

СОЗДАНИЕ БАЗОВОЙ ОНТОЛОГИИ ДЛЯ РОССИЙСКОЙ СИСТЕМЫ ПРАВА НА ОСНОВЕ ОНТОЛОГИИ LKIF-CORE

И. Е. Воронина, Е. А. Пигалкова

Воронежский государственный университет

Поступила в редакцию 01.03.2010 г.

Аннотация. Статья посвящена проблемам онтологического моделирования в правовой сфере. Рассматриваются особенности разработки правовых онтологий и задачи онтологического моделирования. Приводятся результаты работы по адаптации общепринятой европейской юридической онтологии LKIF-Core для российской системы права с использованием русскоязычных обозначений.

Ключевые слова: представление знаний, онтология, Protege, базовая онтология, право
Abstract. The paper is dedicated to the problems of ontological modeling in the legal field. The subject includes specific features of legal ontology development and problems of ontological modeling, covering the results of adaptation of well-known European legal ontology LKIF-Core for use in Russian legal system with native terminology.

Key words. knowledge representation, ontology, Protege, core ontology, law.

ВВЕДЕНИЕ

Для моделирования и описания правовых явлений и процессов актуальна задача создания систем, основанных на знаниях. Такие системы, включающие в себя элементы экспертной обработки и интеллектуального поиска, могут стать незаменимым инструментом в правотворческой и правоприменительной деятельности, стать основой различных семантических приложений. Кроме того, подобные системы можно применять для решения учебных задач в юриспруденции.

Первым шагом в разработке такой системы является принятие единой *онтологии* – документированной совокупности терминов предметной области и правил, согласно которым эти термины могут быть использованы для построения достоверных утверждений о состоянии системы, а также для санкционированного логического вывода новых утверждений.

Понятие онтологии является достаточно широким по степени формализации. Так, к онтологиям можно отнести даже неструктурированный словарь с определениями. Поскольку в текстах правовых актов достаточно широко

практикуется внедрение определений, построить словарь терминов конкретной отрасли права не так сложно, но такая простая онтология не будет эффективна в использовании. Основное преимущество онтологий, о которых пойдет речь далее, – наличие большого количества отношений, причем не только традиционных таксономических (часть-целое, класс-подкласс), но и иных, характерных для предметной области и служащих для отражения сложных явлений социально-правового характера. Более того, для таких онтологий можно задать правила логического вывода, что позволит получить новые знания на основе уже существующих. В. Д. Соловьев и В. В. Иванов [1] приводят перечень элементов, которые свойственны только для ресурсов онтологического характера. Рассмотрим эти элементы.

1. *Формальность*, то есть точное определение иерархического отношения между классами (is-a) и строго соблюдающие транзитивность этого отношения.

2. *Отношения класс-экземпляр*. Такие онтологии могут содержать, помимо имен классов, экземпляры (индивиды) на нижнем уровне.

3. *Свойства концептов*. Концепты могут иметь информацию о свойствах (слотах). Свойства особенно полезны, когда они определены на верхних уровнях и наследуются подклассами.

4. *Ограничения на область значений свойств.* Вводятся дополнительные ограничения на то, экземпляром какого класса может являться значение данного свойства.

5. *Дизъюнктивные классы, обратные свойства.* Дизъюнктивные классы заведомо не имеют общих экземпляров. Обратные свойства представляют собой пару свойств у двух разных классов, которые определяют отношения между их экземплярами

6. *Аксиомы,* позволяющие создавать произвольные логические ограничения.

Очевидно, что такие элементы по мере использования расширяют возможности применения онтологий для моделирования различных процессов и явлений и делают их уникальным инструментом именно в такой предметной области как социально-правовая.

В целом, потребность в разработке онтологий объясняется следующими причинами [2]:

- 1) совместное использование людьми или программными агентами общего понимания структуры информации;
- 2) возможность повторного использования знаний в предметной области;
- 3) создание явных допущений в предметной области;
- 4) отделение знаний в предметной области от оперативных знаний;
- 5) возможности объективного анализа знаний в предметной области.

Данное направление начало интенсивно развиваться сравнительно недавно (в последние десятилетия), однако одна из методологий уже закреплена в соответствующих стандартах семейства IDEF (ICAM/Integrated Definition), которые создавались в США в рамках программы компьютеризации промышленности ICAM.

IDEF5 – Ontology Description Capture – стандарт онтологического исследования сложных систем. С помощью методологии IDEF5 онтология системы может быть описана при помощи определенного словаря терминов и правил, на основании которых могут быть сформированы достоверные утверждения о состоянии рассматриваемой системы в некоторый момент времени. На основе этих утверждений формируются выводы о дальнейшем развитии системы и производится её оптимизация [3].

Помимо IDEF5 среди известных методологий в области онтологического моделирования

можно также перечислить такие, как METH-ONTOLOGY [4], DILIGENT [5], NeOn [6].

1. ЗАДАЧИ ОНТОЛОГИЧЕСКОГО МОДЕЛИРОВАНИЯ В ПРАВОВОЙ СФЕРЕ

Разработка онтологий в последнее время получила довольно широкое распространение в России, однако нет информации о наличии масштабных исследований именно в правовой области. В рамках Университетской информационной системы РОССИЯ (НИВЦ МГУ) используется Тезаурус по общественно-политической тематике АНО Центр информационных исследований [7], однако нельзя сказать, что данный ресурс имеет именно правовой характер.

Онтологическое моделирование в правовой сфере имеет следующие характерные черты [8]:

- 1) большое количество общепринятых понятий с особой спецификой использования;
- 2) различие в структурах различных отраслей права;
- 3) наличие промежуточного общетеоретического правового уровня между онтологией верхнего уровня и онтологией предметной области;
- 4) большое количество теоретических допущений и абстрактных конструкций, зависящих от специфики правовых взглядов.

При разработке правовых онтологий наиболее эффективным является лингвистический подход, основанный на изучении естественного языка и построении онтологий на больших текстовых массивах (корпусах). Это особенно характерно для восходящей стратегии моделирования, но и при нисходящей стратегии результатная онтология в итоге все равно должна содержать термины, которые используются в правовых актах. Это позволяет относить правовые онтологии скорее к лингвистическим, нежели к формальным.

В случае с правовым материалом, особенно при построении отраслевой онтологии, основной текстовый массив совпадает с законодательной базой данной отрасли, т.е. представляет собой четко определенный перечень правовых актов, доступ к которым обеспечен всем гражданам. Несмотря на разнообразие правовых актов и возможные недостатки юридической техники, тексты правовых актов в идеальном случае всегда должны отличаться свойствами,

которые облегчают формализацию предметной области. Среди таких свойств можно назвать следующие: нейтральность (отсутствие экспрессивности), безличный характер, связность, последовательность, точность мысли, наличие большого количества стандартизированных оборотов, сильно выраженная структурированность текста.

Во многих случаях для решения прикладных задач онтологического моделирования в правовой сфере удобно использовать некоторую общую базовую юридическую онтологию, которая разрабатывается без учета конкретной решаемой задачи и которая может быть легко расширена. В настоящий момент нет возможности говорить о существовании такой единой общеприменимой правовой онтологии для российской правовой системы, в то время как опыт зарубежных исследователей показывает, что создание такой онтологии не только возможно, но и необходимо. Такая онтология могла бы быть взята за основу для разработки практически применимых онтологий для решения задач в конкретных отраслях российского права.

Тем не менее, представляется нецелесообразным проводить всю работу по формализации правовой сферы без использования данных других онтологических проектов. Этих проектов несколько, но наиболее привлекательной для интеграции и применимой для данной задачи является базовая правовая онтология LKIF-Core (LKIF, Legal Knowledge Interchange Format – Формат обмена правовыми знаниями [9]), которая разработана и продолжает совершенствоваться в рамках проекта ESTRELLA в Университете Амстердама. Отличие этой онтологии от других заключается в том, что в ней правовые концепции моделируются на основе объективной реальности (прежде всего, физической) и основываются на объектах и процессах реального мира. Для других проектов характерной чертой является склонность к рассмотрению правовых явлений в отдельности от реального мира, в некоем идеальном состоянии.

Таким образом, актуальной является следующая задача: на основе базовой юридической онтологии LKIF-Core создать юридическую онтологию верхнего уровня применительно к системе российского права с использованием русскоязычных концепций. Такая задача ни-

каким образом не может быть сведена к последовательному переводу концептов с одного языка на другой из-за большого количества специфических особенностей российской системы права.

2. РЕАЛИЗАЦИЯ ОНТОЛОГИЧЕСКОЙ МОДЕЛИ

Реализация модели невозможна без выбора соответствующих средств моделирования и определения стратегии разработки.

Выбор средства моделирования не представляет особой сложности, так как родительская онтология LKIF-Core разрабатывается в среде Protégé, а разрабатываемая онтология является естественным продолжением родительской. Таким образом, используется редактор онтологий Protégé 3.4 (Frames). Также для построения диаграмм чрезвычайно удобно использовать плагин Ontoviz на основе генератора диаграмм Graphviz 2.22, который предоставляет широкие возможности для визуализации с использованием различных критериев на основе уже созданного дерева классов и отношений.

Что касается стратегии моделирования, то в используемой методологии *Ontology Development 101* [2], которая разрабатывалась авторами редактора Protégé, существует несколько возможных стратегий для разработки иерархии классов:

- 1) нисходящая;
- 2) восходящая;
- 3) комбинированная.

Для данного проекта целесообразно использование нисходящей стратегии моделирования. Если при проработке дерева классов возникает противоречие между правовыми системами, то данную ветвь необходимо разработать заново с использованием терминов российской теории государства и права, с постепенным переходом от более общих понятий к частным. Необходимо отметить, что данная стратегия является однозначно эффективной только в данном случае, а в случае расширения базовой онтологии до онтологии отрасли права стратегия моделирования меняется с нисходящей до комбинированной, или даже восходящей, поскольку в разработке отраслевой онтологии должна начинаться с анализа конкретных правовых текстов, в первую очередь – с анализа кодифицированного правового акта при его наличии.

3. АНАЛИЗ ПОЛУЧЕННОГО ВАРИАНТА БАЗОВОЙ ПРАВОВОЙ ОНТОЛОГИИ

На сегодняшний день удалось в достаточной степени проработать вариант базовой правовой онтологии для системы российского права, имеющий 8 уровней иерархии и включающий 127 классов и 108 отношений (рис. 1).

Редактор Protégé предполагает следующие виды визуализаций для разработанной онтологии. Концепты (термины) онтологии представлены в дереве классов (рис. 2). Отношения (слоты) концептов также имеют иерархическую структуру и представлены в виде дерева (рис. 3).

При адаптации онтологии LKIF-Core для российской правовой системы были заимствованы основные абстрактные концепции: часть-

целое, пространственно-временные отношения, классификация материальных объектов. Из правовых инструментов использована система правовой квалификации с помощью таких сущностей, как «Суждение», «Отношение_к_Суждению», «Квалификация», «Квалифицированный».

Наиболее явные различия в правовых системах были обнаружены при детализации таких концептов, как «Источник» и «Субъект». Наследники этих классов определены исключительно российской теорией государства и права.

Рассмотрим наиболее интересные фрагменты онтологической структуры. Во-первых, это визуализация отношений между объектами, суждениями и отношениями к суждениям (рис. 4). Суждение, помимо высказывающе-

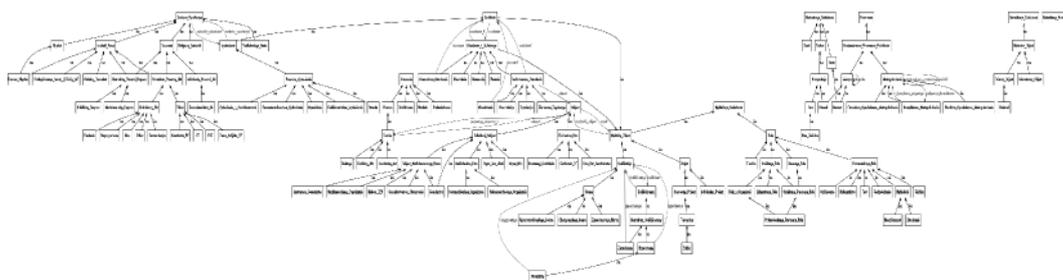


Рис. 1. Общий вид дерева классов полученной онтологии

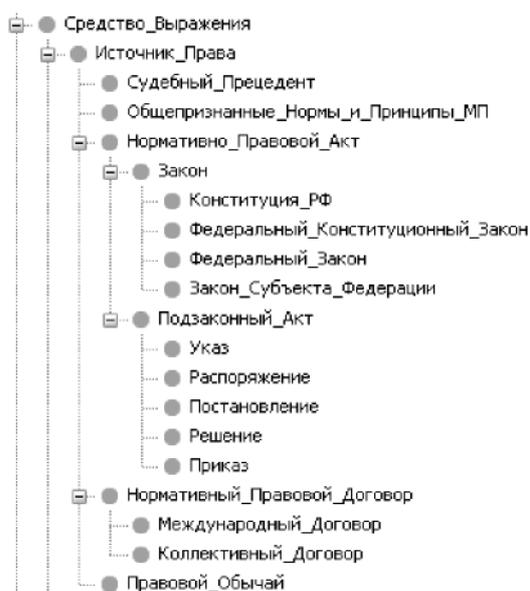


Рис. 2. Фрагмент дерева классов (терминов)

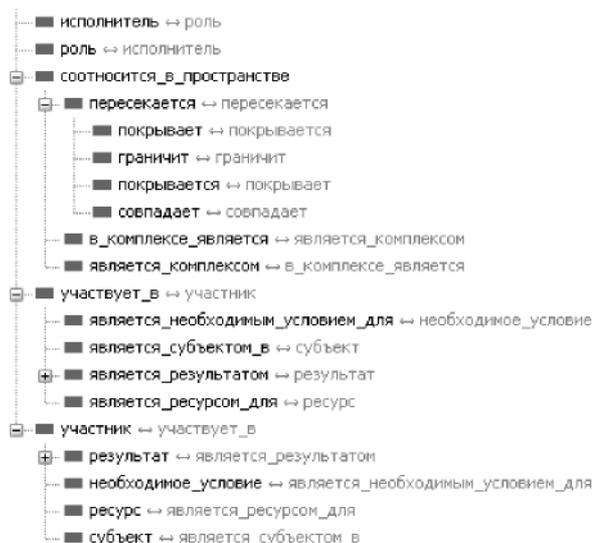


Рис. 3. Фрагмент дерева отношений (слотов)

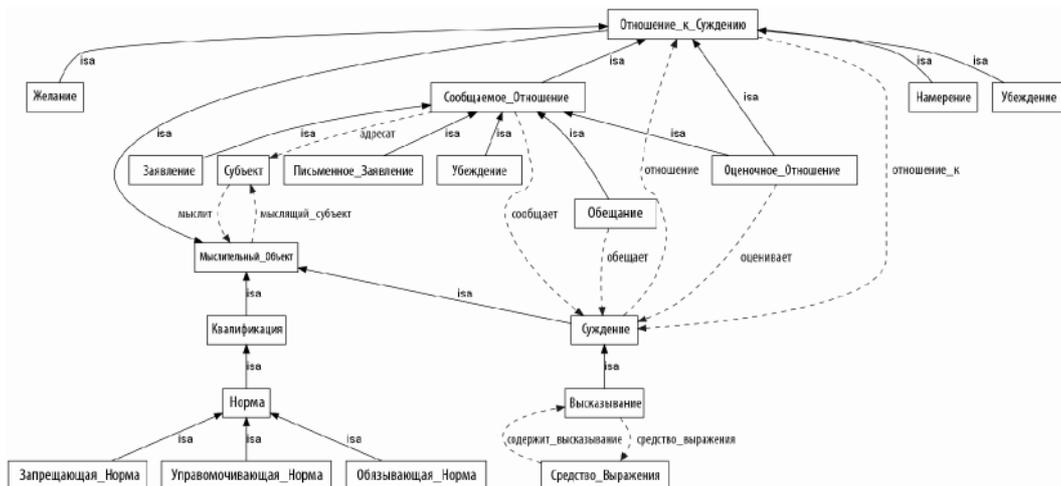


Рис. 4. Отношения между объектами, суждениями и отношениями к суждениям

го его субъекта и способа выражения, также имеет такой важный компонент, как Отношение_к_Суждению, который отображает внутреннее отношение субъекта к суждению, причем это отношение может также выражаться вовне (Сообщаемое_Отношение).

Во-вторых, это визуализация механизма квалификации (рис. 5). Предлагается использовать два больших класса, один – это собственно Квалификация, подвидом которой является Норма, второй – некий квалифицированный объект, явление или действие (Квалифицированный).

Среди проблем, которые были выявлены в процессе разработки, наиболее существенными можно назвать следующие:

- 1) различия в правовых системах и, как следствие, в правовом понятийном аппарате;
- 2) многозначность некоторых терминов и синонимия, как в английском, так и в русском языках, например: {statement, expression, declaration, proposition} и {выражение, суждение, утверждение, высказывание};
- 3) проблемы при обозначении отношений и особенно обратных отношений: широкое употребление пассивного залога глаголов в английском языке; переходность русских глаголов.

ЗАКЛЮЧЕНИЕ

Направление онтологического моделирования является актуальной задачей, в рамках которой большой интерес представляет полученный вариант базовой правовой онтологии

для российской системы права. Результат адаптации представляет собой важный стартовый материал для дальнейшей работы. Кроме того, удалось выявить проблемы адаптации онтологий для другого естественного языка и другой правовой системы.

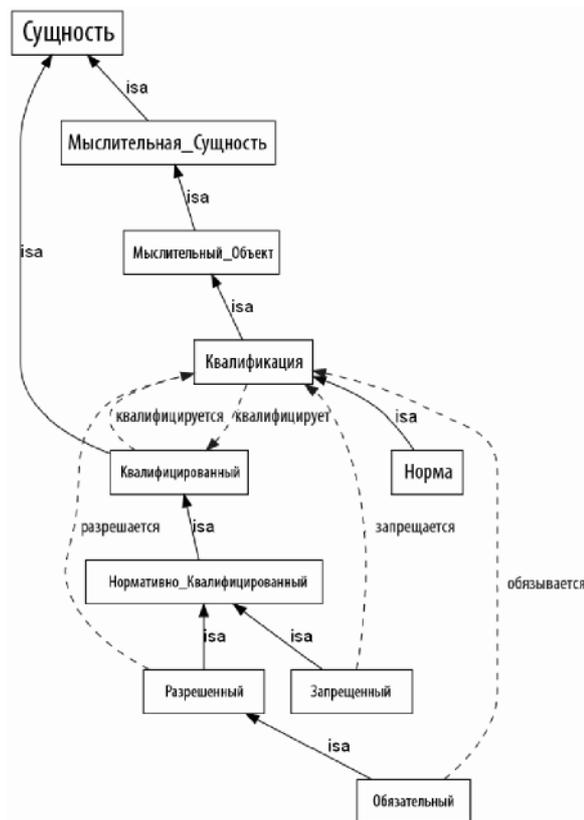


Рис 5. Механизм квалификации

В настоящее время ведется работа над расширением полученной русскоязычной онтологии для работы в конкретной предметной области, то есть более глубокое онтологическое исследование отдельной отрасли права и использование разработанных онтологий для решения прикладных задач при разработке правовых систем, основанных на знаниях.

СПИСОК ЛИТЕРАТУРЫ

1. Соловьев В. Д. Онтологии и тезаурусы / В. Д. Соловьев, Б. В. Добров, В. В. Иванов, Н. В. Лукашевич – Казань: Казанский государственный университет, 2006. – 157 с.

2. *Noy, N., McGuinness, D.* Ontology Development 101. – Stanford Knowledge Systems Laboratory Technical Report KSL-01-05, 2001.

3. IDEF5 Method Report. – Knowledge Based Systems, Inc. for Information Integration for Concurrent Engineering. – 1994.

Воронина Ирина Евгеньевна – к.т.н., доцент кафедры ПОиАИС, факультет прикладной математики информатики и механики, Воронежский государственный университет. Тел. 208-337. E-mail: voronina@amm.vsu.ru

Пигалкова Евгения Александровна – студентка 5 курса факультета ПММ, кафедра программного обеспечения и администрирования информационных систем, специальность «Прикладная информатика в юриспруденции»

4. *Fernández-López, M., Gómez-Pérez, A.* Overview and Analysis of methodologies for building ontologies. Knowledge Engineering Review (KER). Vol. 17[2]. – 2002.

5. DILIGENT Ontology Engineering. – (<http://www.aifb.uni-karlsruhe.de/WBS/cte/ontologyengineering/diligent.htm>).

6. NeOn Project. – (<http://www.neon-project.org>).

7. УИС РОССИЯ. Технология автоматической обработки текстов. – (<http://uisrussia.msu.ru/docs/ips/n/techno/index.htm>).

8. *Деринг Ю.* Онтологии в правовой сфере. / Ю. Деринг, М. Хербергер // Научно-правовой журнал «Проблемы правовой информатизации» – 2006. – № 2(12) .

9. LKIF-Core Ontology. A core ontology of basic legal concepts. – (<http://www.estrellaproject.org/lkif-core>).

Voronina I. Ye. – Candidate of Technical Sciences, Associate Professor, the dept. of Software and Information System Administration, Voronezh State University. Tel. 208-337. E-mail: voronina@amm.vsu.ru

Pigalkova E.. A. – 5th year student of Department of Applied Mathematics, Computer Science & Mechanics, Computer Science in Law Program