

УИС РОССИЯ: онтология предметной области «государственное управление»*

Т.Н. Юдина, А.В. Богомолова

Научно-исследовательский вычислительный центр МГУ имени М.В.Ломоносова,
Экономический факультет МГУ имени М.В.Ломоносова
yudina@mail.cir.ru, anna.bogo@gmail.com

Аннотация

Рассматриваются вопросы разработки технологии организации знаний предметной области «Государственное управление» на базе онтологии. Такой инструмент нужен для формирования информационной среды, поддерживающей задачи управления на всех уровнях власти. Работы ведутся в рамках проекта «Университетская информационная система РОССИЯ» (УИС РОССИЯ). Онтология разрабатывается как терминологический ресурс для поддержки дополнительного инструмента поиска к уже реализованным в УИС РОССИЯ инструментам (тезаурусу и рубрикаторам) и предназначена для детализированной обработки и содержательной интеграции социально-экономических данных и аналитических публикаций по проблемам государственного управления. Аналогичные разработки ведутся во многих развитых странах мира.

1. Введение

Повышение эффективности управления – актуальная проблема на всех уровнях государственной власти и местного самоуправления в России. Отметим, что задача совершенствования управления актуальна для всех стран мира. В свете концепции результативного управления (Performance Management) эффективное управление рассматривается как основное условие повышения конкурентоспособности государства и успешного развития общества. Одно из направлений в решении этой задачи – формирование новой информационной среды для поддержки процесса управления на всех этапах – изучение вопроса, подготовка предложений с учетом возможных рисков,

мониторинг и анализ результатов. Постоянно увеличивающийся объем актуальной для процесса управления информации способствовал возрастанию интереса к автоматическому анализу публикаций и разработке технологий для организации знаний предметной области «государственное управление».

Предметную область характеризует большое число источников разных типов жанра «деловая проза» – официальные документы, статистика, административные регламенты, научные публикации, статьи СМИ, данные опросов и обследований, географические сведения и многие другие. Инструмент классификации публикаций и организации знаний может быть разработан на основе онтологии предметной области «Государственное управление». В данной статье под онтологией понимается «формальное явное описание понятий в рассматриваемой предметной области, свойств каждого понятия, описывающих различные свойства и атрибуты понятия...» [1]. В последнее десятилетие исследования в области онтологий предметных областей динамично развиваются. Обзор проектов и анализ результатов дан в работе Лукашевич Н.В. «Тезаурусы в задачах информационного поиска» [5].

Примеров поисковых инструментов, реализованных на базе онтологий по предметной области «Государственное управление» нет, но исследования активно ведутся. Это направление объявлено приоритетной исследовательской и технологической задачей национальными научными фондами США и Великобритании, а также научным фондом Европейского Союза. Среди известных зарубежных проектов можно выделить, например, инициативу ONTOGOV, тезаурус областей государственного управления Австралии (Thesaurus of Australian Government Subjects, TAGS). Целью европейского проекта ONTOGOV, запущенного в 2004 году, было создание платформы для моделирования предметных онтологий органами государственной власти. Создание тезауруса областей государственного управления Австралии было инициировано в начале двухтысячных для использования государственными ведомствами с

Труды XIV Всероссийской объединенной конференции «Интернет и современное общество» (IMS-2011), Санкт-Петербург, Россия, 2011.

целью облегчения работы с публикуемыми документами [3].

В России один из проектов по исследованиям в области онтологии предметной области «государственное управление» выполняется с 2008 года по грантам РГНФ, с 2011 года по гранту на создание информационной системы для содержательной обработки данных и документов и интеграции знаний предметной области «государственное управление» средствами тематической онтологии. Работы ведутся в рамках проекта «Университетская информационная система РОССИЯ», который также поддерживается грантом РГНФ.

В УИС РОССИЯ реализован поиск по Общественно-политическому тезаурусу и двум рубрикам. [4]. Описываемая онтология создается как ресурс, на основе которого будет разработан дополнительный инструмент поиска в УИС РОССИЯ, ориентированный на предметную область «Государственное управление».

2. Рубрикатор как основа онтологии предметной области «государственное управление»

Первым этапом работы стало создание рубрикатора специалистами-экспертами в области государственного управления. Масштабность выбранной предметной области, сложности в определении ее границ стали ключевой проблемой при формировании рубрикатора. Результатом работы экспертов стало определение трех рубрик верхнего уровня, ограничивающих предметную область: функции государства и политика управления; институты власти; инструменты и методы управления. Выбор данных сфер в качестве рубрик верхнего уровня обусловлен целью создания онтологии - классификация документов и данных, описывающих область государственного управления для упрощения поиска и организации доступа к ним.

В основу рубрики «Функции государства и политика управления» положена классификация государственных функций Р. Масгрейва [2, с., 336-338], модифицированная с учетом цели создания онтологии. Следует отметить, что трехкомпонентная модель Масгрейва (аллокация, распределение и стабилизация) ориентирована на функции, реализуемые государством через бюджетную политику, в то же время для создания онтологии более релевантным является подход, базирующийся на классификации государственных функций в зависимости от сферы воздействия. Таким образом, в рубрике «Функции государства и политика управления» выделены подрубрики «Политическая функция» и «Экономическая функция», что соответствует подходу Р. Масгрейва. Помимо перечисленных, также выделены социальная, внешнеполитическая, экологическая функции и обеспечение безопасности.

Рубрика «Институты власти» включает в себя подрубрики «Институты государственного управления», «Институты судебной власти», «Институты местного самоуправления»; она основана на Общероссийском классификаторе органов государственной власти и управления [7]. Все подрубрики выстроены в соответствии с принципами разделения властей, разделения уровней власти, разделения государственного управления и местного самоуправления, а также выделения трех типов органов исполнительной власти – министерств, агентств и служб, прописанные в Конституции РФ, Указе Президента РФ «О системе и структуре федеральных органов исполнительной власти» от 09.03.2004 №314.

В рубрике «Инструменты и методы управления» выделены традиционные экономические методы управления – балансовый, нормативный и программно-целевой, которые используются для реализации различных управленческих функций – прогнозирования, планирования и контроля [6, с. 14-15].

Первая версия рубрикатора была представлена на экспертизу внешним специалистам. Состав подрубрик уточнен по результатам внешней экспертизы рубрикатора. Итоговая версия рубрикатора имеет сложную структуру и объединяет в себе как иерархический, так и фасетный подходы к построению классификации. В настоящий момент рубрикатор включает в себя 3 рубрики первого и 12 рубрик второго уровня (всего 161 рубрику).

3. Использование социально-экономических показателей для описания рубрик

На втором этапе работы было начато формирование словника и описание рубрик рубрикатора с помощью терминов. В ходе описания рубрик была выявлена дополнительная прикладная задача, чрезвычайно актуальная в свете развития концепции результативного управления (Performance Management) как в российской, так и в зарубежной управленческой практике. Результативное управление предполагает обязательную количественную оценку деятельности органов власти, что обуславливает необходимость увязки показателей, собираемых Федеральной службой государственной статистики (Росстат) и другими статистико-производящими ведомствами с государственными функциями, типами политик и институтами власти. Таким образом, в ходе работы над онтологией была выявлена необходимость классификации статистических показателей, формирования массива терминов, характеризующих данную область, и описание рубрик рубрикатора терминами статистических показателей.

Для выполнения поставленной задачи на первом шаге были изучены нормативные документы, определяющие порядок формирования единого

статистического ресурса для целей государственного управления и обязанности Росстата в части сбора показателей: Федеральный закон от 29.11.2007 №282-ФЗ «Об официальном статистическом учете и системе государственной статистики в Российской Федерации», Постановление Правительства Российской Федерации от 18.08.2008 №620 «Об условиях предоставления в обязательном порядке первичных статистических данных и административных данных субъектам официального статистического учета», Административный регламент исполнения Федеральной службой государственной статистики государственной функции «Предоставление официальной статистической информации Президенту Российской Федерации, Федеральному Собранию Российской Федерации, Правительству Российской Федерации, иным федеральным органам государственной власти, органам государственной власти субъектов Российской Федерации, органам местного самоуправления, судам, ...» (утвержден Приказом Федеральной службы государственной статистики от 18.03.2008 №61). В соответствии с этими документами, за Росстатом закреплено руководство работами по интеграции и межведомственной координации статистических работ по созданию интегрированных информационных ресурсов в рамках федерального плана статистических работ. В процессе работы были изучены публикации Росстата на официальном сайте – Центральная база статистических данных и более 70 тематических сборников. Были также изучены формы федерального статистического наблюдения Росстата.

Область статистической деятельности министерств и ведомств определяется в соответствии с предметом их ведения, но статистическое наблюдение ведется в соответствии с официальной статистической методологией и статистическим инструментарием, утвержденным Росстатом. В федеральном плане статистических работ определен перечень субъектов официального статистического учета и основные направления деятельности для отдельных субъектов. Перечислим лишь несколько ведомств:

МВД России: наличие автотранспорта, безопасность дорожного движения, состояние правопорядка, общественной безопасности;

ФМС России: миграционные процессы;

МЧС России: чрезвычайные ситуации, пожары и последствия от них;

Минобороны России: численность пенсионеров, в том числе инвалидов, состоящих на учете и получающих пенсию в Минобороны России;

Минюст России: нотариальная, адвокатская деятельность, государственная регистрация актов гражданского состояния;

Банк России: платежный баланс РФ, показатели деятельности банковской системы, показатели рынка государственных ценных бумаг, финансовые

операции страховых организаций и негосударственных пенсионных фондов.

В ходе работы был также изучен статистический инструментарий всех субъектов официального статистического учета на основе анализа массива документов Росстата (приказы и постановления по утверждению и изменению статистического инструментария для министерств и ведомств в 2005 - 2009 гг.).

Общее число выявленных статистических показателей около 2000. Показатели классифицированы по дате утверждения статистического инструментария, что позволяет отследить произошедшие изменения, как в организации наблюдения, так и в статистической методологии; по субъектам власти; типам государственных политик, инструментам управления. Такой подход к классификации статистического инструментария позволяет выявить как массивы статистических показателей в названных разрезах, так и конкретные органы власти - субъекты официального статистического учета, поставляющие данную информацию для целей управления. В конечном счете, представляется возможность оценить полноту интегрированных государственных информационных ресурсов как с точки зрения обеспечения различными институтами власти, так и с точки зрения достаточности этих ресурсов при проведении конкретных направлений государственной политики и оценки результатов. Как пример, приведем рубрику «Функции государства и политика управления». Рубрика включает следующие подрубрики: политическая функция, обеспечение безопасности, социальная, внешнеполитическая, экономическая, экологическая функции. Так, например, в описании подрубрики «Жилищная политика» были использованы как показатели, публикуемые в ежегодных сборниках Росстата, так и показатели из всего массива форм федерального статистического наблюдения: 1-жилфонд «Сведения о жилищном фонде»; 22-ЖКХ «Сведения о работе жилищно-коммунальных организаций в условиях реформы»; 4-жилфонд «Сведения о предоставлении гражданам жилых помещений»; 1-КХ «Сведения о благоустройстве городских населенных пунктов»; 1-КР «Сведения о капитальном ремонте жилищного фонда»; 1-тарифы «Сведения о тарифах на жилищно-коммунальные услуги»; 6-ЖКХ «Сведения о строительстве и ремонте объектов благоустройства».

На основе анализа данных типов документов были выделены группы показателей, описывающих деятельность государства в области жилищной политики: жилой фонд и жилищные условия населения; структура жилищного фонда; благоустройство жилищного фонда; жилищные условия населения; реформа прав собственности на жилье и формирование рынка жилья; реформа

управления жилищным фондом; жилищно-коммунальное хозяйство; объем услуг, оказанных населению предприятиями ЖКХ; инфраструктура ЖКХ; расходы населения на оплату услуг ЖКХ; меры социальной поддержки по оплате жилья и коммунальных услуг.

Всего в категорию «Жилищная политика» вошло около 60 показателей.

Выше отмечалось, что в ходе работы были изучены формы федерального статистического наблюдения, публикуемые Росстатом, т.е. первичные документы, заполняемые субъектами экономической деятельности. Изучение более 800 форм продемонстрировало наличие разрыва между составом показателей, официально публикуемых Росстатом и другими ведомствами, и составом показателей, в принципе собираемых и обрабатываемых, но недоступных для анализа широкому кругу исследователей. Такой анализ позволил определить неиспользуемый (или слабо используемый) потенциал существующей системы статистического наблюдения. Более широкое вовлечение этих показателей в практику управления может существенным образом улучшить как качество исследований, затрагивающих различные аспекты государственной политики и управления, так и качество принимаемых управленческих решений при минимальном увеличении (или даже отсутствии) издержек, связанных со сбором информации.

Параллельно с работами по формированию списка статистических показателей велись работы с документами предметной области из коллекций УИС РОССИЯ, содержание которых наиболее представительно с точки зрения тематики онтологии «Государственное управление» - Мониторинг Министерства экономического развития РФ, Вестники Банка России, Бюллетени Счетной палаты, публикации исследовательских центров, журнал «Эксперт», газеты «Коммерсант», «Ведомости», «Новая газета», «Известия».

4. Предварительное тестирование онтологии как поискового инструмента

В 2010 году было проведено тестирование онтологии как инструмента поиска по нескольким коллекциям УИС РОССИЯ, содержание которых наиболее представительно с точки зрения тематики онтологии. Из документов этих коллекций был составлен контрольный массив текстов, на котором тестировалась релевантность отнесения выбранных документов к рубрикам рубрикатора. В публикациях содержалось большое количество таблиц. Анализ результатов свидетельствует о том, что рубрики и большинство подрубрик в основном описаны корректно. Рассмотрим в качестве примера соответствие некоторых типовых документов из контрольного массива рубрикам рубрикатора. Так, например, таблица из Российского статистического ежегодника «Устройство детей и подростков,

оставшихся без попечения родителей» [8, с. 226] отнесена к рубрике «Функции государства и политика управления / Социальная функция/ Социальная политика», так как содержит такие показатели как: всего выявлено и учтено детей и подростков, оставшихся без попечения родителей; устроено в дома ребенка, детские дома и школы-интернаты для детей-сирот и детей, оставшихся без попечения родителей, учреждения социальной защиты населения и иные учреждения на полное государственное обеспечение; устроено под опеку (попечительство), на усыновление и др.

Перечисленные показатели соответствуют рубрике «Функции государства и политика управления / Социальная функция / Социальная политика». Кроме того, данный документ также отнесен к рубрике «Институты власти / Институты государственного управления / Исполнительная власть / Федеральные органы исполнительной власти / Федеральные агентства / Федеральное агентство по образованию». Принадлежность данного документа указанной рубрике определяется Приказом Федеральной службы государственной статистики от 27.07.2009 № 150 «Об утверждении статистического инструментария для организации Рособразования статистического наблюдения за деятельностью образовательных учреждений», в котором утверждена форма федерального статистического наблюдения № 103-РИК «Сведения о выявлении и устройстве детей-сирот и детей, оставшихся без попечения родителей».

Еще одним типом документов являются статьи в средствах массовой информации. Например, статья «Модернизация: элита плюс общество» («Ведомости» № 9(2527) от 21.01.2010) с функциональной точки зрения однозначно отнесена к рубрике «Функции государства и политика управления / Экономическая функция / Научно-техническая политика», что определяется наличием в ней таких терминов как: инновации, технологии, технологические новшества, модернизация, импорт технологий и др., которыми описывается данная рубрика.

Что касается рубрикации данной статьи по институтам власти, то она, в первую очередь, будет отнесена к рубрике «Институты власти / Президент», что определяется многократным использованием в ней соответствующих терминов: президент, послание президента, помощник президента.

Постановление Правительства №248 от 19.04.2010 «О дополнительных мерах по стимулированию жилищного строительства, переселению граждан из аварийного жилищного фонда и капитальному ремонту многоквартирных домов» отнесено к таким рубрикам как: «Функции государства и политика управления / Социальная функция / Жилищная политика», «Институты власти / Федеральные органы исполнительной власти / Правительство РФ», «Институты власти / Федеральные органы исполнительной власти /

Федеральные министерства и ведомства в ведении министерств / Министерство регионального развития». Что определяется наличием в документе таких терминов как: жилищный фонд, аварийный жилищный фонд, капитальный ремонт многоквартирных домов и др., а также Правительство РФ, Министерство регионального развития.

Результаты определения рубрик и подрубрик контрольной коллекции документов свидетельствуют о том, что базовая структура онтологии корректно отражает схему предметной области «Государственное управление», и как поисковый инструмент онтология будет давать удовлетворительные результаты с точки зрения релевантности и полноты.

Вместе с тем отметим, что рубрикатор требует постоянной корректировки. Так, пример 2 демонстрирует ряд проблем, которые являются следствием масштабности предметной области и неустойчивости ряда подрубрик. Во-первых, состав подрубрик рубрики «Институты власти / Институты государственного управления / Исполнительная власть / Федеральные органы исполнительной власти» повторяет структуру федеральных органов исполнительной власти, существовавшую на момент разработки рубрикатора. Регулярные изменения, происходящие в структуре федеральных органов исполнительной власти, требуют постоянной актуализации рубрикатора, чтобы не допустить снижения качества поиска, осуществляемого с помощью онтологии.

Кроме того, изменения в системе федеральных органов исполнительной власти приводят к преобразованию функций и сфер ответственности различных ведомств. Так, упразднение Федерального агентства по образованию (Рособразование) повлекло за собой передачу его функций Министерству образования и науки РФ. На практике это означает, что предлагаемая в качестве примера таблица должна быть отнесена в настоящий момент не к рубрике «Институты власти / Институты государственного управления / Исполнительная власть / Федеральные органы исполнительной власти / Федеральные агентства / Федеральное агентство по образованию», а к рубрике «Институты власти / Институты государственного управления / Исполнительная власть / Федеральные органы исполнительной власти / Федеральные министерства / Министерство образования и науки».

Именно необходимостью обеспечения более устойчивой структуры связей обусловлены изменения, внесенные в следующую версию рубрикатора. В новой версии сокращено количество рубрик за счет агентств и служб, подведомственных министерствам, что хотя и снижает в ряде случаев точность поиска, однако обеспечивает большую стабильность и дружелюбность по отношению к пользователям.

5. Заключение

Коллективом сотрудников в рамках проекта «УИС РОССИЯ» ведутся работы по созданию инструмента для решения одной из актуальных информационных, технологических и лингвистических задач по содержательной интеграции текстовых документов и статистических показателей для исследований процессов в области государственного управления.

Классификационные принципы, положенные в основу разработки рубрикатора как основы онтологии позволяют определять место документа в описании предметной области по функциональному, институциональному и инструментальному признакам. Реализация онтологии как полноценного поискового инструмента – следующий этап работы коллектива.

Литература

- [1] Noy, N.F. *Ontology Development 101: A Guide to Creating Your First Ontology* / Natalya F. Noy and Deborah L. McGuinness // *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880*. March 2001. [Электронный ресурс]. — Режим доступа: http://protege.stanford.edu/publications/ontology_development/ontology101.html
- [2] Musgrave, R.A. *Public Finance and Three Branch Model* // *Journal of Economics & Finance*. 2008. Vol. 32/ P. 334–339.
- [3] *Thesaurus of Australian Government Subjects*. Version 1.1. [Электронный ресурс]. — Режим доступа: http://www.finance.gov.au/e-government/service-improvement-and-delivery/publication-guidelines/docs/TAGS_v1.1.pdf, p. 3.
- [4] Богомолова, А. В. Университетская информационная система РОССИЯ: ресурсы и сервисы для поддержки общественного участия и задач государственного управления / А. В. Богомолова, Т.Н. Юдина, Дышкант Н.Ф. // *Интернет и современное общество: Труды XI Всероссийской объединенной конференции*. Санкт-Петербург, 28 – 30 октября 2008 г. СПб., 2008. С. 196-199.
- [5] Лукашевич, Н.В. *Тезаурусы в задачах информационного поиска*. М.: Изд-во МГУ, 2011.
- [6] *Методы государственного регулирования экономики: Учеб. пособие*. – М.: ИНФРА-М, 2008.
- [7] *Общероссийский классификатор органов государственной власти и управления*. ОК 006 – 93. М., 2010. [Электронный ресурс]. — Режим доступа: <http://arhangelstat.ru/method/classifier/%D0%9E%D0%9A%D0%9E%D0%93%D0%A3.aspx>

- [8] Российский статистический ежегодник . 2010:
Стат.сб. / Росстат. - М., 2010.

**UIS RUSSIA: Public Administration
Domain Ontology for Socio-Economic
Statistical Data and Analytical Publications
Integration**

Tatyana N. Yudina, Anna V. Bogomolova

The paper reports the work undertaken by University Information System RUSSIA team on public administration domain ontology to provide for domain knowledge organization to assist in public administration at federal, regional and local levels. The ontology will serve for statistical data and analytical

publications processing and classification at front end and as a search and query refinement instrument in UIS RUSSIA. Foreign countries teams research projects in the field are also investigated and approaches compared. Public administration is under increasing social attention throughout the world due to its effective performance being crucial for any country position and competitiveness in the future. National scientific communities are mobilized to participate and to apply their fundamental investigations results to assist in innovative information infrastructure composing to empower public administration agencies with the domain knowledge to make evidence-based decisions, to calculate and investigate possible risks.

* Работа выполняется при поддержке РГНФ, гранты 11-02-12012в и 10-02-12116в.