

## ЕДИН ОНТОЛОГИЧЕН ПОДХОД КЪМ ПРОЦЕСА НА ОБУЧЕНИЕ

**гл. ас. Тодорка Терзиева**

Пловдивски университет „Паисий Хилендарски“  
Факултет по математика и информатика  
4003 Пловдив, бул. „България“ 236  
e-mail: dora@uni-plovdiv.bg

**доц. дмн Георги Тотков**

Пловдивски университет „Паисий Хилендарски“  
Факултет по математика и информатика  
4003 Пловдив, бул. „България“ 236  
e-mail: totkov@uni-plovdiv.bg

*В настоящата работа е проведено изследване в областта на извличане и представяне на знания на базата на онтологичен подход. Описано е понятието онтология и основни принципи за нейното построяване. Направен е сравнителен анализ на различни системи и средства за създаване на т. нар. Web-онтологии и се изследват аспекти на тяхното използване в е-обучението. В резултат е предложена методика за създаване на компютърни програми за обучение, базирани на онтологични бази от знания.*

**Ключови думи:** онтология, онтологичен подход, обучение, онтологичен инженеринг

### 1. Въведение

Едно от най-важните и перспективни направления в областта на формализацията на знания, което дава възможност за компютърна обработка на натрупаните знания е онтологията и разработените на нейна база модели. Съвременните тенденции на развитие на Интернет са насочени към преход от документи, които компютъра може да прочете, към документи, които компютъра може смислово да разбере. Този нов подход на развитие на Интернет получи наименованието ‘Semantic Web’.

Основна характерна черта на онтологическия анализ се явява разделянето на реалния свят на части и класове обекти и определяне на съвкупността от фундаментални свойства, които определят тяхното развитие и поведение.

Много изследователи отбелязват перспективността от използване на онтологични системи за обучение в дадена предметна област [5]. При този ‘онтологичен’ подход се разграничават две основни направления. В рамките на първото направление, онтологичният подход осигурява коректен подход при концептуалното описание на елементите и структурата на съответната област на знанието и съществуващите в нея отношения и взаимовръзки. Второто направление е свързано с построяване на програмни системи, базирани на подходящ концептуален модел на предметната област, съставен от класове, съдържащи се в онтологията.

Използването на web-онтологии в учебния процес позволява определяне на основните компоненти на учебните дисциплини, осигуряване на възможност за организация на ефективно разпределение на достъпа до учебни ресурси, създаване на единна база знания, която съчетава в себе си множество учебни дисциплини и която може да бъде споделена в Интернет, и др.

### 2. Понятието ‘онтология’

В съвременната философска литература терминът ‘онтология’ се използва за означаване на определен набор от категории, които са следствие от определена система възгледи за света. Терминът се използва и за специфициране на множество от понятия с цел представяне на дадена предметна област (ПО).

В областта на информатиката понятието ‘онтология’ се използва за пръв път през 1993 г. от Грубер. Според Грубер [3], под ‘онтология’ се разбира експлицитна (явна) спецификация на определена концептуализация (множество от обекти, процеси и връзки между тях) на съответната ПО. От формална гледна точка, онтологията се състои от понятия, организирани в таксономия, придружени с техни описания, атрибути и свързаните с тях аксиоми и правила за извод.

В информатиката онтологията се дефинира като опит за всеобхватно и детайлно формализиране на някаква област на знанието с помощта на концептуална схема. Обикновено тази схема се състои от йерархични структури от данни, съдържащи съответни обекти, връзките между тях и правилата, приети в тази област. Прилага се като форма за представяне на знания за реалния свят или за части от него. За описание на онтологията съществуват различни езици, като за случая най-подходящ е визуалният подход, който дава възможност нагледно да се формулират и обяснят природата и структурата на явленията.

Разпространението на онтологичния подход за представяне на знания довежда до създаване на разнообразни езици за представяне на онтологии и инструментални средства за тяхното редактиране и анализ.

Онтологиите могат да бъдат **класифицирани** в зависимост от различни признаци:

- По степента си на зависимост *от конкретната задача*;
- Според *езика за представяне* на онтологичните знания и неговите възможности;
- Според равнището на *детайлизация*;
- По *предметни области*.

Общите онтологии (онтологиите от най-високо равнище) са предметно независими, съдържащи 'top-level' категории (концепти) като пространство, време, обект, събитие, действие, количество, мярка и т.н., които са независими от конкретната задача или ПО. Затова е целесъобразно те да бъдат унифицирани за големи общности от потребители.

Пример за подобна обща онтология е СУС [7]. Едноименният проект СУС на Сусогр е ориентиран към създаването на мултиконтекстна база от знания и специална машина за извод. Основната цел на този огромен проект е построяването на база от знания за всички общи понятия (започвайки с понятия от типа на време, същност и т.н.), която включва семантичната структура на термините, връзките между тях и множество от аксиоми. Друг пример за обща онтология е онтологията на системата Generalized Upper Model [2], ориентирана към поддръжка на процесите на обработка на естествен език (английски, немски и италиански). Равнището на абстракция на тази онтология се намира между лексическите и концептуалните знания, и се определя от изискванията за опростяване на интерфейсите с лингвистични ресурси. Моделът на Generalized Upper Model включва таксономия, организирана под формата на йерархия между категориите (около 250 понятия) и отделна йерархия на връзките.

Създаването на достатъчно общи онтологии от най-високо равнище е много сериозна задача, която все още няма задоволително решение.

Предметните онтологии и онтологиите на задачи включват речник, свързан съответно с предметната област (медицина, търговия и т.н.) или с конкретната задача/дейност (диагностика, продажби и т.н.) за сметка на специализация на понятията, включени в общата онтология. Примери за онтологии, ориентирани към определена предметна област или задача, са съответно TOVE и Plinius.

Приложните онтологии описват концепти, които зависят както от конкретната ПО, така и от задачите, които ще се решават. Концептите в такива онтологии често съответстват на ролите, които играят обектите в предметната област в процеса на изпълнение на определена дейност. Пример за подобна онтология е онтологията на системата Plinius [13], предназначена за полуавтоматично извличане на знания от текстове в областта на химията. За разлика от другите, споменати по-горе онтологии, тук няма явно зададена таксономия на понятията. Вместо това са определени няколко множества от атомарни концепти (като например химически елемент, цяло число и т.н.) и правила за конструиране на останалите концепти. В онтологията са описани около 150 концепти и 6 правила. Онтологията на Plinius също се описва с помощта на фреймове.

Друга класификация се основава на типа на включения в онтологията речник на понятията. В зависимост от типа на използвания речник, онтологиите се разделят на четири типа: силно неформални (изразени в свободна форма на естествен език), полунеформални (изразени в ограничена и специално структурирана форма на естествен език), полуформални (изразени в термините на изкуствено създаден език) и формални (с точно дефинирани термини и формално описана семантика, с теореми за пълнота и съответни доказателства).

### 3. Сравнителен анализ на системи за създаване на Web-онтологии

Процесът на проектиране и разработване на онтологии е известен под името 'онтологичен инженеринг'. Методиката за построяване на дадена онтология [4] предполага определяне на целите и областта на нейното приложение. Необходимо е предварително да се определи – с каква цел се създава онтологията, за какви типове въпроси се предполага да бъдат получавани отговори (с нейна помощ), как ще се използва и поддържа. За проектиране и създаване на онтология в дадена ПО е необходима съвместна работа на много специалисти. Нейната разработка започва с определяне на предметната област и обхват на онтологията, създаване на модели на реалния свят и на съответните понятия, и т.н. При това се спазват няколко основни принципа:

- *Формализация* – описание на предметната област в единни, строго определени образци (термини, модели и др.);
- Въвеждане на *ограничено количество базови понятия*, на основата, на които се конструират всички останали (съставни) понятия;
- Вътрешна *пълнота и непротиворечивост* на използваните понятия;

- Използване на *инструменти за създаване на онтологии*, поддържащи редактиране, визуализация, документиране, импорт и експорт на онтологии.

Система Ontolingua [1] е разработена в KSL (Knowledge Systems Laboratory) на Станфордския университет. Състои се от сървър и език за представяне на знания. Редакторът на онтологията е Web-приложение на основата на HTML, което включва мрежов компонент Webster (предназначен за определяне на концептите), сървър (осигуряващ достъп до онтологията по протокола OKBC – Open Knowledge Base Connectivity [8]) и Chimaera (инструмент за анализ, обединение и интегриране на онтологията). Всички приложения, освен сървъра OKBC са реализирани на HTML. Системата за представяне на знания е реализирана на функционалния език за програмиране Lisp. Онтологията може да бъде съхранена в различни цифрови формати (за използване от други приложения) или да бъде импортирана от други езици в езика на Ontolingua.

Protégé (създадена от група специалисти по медицинска информатика в Станфордския университет [10]) е свободно разпространявана Java-програма, предназначена за създаване на онтологии. Включва редактор на онтологии, позволяващ разгръщането на йерархична структура от абстрактни и конкретни понятия от съответната ПО. Структурата на онтологията е аналогична на йерархичната структура на обикновен каталог. На нейната основа може да се генерират форми за получаване на знания. Предлага графичен интерфейс, възможности за извеждане на справки и примери. Поддържа езика OWL и генериране на html-документи, представящи структурата на онтологията. Protege е основана на фреймовия модел OKBC за представяне на знания и на редица от плъгини, което позволява да се редактират модели на предметни области, създадени не чрез OWL, а с други езици – UML, XML, SHOE, DAML+OIL, RDF и RDFS и др.

OntoEdit [6] е инструментално средство, осигуряващо редактиране на онтологии. Първоначално е разработен в институт AIFB (университета на Карлсруе), а днес се разпространява под формата на комерсиализираната система Ontoprise. Извършва проверка, преглед, кодиране и модификация на онтологии. OntoEdit поддържа език за представяне Flogis, включващ машина за експорт, езиците OIL и RDFS, собствен език за представяне на знания OXML, базиран на XML и реализиран като автономно Java-приложение.

Система WebOnto [11] също е предназначена за разглеждане, създаване и редактиране на онтологии. Отличава се с простота на използване и предоставяне на средства за мащабиране и създаване на големи онтологии. За моделиране на онтологията използва езика OCML (Operational Conceptual Modeling Language). Има възможност за съхраняване на структурни диаграми, преглед на отношения, класове, правила и др., както и за съвместна разработка от няколко потребители.

SUMO [8] (Suggested Upper Merged Ontology е безплатна, свободно разпространявана онтология от високо равнище, принадлежаща на IEEE (Institute of Electrical and Electronics Engineers). SUMO не е предназначена за решаване на конкретна задача, а цели интегриране на съществуващите онтологии в една обща онтология, която би имала статус на универсален стандарт, и би могла да се използва в различни приложения и изследователски проекти [9]. SUMO съдържа около 1000 понятия и 5000 аксиоми (от тях около 800 правила) и е с възможности за разширяване на своята структура. Конкретните понятия се съхраняват в съответни отраслови онтологии и могат да се включват при необходимост.

CYC [7] е многоцелева база от знания и машина за логически изводи, ориентирана за използване в бизнес-процеси. Базата знания се описва на езика CycL, който е хибриден език, съчетаващ в себе си изразителността на фреймите и логиката на предикати [2]. CycL различава индивиди, класове, предикати и функции. Онтологията се разделя на отделни микро-теории, които се изграждат в йерархия, като равнищата на йерархията достигат 50. Към настоящия момент, CYC съдържа около 400 хиляди концепти и повече от 3,5 милиона аксиоми. Най-важният компонент на базата знания на онтологията е лексикон, построен на основата на WordNET [12]. Лексиконът се описва също на езика CycL и е неразделна част от базата знания.

По-ранни инструменти като Ontolingua, Ontosaurus и WebOnto имат клиент-сървър архитектура, докато Protege, OntoEdit и CYC имат 3-степенна архитектура. Повечето средства за разработка съхраняват онтологията в текстови файлове, докато Protege, WebODE и CYC могат да управляват големи онтологии и да ги съхраняват в бази от данни. Почти всички реализации са на Java.

#### **4. Web-онтологията в процеса на обучение**

Приложението на различни онтологични системи в съвременното електронно обучение е все още в начален етап на изследване. Ще отбележим, че дори и при конвенционалните форми на обучение, не са често явление методически разработки и успешни практики, свързани с използване на подобни системи в учебния процес.

От друга страна, интензивните научни изследвания и постижения в областта на онтологичните системи, а така също ускореното развитие на информационните технологии (специално в областта на електронното обучение), дават добра основа за провеждане на подобни изследвания и експерименти.

Възможни са различни сценарии за използване на web-онтологии в процеса на е-обучение.

**Първият сценарий** се състои в непосредствено използване на съществуваща онтологична система като **допълнителен ресурс в процеса на обучение**. Динамичното свързване на подобен ресурс с конкретна система за е-обучение, би могло да се използва за *получаване на справки* относно изучавана учебна дисциплина или тема (от обучавани, автори и преподаватели), *автоматично създаване на тестови задачи и проверка на отговори* (с използване на съответната машина за извод), и др. В този случай, в зависимост от контекста (цели, образователно равнище, учебна дисциплина и др.) може да се наложи допълване на онтологичната система – с нови понятия, отношения или атрибути. Така например, в случая на обучение за деца (предучилищна или начална образователна степен), е целесъобразно добавяне на мултимедийни атрибути (изображения, видеофилми, и др.) за понятията и отношенията в използваната онтология.

Друго интересно развитие на изследванията в областта е използване на понятията в дадената онтология като основа за създаване на индекс за изучаваната дисциплина. За целта е необходимо провеждане на лингвистичен анализ на текстовете, използвани в процеса на е-обучение с цел – автоматичното им свързване със системата от понятия и отношения на онтологията.

**Вторият възможен сценарий** предполага включване на обучаваните в **развитие на онтологията за изучаваната ПО**. Например, механизмът на езика OWL позволява индуктивно да се построи базата знания, като се започне с отделна дисциплина. Подобен подход би могъл да се използва за създаване на нови и развитие на съществуващи учебни курсове.

Реалистичен е и подход, при който отначало може да се започне със създаване на онтологична система (свързана с преследвани локални цели и конкретна изучавана ПО, вкл. понятия, отношения и атрибути), а на следващ етап да се продължи с експерименти съгласно първия или втория сценарий. Предимствата на подхода се заключават в създаване на 'бързи' (реализуеми без особен разход на проектантски и програмистки труд) и 'леки' (не изискват наличие на голяма онтологична база) прототипи на системи за е-обучение, базирани на онтологичен подход, и изследване на техните възможности и практическа приложимост. На базата на получените резултати е възможно да се пристъпи и към изследване на възможностите за използване на известни онтологични бази от данни и системи за провеждане на е-обучение в конкретни ПО.

### Литература

- [1] Farquhar A., Fikes R., Rice J. The Ontolingua server: A tool for collaborative ontology construction / International Journal of Human-Computer Studies, 46 (6), pages 707–728, 1997.
- [2] Gomez-Perez A., Fernando-Lopez M., Corcho O. Ontological Engineering with examples from the areas of Knowledge management, e-Commerce and the Semantic Web – 2nd. Ed. – London, Springer-Verlag – 2004.
- [3] Gruber T. R. A translation approach to portable ontologies / Knowledge Acquisition, 1993, V. 5(2), P. 199-220.
- [4] Mike Ushold, Michael Gruninger (1996) Ontologies: Principles, Methods and Applications, Knowledge Engineering Review, Volume 11, Number 2.
- [5] Mizoguchi R. Ontology Engineering Environments / Handbook on Ontologies; ред. Staab S., Studer R. – Berlin, Springer-Verlag – 2003.
- [6] OntoEdit <http://www.ontoprise.de/> (5.11.2008 г. последно посетен).
- [7] Overview of Cycorp's R&D, <http://www.cyc.com/cyc/technology/cycrandd> (5.11.2008 г. последно посетен).
- [8] Suggested Upper Merged Ontology (SUMO), <http://www.ontologyportal.org/index.html> (5.11.2008 г. последно посетен).
- [9] The Generalized Upper Model <http://glotta.ntua.gr/StateoftheArt/Ontologies/newUM.html> (5.11.2008 г. последно посетен).
- [10] The Protégé Ontology Editor and Knowledge Acquisition System, <http://protege.stanford.edu/> (5.11.2008 г. последно посетен).
- [11] WebOnto, <http://riverside.open.ac.uk:3000/webonto> (5.11.2008 г. последно посетен).
- [12] WordNet <http://wordnet.princeton.edu/> (5.11.2008 г. последно посетен).
- [13] The PLINIUS Ontology of ceramic materials, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.23.6679> (5.11.2008 г. последно посетен).