



УДК 004.432.4

ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ПОДДЕРЖКИ ЗАПРОСОВ НАЛОГОВЫХ ИНФОРМАЦИЙ

Турмаганбетова Ш.К.

*Евразийский национальный университет имени Л.Н.Гумилева,
г. Астана, Республика Казахстан*

sh_kurmangali@e-mail.address

В рамках диссертационной работы определены основные понятия создаваемой онтологической модели, которая позволит организовать процесс реализации экспертной системы налоговой системы в Республике Казахстан. Таким образом, в рамках данной работы будет предложены три типа моделей, основанных на онтологии, которые структурируют и организуют информацию для поддержки информационного доступа. Извлеченная информация разработана для навигации информационного поиска пользователей, но не ограничения возможностей пользователей.

Ключевые слова: онтологическая модель, SemanticWeb, поиск налоговой информации.

Введение

Информационный портал - это потенциальная важная техника извлечения текстовой информации из базы знаний. Однако, задача информационного поиска сложна, и легко потеряться в сложной сети узлов, представляющих разные понятия. В этом проекте мы рассмотрим, как возможность исследования, может быть поддержана знанием, основанным на онтологии. Упомянутые выше неудобства избегаются, предоставлением услуг исследовательским задачам пользователей, в частности, если они не являются экспертами в этой области. Мы предлагаем информационным ищущим и исследователям, способ преследования информационных объектов, связанных с их задачами и применения действий к задаче для достижения первоначальных целей и намерений. Были предложены три типа моделей основанных на онтологии, которые структурируют и организуют информацию для поддержки информационного доступа. Извлеченная информация разработана для навигации информационного поиска пользователей, но не ограничения возможностей пользователей. Кроме того, мы обсуждаем обеспечение действий к задачам пользователей в целях поддержания повторного использования.

1. Назначение онтологии

Онтология является системой понятий некоторой предметной области. Назначение

словаря онтологии - предоставить терминологию для описания знаний о конкретной области.

Онтология структурирует понятия словаря. В самом простом случае онтология описывает иерархию связанных представлений. В более сложных случаях добавляются подходящие аксиомы (или ограничения), выражающие связи между представлениями и ограничивающие их интерпретацию.

Как правило, при описании знаний представления используются для обозначения типовых множеств объектов, обладающих общими признаками, или множеств объектов, существующих в данной предметной области, но не являющихся типовыми. Типовые множества объектов организуются в классы объектов, каждый из которых характеризуется набором типовых свойств (или атрибутами класса). Допустимая область применения знаний задаётся посредством аксиом или ограничений. Перечисленные составляющие - классы, атрибуты, отношения, аксиомы/ограничения - рассматриваются как элементы онтологии.

Онтология является составной частью базы знаний, которая, помимо онтологии, включает в себя логику (или правила вывода), а также может содержать неструктурируемую (или неформализованную) информацию, выраженную средствами естественного языка.

Онтологии позволяют оригинально построить web-ресурсы на основе формальных описаний тех областей знаний, что составляют суть контента ресурса.

Факт создания информационного портала для поддержки исследовательской (браузерной) функции может быть оспорим. Однако, поиск может оказаться задачей отнимающей много времени, как только число потенциально соответствующих страниц становится больше. Исследующие пользователи могут испытывать затруднения в адресации своих потребностей, если они не эксперты в этой области, например, просматривая информацию о налогообложении. При таких обстоятельствах пользователь может захотеть организовать понятия в кластер и обусловить поперечные связанные ссылки, чтобы выполнить определенную задачу и исследовать нужную информацию. Поэтому, требуется, чтобы пользователь нашел нужную информацию в пределах информационной структуры.

Однако, пользователям трудно сгруппировать понятия, если они не эксперты в интересующей области. Один из самых часто возникающих проблем у пользователей является сомнительная информация, встречающаяся в многократных источниках, использующая похожую терминологию. Исследование показало, что вероятность совпадения терминов в домене и словарного запаса пользователя равна <0.2 (Furnasetal, 1987).

Чтобы рассмотреть эти проблемы по отдельности, была предпринята попытка моделирования поведения исследования, сопоставляя исследовательские задачи пользователей с необходимой информацией, требуемой для выполнения этих задач. Поэтому извлеченная информация может быть разработана так, чтобы она могла вести пользователей и избежать путаницы в сложной сети узлов, представляющих понятия. К примеру, если пользователь исследует информацию о «налоге», необходимый вход, требуемый для такой информации, может быть поддержан, показывая ставку налога, назначения налога, стоимость и требуемое время для оплаты налога. Если пользователь исследует информацию «о земельном налогу», то необходимые документы и условия подачи заявки могут предоставить пользователям необходимую информацию о назначении и периодичности налога, таким образом, осуществляя поисковую задачу.

Поскольку Сеть экспоненциально расширяется в размере, недостатки в организации затрудняют эффективный подбор знания из Сети. Для того, чтобы обеспечить, средства навигации, в информации исследовательских задач пользователей в Сети, важно обеспечить семантическую структуру информационной окружающей среды. Онтология была предложена для использования в различных областях: системы, основанные на знаниях, информационный поиск, стандартизация, и т.д. Онтология была рассмотрена как общее организационное устройство для спецификации концептуального знания. Онтология

предоставляет базу для ясного раскрытия понятия во всех отношениях. Высшие категории могут быть полезным гидом в разложении сложной информационной проблемы и в формировании доступного и эффективного метода доступа к информации. Поскольку потребность в использовании больше чем одного типа онтологии была вынужденной из-за возрастающего размещения информации, мы предложили три типа моделей, основанных на онтологии для классификации и поисковых целей.

2. Модели, основанные на онтологии

Наши усилия были сосредоточены на использовании онтологии для поддержания информационных потребностей в поиске информации о правительстве. Сценарии, возникающие в приложениях, мотивируют развитие онтологии. Большинство информационных порталов интересуется тремя типами информации, исследуя правительственное содержание: организационная структура, бизнес функции, и деятельность человека. В этом проекте мы сконцентрируемся на определении организационной структуры и предпринимательской деятельности большинства правительственных организаций.

Два типа пользователей вовлечены в развитие компьютерной системы, Интранета: Инженеры Онтологии/Менеджеры Отчетов и конечные пользователи. Главная ответственность Инженеров Онтологии/Менеджеров Отчетов состоит в том, чтобы обеспечить контроль и последовательность в словарном запасе, используемого для классификации и индексирования отчетов. Классификация требует понимания контекста и содержания отчета, знания существующих отчетов в системе и индексация, которая включает прикрепление одного или более поисковых слов, к каждому записанному отчету, такому как файлу печатного или электронного документа. Это облегчает непринужденность поиска конечных пользователей и извлечения отчетов. Основанные на информационных потребностях, три типа моделей онтологии предложены: организационная онтология, онтология действий домена и пользовательская онтология действий. Организационные методы и методы деятельности домена предназначены для индексирования изаголовков, тогда как пользовательская онтология деятельности облегчает поиск и извлечение информации.

2.1. Организационная онтология

Организационная онтология предоставляет семантическую информацию правительственной структуры. Сложная структура часто решается разложением структур в иерархические единицы в компьютерной системе. Рисунок 1 показывает иерархическое дерево организационной онтологии правительства. Дерево представляет понятия

предприятий. Например, правительство состоит из государственных и местных департаментов. Государственные департаменты включают в себя ЗАКОН, Республики Казахстан и нормативные акты и т.д. Местные департаменты включают в себя Генерального прокурора Департамента, Министерство обороны, Министерство здравоохранения и помощи семьям, и т.д. Каждое предприятие связывается с соответствующими отношениями и признаками. Например, определение налог – ставка налога, и его функция определяется термином «назначается».

Организационная онтология в состоянии ответить – вопросы, базируемые на сложное знание, такие как нахождение «нарушение налогоплательщика». Поиск в существующей сети идентифицирует тысячи страниц о налогоплательщике в налогах на слово «нарушение налогоплательщика». В этом примере онтология включала бы такой признак как «фамилия», классификацию как «человек», и связь со словом «налогоплательщик».



Рисунок 1 – Иерархическое дерево организационной онтологии правительства

2.2. Онтология действий домена

Онтология действий домена разработана для организации и структурирования информационной функции и действия в домене. В отличие от традиционной классификации библиотек, где книги классифицированы согласно предметам, онтология действий домена предлагает иерархическую структуру для классификации отчетов согласно информационным функциям и действиям, зарегистрированные в отчетах для классификации и индексации задач. Таким образом, последовательность информационной действий может сопровождаться группой запросов, которые были классифицированы вместе.

Развитие онтологии действий домена основано на CSIROKeywordAAA (CSIRO 1996), который является алфавитным списком терминов.

Роль онтологии действий домена - это определение запроса по их отношениями с

деятельностью, которую они документируют. Поэтому, отчеты, классифицированные и индексированные при помощи онтологии действий домена, будут иметь значение и будут размещены в пределах контекста, в котором они были созданы. Это также означает, что все файлы на одной функции или действиях группируются вместе в индексе.

2.3. Пользовательская онтология действий

Пользовательская онтология действий связана с информационными действиями поиска, выполненными конечными пользователями. Задача информационного поиска сложна. С точки зрения анализа основанного на знания, мы должны рассмотреть, каким пользователям нужно знать об информационных объектах и действиях, вовлеченных в информационную задачу поиска, и как это знание организовано. Действия пользователя обычно могут быть описаны, выполнением действий задачи информационными объектами, связанными с информационными потребностями. Например, пользователь ищет информацию в музыкальной библиотеке. Он или она выполняет действия, такие как вход в коллекцию, поиск индекса к полке и слушанию песни.

Развитие пользовательской онтологии действий начинается с построения таксономической классификации задачи и информационных объектов. Знание задачи включает словарный запас для процесса выполнения действий, как, вход, поиск и т.д. Рисунок 2 показывает концептуальную модель для пользовательской онтологии действий. Задача - это информационные действия поиска пользователей. Родовой глагол или фразовые глаголы используются для представления словарного запаса действия задачи, которое независимо от домена и может быть составлен вручную.

Сборник родовых глаголов используется, чтобы описать информационные задачи поиска пользователя, такие как:

- Доступ к: достижению, использованию или приближению к чему-то.
- Обратиться: формально попросить что-то.
- Иммигрировать: стать поселенцем, не туристом или посетителем.
- Дом: поместить (заявление, и т.д.) надлежащими властями.
- Поиск: искать, попытаться найти что-то.

Информационный объект представляет ряд объектов, связанных с информацией. С точки зрения конечного пользователя, они могут быть незнакомы с организационной структурой и терминологией, используемой для бизнес функций и действий домена. Поэтому, информационные объекты включают в себя высокоуровневое описание организационной онтологии и онтологии действий домена, например, они только включают

ключевые слова и описатели действия в онтологии действий домена. Список существительных или фразовых существительных используется, чтобы описать информационные объект, таких как:

- Налог (например, налог на печать, земельный налог, дебаты, налог, и т.д.)

Комбинация родового глагола и существительного, как «глагол + существительное» может быть предпосылкой на процесс решения общей проблемы. Такая комбинация облегчает процесс информационного поиска.

Комбинация задачи и информационных объектов относится к поисковым действиям пользователя, как:

- Налог Заявления (например, налог на печать, земельный налог, налог дебетов, и т.д.)

Вышеупомянутые подчеркивающие линии указывают на родовые существительные информационного объекта.



Рисунок 2 – Концептуальная модель онтологии действий пользователя

Заключение

Мы предложили два типа моделей онтологии, которые структурируют информацию путем поддержки средств навигации при доступе к информации. Мы показали, что исследовательский просмотр может быть объединен с основанным на онтологии знанием во избежание потери в киберпространстве, в особенности, если пользователи не являются экспертом в домене. Кроме того, мы рассмотрели проведение действий задачи пользователей, чтобы поддержать информационное повторное использование.

В рамках нашей будущей работы, важное рассмотрение информационной навигации явилось вызовом, поставленным Познавательным Движением. Познание основано на идее, что использование знания, полностью зависит от ситуации. Мы признаем, что вопросы навигации информации должны быть исследованы варьированием ситуаций, таким образом, мы

смогли бы отследить необходимые изменения в онтологических моделях. Это также говорит о том, что мы нуждаемся в системах, которые поддерживали бы, необходимые изменения в моделях онтологии, поскольку ситуации изменяются. Необходимо связать модели онтологии к расположенной парадигме познания.

Библиографический список

- [Berners-Lee, 2001] T., Hendler, J., Lassila, O.] The Semantic Web // Scientific American, May 17, 2001.
 [Stewart S.] Karlinsky and Daniel E. O'Leary. Tax-based expert systems: a first principles approach. Expert Systems in [Finance D.E. 1992] O'Leary and P.R. Watkins (Editors) 1992 Elsevier Science Publishers B.V.

ONTOLOGY MODELS FOR SUPPORTING EXPLORATORY INFORMATION NEEDS

Turmaganbetova Sh.K.

Евразийский национальный университет имени Л.Н.Гумилева,

г. Астана, Республика Казахстан

sh_kurmangali@ e-mail.address

Conclusion

We have provided three types of ontology models that structure information in a way to support navigation aids for information access. We have shown that exploratory browsing can be coupled with ontology-based knowledge to avoid getting lost in a not experts in the domain. In addition, we have discussed the main tenancy of users' task actions to support information reuse purposes.