

ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ INTERNET РЕСУРСОВ ПОСВЯЩЕННЫХ ИСТОРИЧЕСКИМ ИССЛЕДОВАНИЯМ

Ontological model internet resources dedicated historical research

Ручкин Владимир Николаевич, доктор технических наук, профессор кафедры информатики и ВТ ФГБОУ ВПО «Рязанский государственный университет имени С.А. Есенина».

 v.ruchkin@rsu.edu.ru

Махмудов Марат Наильевич, кандидат физико-математических наук, доцент, проректор по научной работе ФГБОУ ВПО «Рязанский государственный университет имени С.А. Есенина».

 m.mahmudov@rsu.edu.ru

Ручкина Екатерина Владимировна, кандидат исторических наук, доцент кафедры философии и истории ФГБОУ ВПО «Рязанский государственный радиотехнический университет».

 ek-ruchkina@yandex.ru

Фулин Владимир Андреевич, начальник центра дистанционного обучения и мониторинга качества образования ФГБОУ ВПО «Рязанский государственный университет имени С.А. Есенина».

 v.fulin@rsu.edu.ru

В статье рассматриваются возможности онтологического представления концептуального моделирования ресурсов интернет для исторических исследований.

The possibility of an ontological representation of conceptual modeling online resources for historical research.

Ключевые слова: **электронная библиотека, онтологии, обработки структур данных, анархизм, история, Moodle, Protégé.**

Keywords: **e-library, ontology, processing data structures, anarchism, history, Moodle, Protégé.**

Бурное развитие информационных технологий привело к росту образовательной информации окружающей современного человека. Это, в свою очередь, приводит к необходимости повышения эффективности использования образовательных ресурсов, поиска новых способов хранения, представления, систематизации и структуризации, а также автоматической обработки знаний. При этом, речь идет не только о ресурсах взаимосвязанных компьютерных систем, а об особых информационных пространствах сети Интернет.

Телекоммуникационные средства значительно увеличивают диапазон и расширяют границы любого исследования, в том числе и исторического. Возможности Интернет, простота создания web-сайтов, электронных архивов и библиотек, а также форумов и порталов посредством различных языков разметки (HTML, XML, JavaScript, PHP и др.) позволяют осуществлять телекоммуникационный доступ, дают возможность подключения, по-

иска и сбора данных из архивов и библиотек по всему миру. Систематизация полученной информации необходима для дальнейшего исследования и анализа. В результате появляется возможность «путешествовать» по всемирной сети, не обращая внимания на физическое нахождение интересующих ресурсов, и получать к ним доступ. Однако при возникновении и постоянном увеличении спроса необходимо решить проблему расширяющегося предложения, т. е. обеспечить структуризацию ресурсов с целью быстрого, достоверного и надежного доступа к ним. Для решения такого рода проблем авторами предлагается использовать онтологический метод описания ресурсов вследствие его абстрактного-обобщающего свойства и достояния для различных предметных областей.

Для удобства поиска требуемой информации в глобальной сети разработаны специальные механизмы, называемые поисковыми системами. Они помогают быстро производить поиск нужной информации среди экспоненциального роста объема информационных ресурсов.

Онтология представляет собой формальное описание концептуального моделирования конкретной предметной области, представленное в удобной форме, как исследователю, так и компьютерной системе в Интернет.

Потребность в разработке онтологий объясняется следующими причинами [4]:

1. Совместное использование исследователями или программными агентами общего понимания структуры информации;
2. Разработка и управление терминологией;
3. Возможность повторного использования знаний в предметной области;
4. Получение надежного семантического базиса в определении содержания;
5. Отделение знаний в предметной области от оперативных знаний;
6. Получение логической теории, состоящей из словаря и набора утверждений на некотором языке логики, что позволяет на основе этой теории получать вывод новых знаний, явно не заложенных в онтологии;
7. Возможность использования онтологий для поддержки функционирования и роста нового вида цифровых библиотек, реализованных как распределенные интеллектуальные системы.

Предлагаемая авторами онтология историческое «исследование» содержит описание универсальных и специализированных ресурсов Интернет, необходимых в процессе исследования (рис.1). Все онтологии были встроены в СДО Moodle, их можно просматривать как рисунки графов, а также работать с ними, используя редактор онтологий Protégé.

Справочники, или тематические каталоги, представляют собой базу данных, рассчитанную на ручной ввод данных. Для регистрации в справочнике автору необходимо послать заявку с указанием адреса URL, в какой раздел следует поместить страницу, краткое описание и список ключевых слов для поиска страницы в справочнике. Через некоторое время заявка будет рассмотрена редактором или модератором. Модераторы справочников оставляют за собой право изменить описание страницы или ключевые слова, поместить страницу в другой раздел, который, по их мнению, более подходит для нее, или вообще отказать в регистрации. Очень редко справочник имеет программу, которая способна считать документ из сайта, хотя некоторые справочники имеют простой «робот», который контролирует достоверность сообщенного адреса URL. Таким образом, справочник напоминает тематический указатель в библиотеке, позволяя найти наиболее значимые документы по заданной теме - путеводитель. Считается, что результаты поиска в справочниках оказываются более эффективными, чем в машинах поиска. Самый известный зарубежный справочник – Yahoo! (<http://www.yahoo.com/>) – состоит из более 20 000 тематических рубрик.

Что касается источников по истории анархизма, необходимо упомянуть, что наиболее важное значение для ученого-историка имеют архивы. Только просмотрев, исследовав и проанализировав значительное количество исторических документов исследователь может наиболее полно представить себе картину исторической действительности и сделать действительно верный вывод о причинах, следствиях и результатов того или иного исторического события. Зная структурные связи между отдельными элементами особенно важно и необходимо на основании полученных данных и знаний выявить их фальсификацию, т.е. осуществить своеобразную защиту информации и не позволить внести сомнительные данные. Например, 1-й Украинский, 4-й Украинский и другие фронты формировались по территориальному признаку, а не по национальному. А в составе фронтов воевали все национальности. Так, например, в состав Украинского фронта входила польская дивизия имени Костюшко под командованием Ярузельского, сформированная на Рязанской земле. Поэтому многие центральные архивы подключены к сети Интернет и имеют в распоряжении не только электронный каталог, но и электронные библиотеки документов. Пользование такими библиотеками ведется через электронную почту. Электронная почта (e-mail) является в данном случае не просто средством общения с другими людьми, но и универсальным инструментом доступа к базам данных и пересылки файлов [3].

Возможности использования архивов сети Интернет можно рассмотреть на примере историографии и источниковой базы такого исторического явления, как анархизм. Как оказалось, теории и практике анархизма в Интернет посвящено множество ресурсов и можно произвести достаточно глубокое исследование.

Наиболее важной и значимой категорией ресурсов InterNet являются, без сомнения, web-сайты архивов и библиотек, которые облегчают поиск исторических источников.

Самым крупным web-центром по поиску архивных документов является портал Федерального архивного агентства «Архивы России»[5]. На этом портале можно найти адреса и телефоны как центральных, так и региональных российских архивов, информацию о конференциях и семинарах, ознакомиться с изданиями и публикациями. На главной странице портала есть ссылки на сайты центральных архивов, откуда можно продолжить поиск по электронным описям документов. Наиболее полезными для исследователя будут раздел архивных справочников, каталог тематических баз данных федеральных и региональных архивов, а также поисковая система, позволяющая искать нужные фонды в электронном Центральном фондовом каталоге. Для данной системы существует возможность расширенного поиска, в которой можно выбрать для поиска конкретный архив.

В файлах анонимных ftp-архивов (обычно библиотеки таких архивных файлов находятся по адресам крупнейших университетов США и Западной Европы) существуют основные работы и статьи крупнейших исследователей, работающих в этих университетах по исследованию данной проблематики. Например, в электронной библиотеке Оксфордского университета размещены и доступны для скачивания либо в HTML-версии, либо в формате .pdf исследования профессора университета Джеймса Джолла [6]. Они посвящены истокам, происхождению и развитию анархизма в Европе и в России, психологии людей, увлеченных этой идеей, анализу феномена анархической «пропаганды действием», начатой в России приверженцами М.А. Бакунина и с энтузиазмом воспринятой европейскими анархистами, а также взаимодействию и взаимосвязи российского и зарубежного анархизма.

Существенный объем информации предоставляют также исследовательские онлайн-центры, существующие как на базе высших учебных заведений, так и поддерживаемые группами энтузиастов-единомышленников. Среди первых нужно отметить самодокументируемый исследовательский онлайн-центр «Архивы анархии» (“Anarchy Archives”), который создал в 1995 году профессор Питцеровского Колледжа Дейна Уард и в пополнении которого документами участвуют студенты нескольких колледжей [7].

Цель этого ресурса – собрать на одном сайте собрания сочинений самых известных анархистов и историю анархизма. На сайте размещены как труды анархистов, так и статьи исследователей, посвященные культурному влиянию анархизма на общество.

Так, на сайте «Архивы анархии» размещены мемуары американских анархистов Эммы Гольдман и Александра Беркмана уроженцев России и американских эмигрантов, высланных из США обратно в Россию в 1919 году. Во время гражданской войны они находились в центре российского анархического движения, поддерживали связь с московскими и пет-

роградскими анархистами и членами конфедерации «Набат», поддерживали связь с Н.И. Махно. Эти работы вышли в середине 1920-х гг. и в России не издавались, поэтому русскому читателю эти работы доступны только как англоязычный ресурс сети Internet.

Второй по значимости категорией Internet-ресурсов, по нашему мнению, являются неофициальные web-сайты и порталы, поддерживаемые исследователями-энтузиастами. Среди ресурсов данной категории заслуживает особого упоминания официальный сайт Н.И. Махно, организатор которого, Сергей Шведов, проделал большую работу по сбору документов и фотографий, касающихся махновского повстанческого движения и сведений о махновцах. Создатель сайта постоянно отслеживает публикации о Н.И. Махно и его соратниках и ведет регулярную рассылку для подписчиков [8].

С этой же точки зрения интересен сайт «Наш Баку», посвященный истории города и горожан. Его раздел «История города» содержит подраздел «Политические партии и движения», в котором размещена статья об анархистах в Баку, приведены газетные сообщения, касающиеся анархистов. Эти сообщения, представленные с грамотными ссылками на газеты, откуда они взяты, представляют большую ценность для историка [9].

Обмену мнениями способствуют также форумы, которые дают ученым возможность обмениваться мнениями. Как правило, крупные сайты и порталы имеют такие форумы, цель которых – обеспечить удобное общение и эффективную обратную связь. В нашем случае, это такие крупные форумы, как «Единый форум анархистов» [10] и форум сайта «Российские социалисты и анархисты после октября 1917 года» [11].

Третьей по релевантности категорией являются сайты анархических организаций – «Автономное действие», Конфедерация революционных анархо-синдикалистов (КРАС), Межпрофессионального союза трудящихся (МПСТ), объединенный сайт анархических организаций «Bakunista», революционной конфедерации анархистов-синдикалистов (РКАС). Они представляют определенный интерес для историков, потому что наряду с очерками по истории анархизма, там публикуются переводы статей по истории анархизма с других языков, а также ведутся дискуссии по современному состоянию анархизма, что позволяет проследить развитие этого политического направления.

Так, организация «Автономное действие» является международной. Поэтому основной ее сайт располагает обширной библиотекой, состоящей не только из трудов теоретиков анархизма, но и из работ современных исследователей. Часто эти работы публиковались очень маленькими тиражами и доступны широкому кругу читателей только в электронном виде. Среди таких работ – книга «Пепел наших костров», автором которой является Моше Гончарок, вышедшая в Иерусалиме в 2002 г [12].

Кроме того, сайты анархических организаций интересны тем, что на их страницах размещены материалы по истории зарубежного анархизма. Так, на сайте КРАС размещен очерк по истории анархизма в Мексике,

в библиотеке сайта «Bakunista» находится статья по истории анархизма в Бразилии. Тем не менее, в политической жизни этих стран анархические взгляды традиционно получили большую популярность и широкое распространение. Часто среди материалов на таких сайтах можно встретить редкую информацию, зачастую недоступную широкому кругу исследователей. В качестве примера можно назвать очерк Марко Вуковича, посвященный югославскому анархизму и подготовленный им на основе личных воспоминаний и работ Триво Инджича и Горана Ивановича. Работа носит название «Очерк истории анархизма в бывшей Югославии». [13]

Подобные web-сайты и порталы являются хорошим подспорьем как начинающим исследователям, так и ученым со стажем, потому что позволяют им следить за научными новинками.

Иногда ценные материалы можно найти в так называемых «живых журналах», или блогах. Обычно информация, приводимая в блогах, требует самой тщательной проверки, поэтому исследователю необходимо критично относиться к приводимой информации и выбирать только те блоги, которые заслуживают доверия.

Из выше представленного материала выделим следующие функциональные блоки:

- *форумы*: «Единый форум анархистов» «Российские социалисты и анархисты после октября 1917 года» и др.;
- *архивы*, включающие в себя архивные справочники; архивы России: порталы, конференции, семинары, издания, публикации; центральный архив; FTP архивы: ведущих университетов Европы, США и Канады, колледжа Дейна Уард; онлайн центры и т.д.;
- *библиотеки*, включающие в себя разнообразные прикладные ресурсы справочного характера: сетевые версии словарей, справочников, энциклопедий и другие информационные ресурсы образовательной направленности).
- *справочники*, содержащего систему контрольно-тестовых заданий, имеющих своей целью осуществление мониторинга учебной деятельности каждого учащегося;
- *FTP архивы* - включающие в себя ведущие университеты Европы, США и Канады, профессора Джеймса Джолла в его состав входит персональная веб-страница учащегося и неструктурированные файловые архивы с результатами учебной деятельности учащихся.
- *исследователи-энтузиасты* - в него входит описание навигации по учебному курсу и ресурсам учебного компонента, методические указания для учащегося и преподавателя по работе с ресурсами.
- *анархические организации* – содержит элементы для организации общения между преподавателями и студентами: форумы, чаты и видео-конференции.
- *живые журналы (блоги)*.

Представленный набор составляющих блоков является инвариантным для формирования учебного контента. Принципы комплектации кон-

кретного учебного курса будут зависеть от требований педагога, который структурирует курс в зависимости от целей и задач обучения.

Таким образом, в работе дается сравнительно полная классификация встроенных информационных и телекоммуникационных ресурсов современного мирового анархического движения, использование в гуманитарных дисциплинах компьютерных технологий и электронных архивов глобальной сети. Перечисленное и составляет основу онтологии. В результате получается онтологическая модель.

Следующим этапом развития является переход к представлению историческим исследованиям в виде, доступном для машинной обработки, а именно – применение семантического подхода. Суть подхода состоит в описании информационных ресурсов с помощью моделей знаний, которые должны быть стандартизованы и понятны широкому кругу профессионалов в различных предметных областях [5]. Систему понятий можно представить как дерево понятий, то есть как ориентированный граф, вершины которого – понятия, а дуги – родовидовые отношения между ними. Иерархически структурированное множество терминов, описывающих предметную область – это и есть онтология, которая может быть использована как исходная структура для базы знаний в дистанционном обучении и любом исследовании. Таким образом, онтология как модель выражает определенный взгляд на некоторую предметную область (subject area) и формально может быть описана следующим набором множеств:

$O_{SA} = \langle X, R, F \rangle$, где X – конечное множество концептов предметной области, которую представляет онтология OSR , например для предметной области «историческое исследование» такими понятиями являются форумы, архивы, FTP архивы, библиотеки, справочники, исследователи-энтузиасты, анархические организации, живые журналы (блоги) и т.д.; R – конечное множество отношений между концептами заданной предметной области. Например, «форумы»_ «архивы», «FTP архивы»_ «библиотеки», «справочники»_ «исследователи-энтузиасты» «анархические организации» «живые журналы (блоги)» и т.д. Здесь имя отношения формируется из имен связываемых этим отношением сущностей и знака подчеркивания; F – конечное множество функций интерпретации, заданных на концептах и/или отношениях онтологии O .

Построение онтологии понятийного состава знаний дает возможность для последующего использования: создания обучающей и тестирующей системы интерактивных курсов.

Использование программ структуризации, оптимизации и конвертирования онтологических описаний (Protégé, Chimaera, ODE, GrOWL) используют оригинальные подходы к структуризации и унификации онтологических описаний.

В Рязанских государственном университете имени С.А. Есенина и государственном радиотехническом университете авторами были разработаны онтологии на Open Web Language следующих курсов «Системы ис-

кусственного интеллекта», «Internet-Intranet», «Экспертные системы и базы знаний», «Культурология», «История», соответствующих требованиям ФГОС ВПО третьего поколения. Онтологии понятий курсов используются для самостоятельного изучения понятий предметной области, а для закрепления изучаемого понятия студенту предлагаются разнообразные тестовые задания с выбором уровня сложности с целью автоматического выставления рейтинговой оценки в базу данных.

Проведенные исследования и полученные результаты подтвердили возможности онтологического представления концептуального моделирования ресурсов Интернет для исторических исследований в виде форумов, библиотек, архивов, FTP архивов, ведущих университетов мира с целью создания современных интеллектуальных поисковых систем на основе OWL, обеспечивающих структуризацию ресурсов для быстрого, достоверного и надежного доступа к таким ресурсам.



БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ручкин В.Н., Фулин В.А., “Использование онтологического метода структуризации учебного контента”, Вестник ТулГУ, Технические науки, Том 6, Тула : Издательство ТулГУ, 2014, с. 65-71.
2. Каширин Д.И. Формализм ICF-онтологии для представления знаний в глобальной сети нового поколения Semantic Web (С.57-64) Вестник РГРТУ, № 22, Рязань: РГРТУ, с. 23-29.
3. J. Davies, R. Studer, P. Warren, “Semantic Web technologies: trends and research in ontology-based systems” / Davies, J. (N. John), Chichester, 2006, 312 p.
4. Ручкин В.Н., Романчук В.А., Фулин В.А. Когнитология и искусственный интеллект. Рязань, 2012.
5. Портал “Russian Archives” - <http://www.rusarchives.ru/index.shtml>
6. <http://solo.bodleian.ox.ac.uk/>
7. www.dwardmak.pitzer.edu/anarchist_archives/
8. <http://www.makhno.ru>
9. <http://www.ourbaku.com/index.php5>
10. <http://anarhia.org>
11. <http://socialist.memo.ru/>
12. <http://avtonom.org/old/lib/theory/goncharok/idish-anarchism0.html>
13. <http://anarhia.org/forum/viewtopic.php?p=306978&#p306975>