последних, особо важное значение имеют такие направления интеллектуализации как: автоматическое распознавание семантики текста и речи, извлечение и структуризация знаний, создание баз образовательных ресурсов, поиск информации в хранилищах большого объема, системы поддержки управленческих решений, основанные на интеллектуальном анализе данных.

Для структуризации знаний и поиска нужных сведений разрабатываются классификационные системы и способы представления знаний в виде семантических сетей – Semantic Web. Известно, насколько важную роль в естествознании сыграла первая научная классификация К. Линнея (1735 г.). Сегодня семантические сети получают все более широкое распространение в различных приложениях. Наиболее популярными направлениями в этой области являются: Data Mining (DM), Knowledge Discovery in Databases (KDD), Machine Learning (ML), Business Intelligence (BI). Они предоставляют теоретическую и методологическую базу для изучения, анализа и понимания огромных объемов данных [1]. Однако этих методов недостаточно, если сама структура данных не будет распознаваться машинными алгоритмами.

На сегодняшний день существует большое количество определений понятия онтология, например, *онтология* – это:

- ✦ набор базовых понятий и отношений между ними;
- ✦ эксплицитная спецификация концептуализации, где в качестве концептуализации выступает описание множества объектов предметной области и связей между ними (по Т. Груберу);
- ✦ формальная спецификация разделяемой концептуализации, которая имеет место в некотором контексте предметной области;
- ✦ база знаний, описывающая факты, которые предполагаются всегда истинными в рамках определенного общества на основе общепринятого смысла используемого словаря;
- ✦ конструкция для представления знаний в Internet, представляющая собой множество объектов, классифицированных в соответствии с некоторыми критериями, а также содержащими описание свойств этих объектов (по А. В. Манциводу);
- ✦ попытка всеобъемлющей и детальной формализации некоторой области знаний с помощью концептуальной схемы.

**О**бычно онтологии строятся из экземпляров, понятий, атрибутов и отношений. Экземпляры – элементы самого нижнего уровня; главной целью онтологий является именно классификация экземпляров. Понятия – абстрактные наборы, коллекции объектов. Атрибуты используются для хранения специфичной информации об объекте. Отношения – позволяют задать зависимости между объектами онтологии.

Формальная модель онтологии  $O = \langle T, R, F \rangle$  – это упорядоченная тройка конечных множеств, где:

$T$  – термины прикладной области, которую описывает онтология  $O$ ;

$R$  – отношения между терминами заданной предметной области;

$F$  – функции интерпретации, заданные на терминах и/или отношениях онтологии  $O$ .

В инженерии знаний под онтологией понимается детальное описание некоторой проблемной области, которое используется для формального и декларативного определения ее концептуализации. Онтологии позволяют представить понятия в виде, пригодном для машинной обработки; нередко они используются в качестве посредника между пользователем и информационной системой, позволяют формализовать договоренности о терминологии между членами сообщества, например, между пользователями некоторого корпоративного хранилища данных.

Простейший алгоритм онтологического инжиниринга включает: выделение концептов – базовых поня-

тий данної предметної області; визначення «висоти дерева онтологій» – кількість рівнів абстракції; розподілення концептів по рівнях; побудова зв'язів між концептами – визначення відносин і взаємодій базових понять; консультації з різними спеціалістами для виключення протиріччя і неточностей.

**А**ля того щоб реалізувати різні онтології, необхідно розробити мови їх представлення, які мають достатню виразну силу і дозволяють користувачеві уникнути «низкорівневих» проблем. Розповсюдження онтологічного підходу до представлення знань сприяло розробці різних мов і інструментальних засобів, призначених для створення, редагування і аналізу онтологій.

Серед традиційних мов специфікації онтологій слід виділити Ontolingua, CycL, а також засновані на описувальній логіці (наприклад, LOOM) і мови, засновані на фреймах (OKBC, OCML, Flogic). Більш пізні мови засновані на Web-стандартах (XOL, SHOE, UPML). Спеціально для обміну онтологіями через Web були створені RDF(S), DAML, OIL, OWL.

В межах проекту семантичної інтерпретації інформаційних ресурсів Internet (Semantic Web) було запропоновано стандарт опису метаданих про документ Resource Description Framework (RDF), який використовує XML-синтаксис:

- ✦ XML – надає можливість створення структурованих документів, але не пред'являє до них жодних семантичних вимог;
- ✦ XML Schema – визначає структуру XML документів і додатково дозволяє використовувати конкретні типи даних;
- ✦ RDF – надає можливість описувати абстрактні моделі даних деяких об'єктів і відносин між ними. Використовує просту семантику на основі XML синтаксису;
- ✦ RDF Schema – дозволяє описувати властивості і класи RDF – ресурсів, а також семантику відносин між ними;
- ✦ OWL – розширює описувальні можливості попередніх технологій. Дозволяє описувати відносини між класами (наприклад, непересекаємість), кардинальність (наприклад «точно один»), симетрія, рівність, перераховувані типи класів.

Онтології OWL – це послідовності аксіом і фактів, а також посилань на інші онтології. Вони містять компоненту для запису авторства і іншої детальної інформації, є документами Web, на них можна посилатися через URI (Uniform Resource Identifier) – уніфікований ідентифікатор ресурсу. На даний момент мова OWL є основним інструментом опису онтологій; за ступенем виразності виділяють три її діалекти – OWL Lite, OWL DL і OWL Full.

Одним з важливих переваг онтологій є наявність інструментального програмного забезпечення, що підтримує їх аналіз: редагування, візуалізацію, документування, імпорт і експорт онтологій різних форматів, їх представлення, об'єднання, порівняння. Серед існуючого ряду інструментів окремо варто виділити наступні редактори: Ontolingua, Protégé, DOE, OntoEdit, OilEd, WebOnto, WebODE і др.

Реальне навчання відбувається на рівні засвоєння (формування) понять предметної області і встановлення зв'язів між ними. Процес структуризації предметних знань на понятійному рівні є процесом побудови предметної онтології. Практична розробка онтології включає: визначення класів в онтології; розташування класів в таксономічній ієрархії; визначення слотів і описування допустимих значень цих слотів; заповнення значень слотів екземпляром.

**О**сновні цілі застосування інформаційних освітніх технологій полягають у підвищенні якості освіти. Сучасне навчання породжує більш суворі вимоги до технологій інформатизації освіти, і зокрема, ці вимоги стосуються проблем управління навчанням і створення навчального контенту. Першою освітньою системою, що використовує в освіті електронні освітні ресурси онтології предметних областей, стала система CTS, розроблена в МГТУ ім. Н. Е. Баумана в 1993 р. В сучасному вигляді технологія розділяється на одиниці контенту на основі онтологічного підходу представлена в системі Базис і генератора освітніх ресурсів (БіГОР). Онтологічний підхід в поєднанні з концепцією модульності дозволяє реалізувати синтез індивідуальних маршрутів навчання, а також розробляти нові сучасні системи і онтологічні портали знань [2].

Поняття онтології і онтологічного аналізу ввійшли сьогодні в процедури і стандарти моделювання бізнес-процесів. Опис бізнес-процесу – це, по суті, структуризація даних і знань. Онтологічне дослідження складних систем дозволяє накопичити цінну інформацію про їх роботу, результати аналізу якої будуть мати вирішальне значення при проведенні процесу реорганізації існуючих і побудові нових систем.

Для моделювання складних систем розроблено ряд методологій, наприклад, методології родини IDEF (Integrated DEFinition). IDEF містить 14 державних стандартів США, створених в межах запропонованої ВВС США програми комп'ютеризації промисловості ICAM. Вони призначені для аналізу процесів взаємодії в виробничих системах. Для підтримки онтологічного аналізу призначена методологія IDEF5. Процес побудови онтології в IDEF5 підтримують спеціально розроблені онтологічні мови: схематична мова (Schematic Language-SL) і мова допрацювань і уточнень (Elaboration Language-EL).

Существуют четыре основных вида схем, которые используются для накопления информации об онтологии в графической форме: диаграмма классификации (Classification Schematics) – обеспечивает механизм для логической систематизации знаний, накопленных при изучении системы; композиционная схема (Composition Schematics) – механизм графического представления состава классов онтологии, позволяющий описывать, что из каких частей состоит, т. е. наглядно отображать состав объектов, относящихся к тому или иному классу; схема взаимосвязей (Relation Schematics) – инструмент визуализации и изучения взаимосвязей между различными классами объектов в системе; диаграмма состояния объекта (Object State Schematics) – средство документации процессов с точки зрения изменения состояния объекта.

Таким образом, диаграммы состояния в IDEF5 наглядно представляют изменения состояния или класса объекта в течение всего хода процесса. При построении концептуальной модели используются предметные знания в виде набора понятий и связывающих их отношений.

**С**троение и свойства любой системы могут быть эффективно исследованы при помощи словаря терминов, используемых при описании характеристик объектов и процессов, имеющих отношение к рассматриваемой системе, точных и однозначных определений всех терминов этого словаря и классификации логических взаимосвязей между этими терминами. Набор этих средств и является онтологией системы, а стандарт IDEF5 предоставляет структурированную методологию, с помощью которой можно наглядно и эффективно разрабатывать, поддерживать и изучать эту онтологию [3].

Для современного бизнеса характерно постоянное изменение среды, в которой адаптируются и функционируют сообщества, организации, люди. В условиях конкуренции судьба бизнеса во многом определяется скоростью и точностью реакции компании на внешние дестабилизирующие воздействия. Деятельность, как отдельных людей, так и организаций все в большей степени зависит от имеющихся у них знаний – одного из самых ценных ресурсов – и способности их эффективно использовать.

Управление знаниями может рассматриваться сегодня как мощное конкурентное преимущество. Почти все компании обладают огромным исходным багажом данных и практического опыта, но пока эта информация рассредоточена в базах данных, хранилищах документов, сообщениях электронной почты, отчетах о продажах, в головах сотрудников, проблемой остается организация доступ к этим данным, придание им формы, удобной для использования.

Таким образом, управление знаниями – это стратегия, цель которой – выявить и обратить на пользу всю имеющуюся информацию, опыт и квалификацию персонала. Одним из важнейших и перспективных направлений в области формализации знаний, которое дает возможность использования накопленных знаний для компьютерной обработки, являются онтологии. Онтологический подход к проектированию систем управле-

ния знаниями позволяет создавать системы, в которых знания, накопленные внутри организации, становятся доступными для большинства пользователей.

Факторы разрозненности информации и использования разнообразных терминов в различных предметных областях значительно усложняют взаимопонимание, поэтому необходимо разрабатывать формализованные модели представления знаний, которые обеспечивали бы обработку информации на семантическом уровне в системах управления знаниями и бизнес-процессах.

В настоящее время существует значительный интерес к системам управления знаниями со стороны промышленных компаний, которые осознают высокий прикладной потенциал систем, базирующихся на знаниях и используемых для решения целого ряда практических задач. Вопросы управления знаниями имеют решающее значение для развивающейся экономики, где знание капитализируется и, поэтому, приобретает совершенно иной статус; применение онтологии в этой связи может играть ключевую роль.

**В** условиях экспоненциального роста объема данных в мире вопросы быстрого релевантного поиска необходимой информации, ее структуризации и на этой основе извлечения новых знаний и принятия решений можно эффективно осуществлять привлекая интеллектуальные технологии, в частности, Semantic Web, онтологии. ■

#### ЛИТЕРАТУРА

1. **Гладун А. Я., Рогушина Ю. В., Петрухина Л. В.** Использование технологии Semantic Web для управления знаниями в системах Business Intelligence // Штучний інтелект.– 2010, № 2.– С. 15 – 23.

2. **Шевченко О. Ю., Климова М. В.** Онтологічна система менеджменту національними та науковими ресурсами // Системи обробки інформації.– 2009, Вип. № 4 (78).– С. 146 – 151.

3. **Шевченко О. Ю., Климова М. В.** Метод побудови інтелектуальних систем обробки інформації та документообігу за допомогою онтологічної бази знань // Штучний інтелект.– 2009.– № 2.– С. 91 – 97.