

Г.В. КАНЫГИН, М.С. ПОЛТИННИКОВА, В.С. КОРЕЦКАЯ

ОПЫТ ПОСТРОЕНИЯ СОЦИАЛЬНОГО ЗНАНИЯ НА ОСНОВЕ КОМПЬЮТЕРНЫХ ОНТОЛОГИЧЕСКИХ МЕТОДОВ

Аннотация. В статье обсуждаются проблемы создания инструментальных средств коллективного описания социальных процессов. Продемонстрированы плюсы и минусы естественного языка в роли инструмента такого описания. Проанализированы преимущества графовых контекстно-ориентированных онтологических (ГКОО) методов управления знаниями, позволяющих создавать и интегрировать знания, возникающие в социальной коммуникации разного уровня. Показаны достоинства ГКОО методов в качестве инструмента, ассистирующего при естественно-языковом описании социальных процессов их рядовыми участниками: наглядность и доступность для неспециалиста, генерация структурного представления знаний любого объема в виде графов, составленных из естественно-языковых обозначений; контроль связности аналитического представления знаний, выраженных на естественном языке; возможности алгоритмического соединения результатов разных концептуализаций и др. Прикладные преимущества ГКОО методов проанализированы на примерах концептуализации текстов разной направленности.

Ключевые слова: социальное знание; управление знаниями; анализ качественных данных; графовые контекстно-ориентированные онтологические метод; метафора.

Для цитирования: Каныгин Г.В., Полтинникова М.С., Корецкая В.С. Опыт построения социального знания на основе компьютерных онтологических методов // Социологический журнал. 2017. Том 23. № 3. С. 125–143. DOI: 10.19181/socjour.2017.23.3.5367

Каныгин Геннадий Викторович — доктор социологических наук, ведущий научный сотрудник, Социологический институт Федерального научно-исследовательского социологического центра Российской академии наук.
Адрес: 190005, Санкт-Петербург, ул. 7-я Красноармейская, д. 25/14, к. 505.
Телефон: +7 (921) 352-14-41. **Электронная почта:** g.kanygin@gmail.com

Полтинникова Мария Сергеевна — кандидат физ.-мат. наук, старший научный сотрудник, Социологический институт Федерального научно-исследовательского социологического центра Российской академии наук.
Адрес: 190005, Санкт-Петербург, ул. 7-я Красноармейская, д. 25/14.
Телефон: +7 (921) 648-47-64. **Электронная почта:** maria.poltinnikova@gmail.com

Корецкая Виктория Станиславовна — младший научный сотрудник, Социологический институт Федерального научно-исследовательского социологического центра Российской академии наук.
Адрес: 190005, Санкт-Петербург, ул. 7-я Красноармейская, д. 25/14.
Телефон: +7 (812) 316-24-96. **Электронная почта:** si_ras@mail.ru

Введение

В попытках понять, что такое социальное знание и каким образом изучение законов его функционирования может помочь решению актуальных проблем развития общества, мы в первую очередь заинтересовались трудами классиков социологии — Э. Дюркгейма, Г. Зиммеля, К. Маркса, М. Вебера, Т. Парсонса, А. Шюца, Дж. Мида, Дж. Хоманса, П. Бурдьё, Э. Гидденса и других. Затем мы обратились к работам социологов, развивающих тематику обыденного и научного знания об обществе, которые привлекли наше внимание благодаря коллективной монографии [6]. Но в какой-то момент изучения взглядов известных авторов мы ясно осознали, что знакомство с большим числом научных мнений не ведет к лучшему пониманию того, что такое социальное знание.

Мы ощутили себя в ситуации социолога-качественника, перед которым лежат тексты многочисленных информантов, а он не знает, как соединить содержащиеся в них суждения. И вслед за социологом-качественником мы почувствовали необходимость концептуализировать научные тексты своих социологических информантов так, как это принято в качественной социологии, т. е. воспроизводя смыслы, закладываемые самими информантами. Но, хотя мы знакомы с техниками концептуализации, применяемыми в качественном исследовании, к сожалению, нам не удалось осуществить задуманное относительно научных текстов, раскрывающих тематику социального знания.

Анализируя причину неудачи, мы обратили внимание, что методы качественного исследования содержат специальный аппарат концептуализации, называемый функциями кодирования и реструктурирования текста [15; 16]. А разработчики социального знания, труды которых мы изучали, свои построения осуществляют с помощью естественного языка. Они не предлагают читателям какого-то концептуализирующего инструмента, подобно тому, как социологу предлагается компьютерный пакет, скажем, NVivo или Atlas.ti, чтобы он мог лучше понять своих информантов.

Оказавшись в затруднительном положении, мы обратили внимание на то, что сегодня социальное знание и его модели развиваются не только в виде теорий социологов, но и разнообразных информационных технологий. На пике этого развития находятся методы управления знаниями [5; 7; 13].

Получается, что имеются два разных научных подхода, взявших на себя миссию развивать знания об одном и том же объекте — обществе. В чем же различие в исходных принципах и инструментальных средствах, которые характеризуют эти отдельные направления исследований в области социального знания? Не останавливаясь на деталях, дадим ответ по существу. В трудах социологов в качестве основного

инструмента описания того, что такое социальное знание, используется естественный язык. Компьютерные технологии основаны на отношениях и методах объектно-ориентированного программирования, которые мы далее будем называть компьютерной функциональностью.

Нельзя ли соединить в одном научном подходе такие разные методы создания знания — с помощью естественного языка и с помощью компьютерной функциональности? Создать новые методы построения знаний об обществе, которые способны вобрать в себя доступность и универсальность естественного языка и потенциал аналитического аппарата компьютерных технологий. Такое соединение позволяет по-новому поставить задачи сбора и поддержания знаний об обществе. В частности, открываются новые возможности для разработки краудсорсинговых технологий [14] описания общества, то есть широко-масштабного построения знания о социальных процессах силами их участников.

Представляется, что при разработке инструментальных средств со столь масштабными возможностями применения нужно начинать с анализа конкретных случаев формирования знания. Такой подход давно пропагандируется в качественной социологии.

На наш взгляд, общение между социологом и его информантами по поводу обыденных ситуаций уже содержит ключевое несоответствие. Изучаемый феномен представлен двояко: с одной стороны, информантами с помощью естественного языка (преимущественно в виде текстового описания), а с другой — категориальным аппаратом социолога, выраженным с помощью компьютерных инструментов. И так же, как в случае магистральных подходов, возникает проблема соотношения знания, выраженного на естественном языке, и знания, представляемого аналитически.

В литературе можно встретить выделение отдельных видов знания об обществе, например обыденное и научное [1; 17]. Мы исходим из того, что специфика знания об обществе в любом его виде — философском, социологическом, обыденном и т. д. — заключается в построении знания его носителями с помощью естественно-языковых метафор. Даже основа современного электронного управления — документы — это прежде всего понимаемые человеком метафоры, а только потом — структуры, создаваемые в информатике [10]. Поэтому нужно осознать плюсы и минусы функционирования социального знания, обусловленные применением речи в качестве инструмента социальной коммуникации.

Плюсы и минусы описания социальных явлений с помощью естественного языка

Принципиальные трудности при описании с помощью естественного языка предметных областей известны лингвистам [9]. Трудности

возникают из-за использования в качестве средств социальной коммуникации неясных механизмов смыслообразования — метафор [8]. В результате знания об обществе оказываются представленными в виде «литературной» теории [11], которую любой участник социальных процессов имеет возможность трактовать по своему усмотрению.

Попытаемся увидеть плюсы и минусы метафор в разных ситуациях их употребления при описании социальных процессов. Рассмотрим в качестве примера такого описания естественно-языковое утверждение «население сегодня потребляет столько-то продуктов питания». Разные варианты этого утверждения можно часто слышать в эфире и читать в социологических отчетах. Прежде чем разбираться, сколько потребляет население в тоннах или граммах, мы хотели бы понять, какие именно социальные явления или отношения имеют в виду авторы подобных фраз.

Интуитивно ощущаемая трудность подобного понимания состоит в том, что социальный опыт, которым пользуется каждый из нас, всегда соотнесен с наблюдаемыми процессами, осуществляемыми людьми из плоти и крови. В социальных реалиях существует конкретный человек, что зафиксировано, например, наличием у него паспорта или знакомством с ним других людей. Человек пошел в магазин — это также не происходит «вообще», а всегда конкретно, и могут быть установлены, скажем, дата и время. Человек купил что-то (см. чек). Отвез это что-то куда-то на машине. Вот реальные процессы, переплетенные между собой.

Население же — это обобщенное представление о некоей подразумеваемой группе лиц. Да, каждый из нас способен пользоваться такого сорта обобщениями, но подобное словотворчество вовсе не обязано соотноситься с социальными реалиями. Оно всегда опосредовано автором метафорического описания, который, как известно, может исходить из собственных социально обусловленных принципов.

«Население потребляет...» — это так называемая стертая метафора (см. [8]). Оставляя лингвистам выяснение того, как теоретически строго описать феномен метафоры, мы сосредоточимся на последствиях применения метафор, которые важны для практики социального взаимодействия. Во-первых, некий социальный феномен, названный с помощью естественного языка «населением», является конструкцией, создаваемой в рамках самого языка. Социологическое понятие «население» — это фигура речи, которая, безусловно, узнаваема каждым из носителей языка. Но вопрос о том, к какому реальному процессу относится такая «узнанная концептуальная единица», требует неочевидной конкретизации ответа. Во-вторых, операция «потребления», осуществление которой приписывается «населению», как следствие неопределенности этого понятия, также оказывается конструкцией языка.

Почему столь популярны метафоры при описании социальных явлений? Ответим на примере известной фразы «председатель продирался сквозь собрание». Вместо того чтобы детально сообщить, что именно председатель собрания делал — открывал собрание при отсутствии кворума, успокаивал спонтанные аплодисменты, отвечал на странные реплики и т. п., говорящий лишь уподобляет одну группу коннотаций, стоящих за словом «собрание», другой группе коннотаций, обозначенных словом «продирался».

Этот ясный для любого носителя языка пример предъясвляет прикладное решение общеизвестной проблемы сложности социальных явлений. Метафора позволяет узнаваемое явление — «собрание людей» — изображать не как реальный процесс, оказавшись в котором человек способен осуществить те или иные действия, а ограничиться уподоблением с помощью языка другому узнаваемому феномену — «продираться». Такое уподобление понимаемо носителем языка и может применяться им для построения естественно-языковых конструкций, но никак не поясняет связи этих конструкций с чем-либо существующим вне головы носителя языка.

При высказываниях, подобных приведенному, говорящий всегда подразумевает массы предопределений. Население — не вообще, а применительно, скажем, к городу. Продукты питания — это то, что санкционировано в пищу социальным опытом, национальной традицией и т. п. еще до того, как говорящий упомянул о продуктах. Но, говоря о питании, что человек имеет в виду: людей или домашних животных, которых эти люди содержат?

Все эти и многие другие «сущности», используемые говорящим, уже кто-то, включая его самого, когда-то и где-то определял. Иначе они не были бы узнаваемы. Без такого предопределения само говорение оказывается бессмысленным для всех участников этого процесса. Однако, внимая словам, мы остаемся без устраивающего нас ответа на вопрос, кем является этот кто-то — скажем, «коллективной практикой», конкретными государственными инстанциями или общественными организациями?

Сегодня предопределенные описания (население, питание, город и многие другие) существуют в практике публичной социальной коммуникации как в виде документов, так и в «головах людей» [19]. Хотелось бы, чтобы естественно-языковые утверждения, возникающие в публичной коммуникации, могли быть соотнесены каждым ее участником с теми реальными социальными объектами, отношениями, событиями, которые различимы в речи говорящего. Но для этого приходится задумываться, чем и как заменить метафоры в качестве инструмента социальных взаимодействий.

Ход мыслей авторов не отличается ни революционностью, ни утопичностью: именно такой подход реализуется в онтологических ме-

тодах, развиваемых в информатике и лингвистике. Например, лингвисты по корпусу естественно-языковых текстов предварительно строят онтологию, которую применяют для семантического анализа текстов, не входящих в корпус.

Однако для того, чтобы реализовать эту идею, нужно предложить инструментальный аппарат, который заменил бы метафору при описании сложных социальных явлений. В качестве искомых методов в статье рассматриваются разработанные нами ранее графовые контекстно-ориентированные онтологические (ГКОО) методы [3]. Эти аналитические процедуры относятся к классу информационных онтологических методов, называемых также концептуальным моделированием. Основания и особенности применения онтологических методов в области социальных наук рассмотрены в [4].

Создание знания графовыми контекстно-ориентированными онтологическими методами

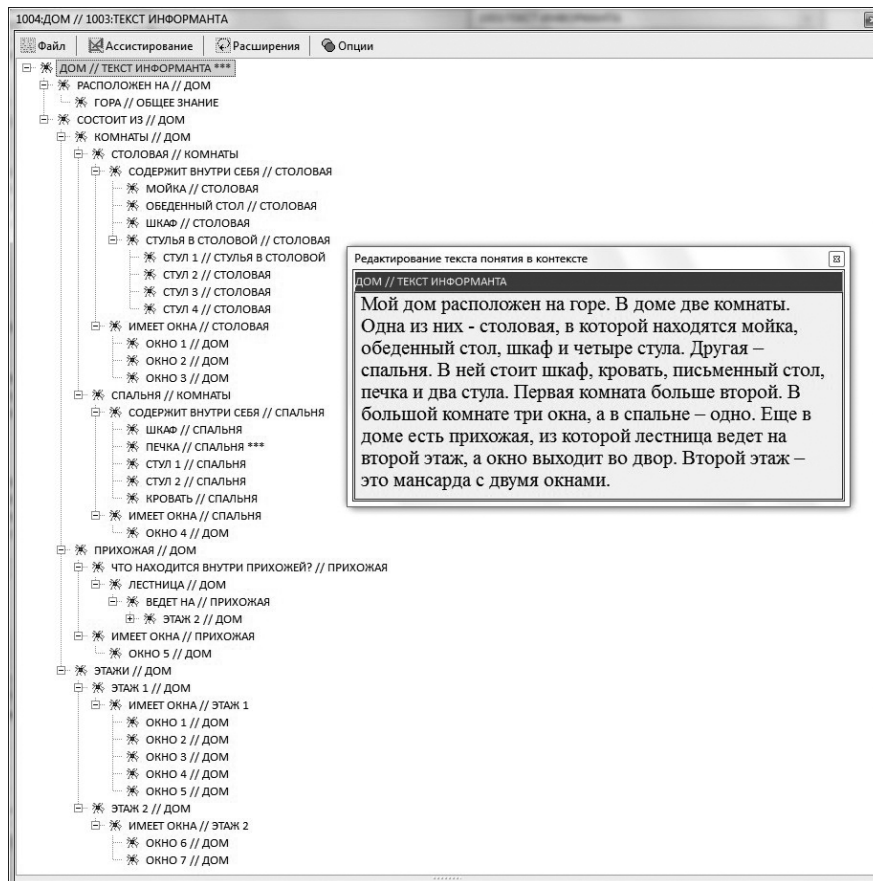
Рассмотрим типичный социологический случай, который мы анализируем в качестве модели построения любого социального знания. Речь идет о полевой работе социолога. Эта модельная ситуация характеризуется, с одной стороны, наличием информантов, которые формулируют свои знания на естественном языке и представляют их в виде текста. С другой — социолога (аналитика), который переформулирует знания информантов в аналитическом виде. Целью переформулирования является интеграция социологом знаний информантов. Принципиально важно, что, во-первых, информанты и социолог не могут использовать какие-то специальные знания, относящиеся к определенной предметной области. Во-вторых, концептуализируемые социологом знания имеют сложную структуру и огромные объемы.

Мы полагаем, что если удастся предложить инструментарий, который позволяет социологу аналитически представить знания информантов в приведенном модельном случае, то можно рассчитывать, что этот инструментарий окажется востребованным при построении знания в других ситуациях социального взаимодействия (бизнесе, государственном управлении, интеграции социологических методов и т. д.).

Мы демонстрируем достоинства предлагаемых ГКОО методов путем концептуализации двух разнородных текстов. Один текст, условно названный «естественнонаучным», описывает материальный объект — дом. Этот текст и результат его концептуализации в виде графа представлен на рисунке 1. Мы используем этот граф при пояснениях общих особенностей ГКОО инструментария.

Второй текст, условно квалифицируемый как «социологический», выражает переживания, которые испытывает информант, проживая

в собственном доме (известен в учебных курсах по анализу качественных данных как текст Терри; показан ниже).



**Рис. 1. Концептуализация «естественнонаучного» текста.
Описание материального объекта ДОМ**

Прежде чем перейти к анализу прикладной концептуализации, поясним терминологию, архитектуру, методiku и техники применения ГКОО методов.

Любая концептуализация, выполняемая аналитиком с помощью ГКОО инструментов, приводит к появлению концептуальной модели (иначе говоря, онтологии). Онтология представляет собой структуру, состоящую из словаря, тезауруса и алгоритмов. Словарь — это множество понятий. Тезаурус — это множество локальных связей каждого понятия с другими понятиями. Алгоритмы — это способы преобразования тезауруса (а не текста информанта!). В частности, граф, показанный на рисунке 1, сгенерирован алгоритмами.

Построение концептуальной модели осуществляется аналитиком с целью описания какого-либо сюжета или объекта любой сложности. Этот сюжет, в общем случае, может быть представлен в тексте информанта или подсказан опытом самого аналитика. В примерах все действия с понятиями, необходимые в процессе концептуализации (создание, введение связей и др.), выполнены авторами статьи с помощью специальной компьютерной программы *Diagogue*, созданной ими и относящейся к классу онторедкторов.

Логика и действия, описываемые в дальнейшем изложении, станут понятнее, если предварительно разобраться в основных техниках ГКОО моделирования, подробно объясненных на примере концептуализации упомянутого «естественнонаучного» текста в рукописи [12]. В данной статье мы стараемся воспроизвести ключевые объяснения этой рукописи, но, будучи ограничены объемом, сосредоточиваем основное внимание на структурных особенностях концептуальных моделей, наблюдаемых на рисунках.

Этапы построения концептуальной модели

Построение любой концептуальной модели предлагаемыми методами можно условно разделить на три взаимно переплетающиеся части. Во-первых, экспертное выявление и обозначение в виде слово-сочетаний естественного языка отдельных относительно независимых смысловых аспектов, различимых экспертом в тексте или представляющихся ему значимыми по собственному опыту. Эта операция в случае работы с текстом фактически представляет собой первичное кодирование свидетельства информанта, осуществляемое в анализе качественных данных [18].

Однако в случае ГКОО методов такое кодирование имеет принципиально важное отличие: обозначение проводится с учетом контекста. Употребляя несложную нотацию, аналитик должен каждое свое обозначение снабдить явным указанием обстоятельств, относительно которых он делает обозначение. Например, описывая комнаты в доме, о которых идет речь в первом тексте, аналитик пишет КОМНАТЫ//ДОМ. Такое двойное обозначение говорит, что аналитик, передавая смыслы информанта, рассматривает КОМНАТЫ относительно контекста ДОМ. А не в каких-либо других многочисленных обстоятельствах, которые он способен выявить в свидетельстве и обозначить с помощью понятия!

Во-вторых, аналитик разъясняет введенные им парные обозначения одни через другие в виде ответов на простой вопрос: «Что такое именно это?». Основания для такого ответа он черпает в свидетельстве информанта или собственном опыте. Такие ответы на естественном языке выглядят следующим образом. Что такое комнаты именно в доме, описанном в тексте? Это — спальня и столовая, но рассматри-

ваемые в контексте тех условий, которые чуть раньше обозначены как комнаты именно в доме. С помощью уже использованной нотации такой ответ можно записать короче. Что такое КОМНАТЫ//ДОМ? — это СПАЛЬНЯ//КОМНАТЫ и СТОЛОВАЯ//КОМНАТЫ. И человек продолжает этот процесс разъяснений одних пар понятий через другие до тех пор, пока не исчерпает все смыслы, передающие, с его точки зрения, концептуализируемый сюжет.

Такие действия не требуют от пользователя каких-то особых навыков сверх того, что ему уже известно из опыта создания структур папок в компьютере. Интерфейсы программы Diagogue предоставляют все операции, необходимые для связывания понятий (в виде пар) в режиме drag&drop.

В-третьих, по результатам выполнения любого концептуального действия (ввод понятия, указание для него контекста, определение и переопределение связей между понятиями и др.) ГКОО методы позволяют аналитику буквально увидеть, какая концептуальная модель получилась по словарю и тезаурусу, созданному им на данный момент. Такое визуальное представление является методом контроля аналитиком собственных концептуальных усилий (см. рис. 1 и примеры ниже).

Основы компьютерной функциональности ГКОО методов

Обратим внимание, что смыслы, не имеющие структурной определенности в свидетельстве информанта, в концептуальной модели воссоздаются в виде системы контекстов, возникающей благодаря двойным обозначениям. На рисунке 1, видно, что каждый узел графа представлен парой понятий. Первое — традиционное обозначение или термин (в терминологии качественного исследования — код), второе — контекст, то есть указание условий, при которых введен термин.

Создание системы контекстов — исключительно творческая задача аналитика. Но при ее решении человек вынужден систематически руководствоваться некоторыми правилами. Иначе ему не удастся разумно упорядочить огромную массу понятий, неизбежно возникающую при описании социальных феноменов. Эти правила управляют генерацией структур знания, показанных на рисунках.

В приведенных примерах мы воспользовались правилами динамического соподчинения, определяющими, является ли очередная пара допустимой в составе графа на том месте, где ее ожидает увидеть аналитик. Чтобы пара понятий оказалась узлом в составе графа (см. рисунки), второе понятие пары — контекст — должно уже входить в один из узлов, введенных «выше» по графу. Выполнение этого правила может быть прослежено по графу на рисунке 1.

Подобные правила давно доказали свою полезность в программировании. Любое обозначение пользователя, которое из-за явного указания контекста мыслится им в границах конкретного случая, не

должно выходить за рамки этого случая. ГКОО инструментарий побуждает пользователя следовать этому естественному принципу при каждом (!) обозначении. Более того, генерируя общие структуры знания, методы контролируют, насколько успешно пользователь соблюдает границы всех ранее сделанных им обозначений.

Особенности ГКОО концептуализации естественнонаучного текста

Теперь проанализируем, в чем и каким образом проявляются достоинства предлагаемого инструментария при концептуализации текста, описывающего материальный объект — дом. Такая концептуализация осуществлена нами путем составления по тексту информанта соответствующего словаря и тезауруса [12]. Ее результаты представлены в виде автоматически построенного графа, показанного на рисунке 1.

Для социолога концептуализация такого «естественнонаучного» текста может показаться надуманным упражнением. Действительно, не вызывает сомнения то, что для общепринятого описания материальной сферы в информатике или естествознании уже разработаны адекватные методы. В этой связи можно подумать, что для разработки социологических методов нет смысла анализировать концептуализацию «естественнонаучного» текста.

Однако, рассчитывая строить социальное знание на принципах краудсорсинга, мы вынуждены учесть ряд практически важных условий. Во-первых, информанты в своих свидетельствах не будут различать, говорят они о физическом предмете, для которого существуют соответствующие естественнонаучные теории, или о своих переживаниях относительно этого предмета. Например, Терри говорит о доме, в котором он живет один. И этот материальный объект представляет собой условия существования эмоций информанта.

Во-вторых, информант — это только носитель естественного языка, а не специалист в естественнонаучной области. Это означает, что любые характеристики явлений или объектов, моделируемых с помощью естественнонаучных методов, должны предварительно выражаться информантом на естественном языке. В частности, все числовые характеристики — это не конструкторы, разрабатываемые в виде математической теории, а числительные, которые информант употребляет на основе своих знаний, полученных в школе и закрепленных житейским опытом.

В-третьих, текстовые разъяснения материальных объектов в свидетельствах информантов могут относиться не только к количественным характеристикам этих объектов, но и к структурам. И хотя сложность таких структур может оказаться значительной, они должны быть наглядными для человека, обладающего квалификацией не выше пользователя офисного пакета.

ГКОО методы удовлетворяют всем трем условиям. Наши методы используют для обозначения только словосочетания естественного языка. При этом полагается, что методы позволяют упорядочивать именно обозначения информанта, а не его мотивы или побуждения что-то обозначить. Не делая различий между тем, что служит прообразом обозначений информанта — материальный объект, внутреннее переживание, оценка социального действия или что-то еще, ГКОО инструментарий, также как естественный язык, может применяться человеком для описания чего угодно.

В частности, как демонстрирует рисунок 1, не только сами понятия, обозначающие конкретные предметы или объекты — ДОМ, КОМНАТЫ, СТУЛЬЯ и другие, — выражены на естественном языке, но и отношения между ними создаются и обозначаются пользователем с помощью речи. В числе таких отношений СОСТОИТ ИЗ, ВЕДЕТ НА и др.

Тем самым предлагаемые методы дают возможность аналитику, в роли которого может выступать рядовой участник социальных процессов, самому формировать смысловые отношения между наблюдаемыми в тексте «сущностями» — предметами, свойствами и т. п. Такой подход позволяет отказаться от завышенных требований к квалификации пользователя, рассчитывая только на его умение общаться на естественном языке.

ГКОО методы дают в распоряжение своего адепта мощный структурный аппарат современного программирования. Некоторое начальное представление о воспроизведении структур, описываемых информантом с помощью текста, можно получить из рисунка 1. На нем структурно воспроизведены отношения между отдельными частями дома, увиденные нами в потоке слов информанта. Подчеркнем, что структура рисунка 1 сгенерирована алгоритмами, работающими при концептуализации любых текстовых описаний, а не только применительно к какой-то отдельной отрасли, скажем, домостроению. Таким образом, путем применения оригинального социологического инструментария удается в принципе преодолеть нестыковки методов, применяемых в разнообразных естественнонаучных предметных областях.

Обратим внимание на «количественную» масштабируемость концептуальных описаний, получаемых с помощью ГКОО инструментов по «естественнонаучному» свидетельству. Допустим, нам стали известны точные размеры окна, которое мы обозначили при концептуализации как ОКНО1//ДОМ. И мы хотим внести соответствующие изменения в первоначальную модель. Другими словами, нам нужно масштабировать уже построенную концептуальную модель сообразно новым требованиям.



**Рис. 2. Концептуализация «естественнонаучного» текста.
Масштабируемость свидетельства информанта
о материальном объекте**

Отметим конкретные достоинства такого масштабирования в случае прикладного применения наших методов. Во-первых, узел ОКНО1//ДОМ, который требует масштабирования, легко идентифицируется визуально по уже построенному графу. Во-вторых, все изменения вносятся локально относительно узла ОКНО1//ДОМ. Эти изменения состоят во введении соответствующих понятий ШИРИНА В САНТИМЕТРАХ и других, а также связей между ними (см. рис. 2, на котором локальный фрагмент, масштабирующий узел первоначально построенного графа, обведен прямоугольником). В-третьих, числовые значения вводятся как числительные (100 и 120 на рис. 2 — это текстовые обозначения), то есть пользователь по-прежнему применяет только естественный язык. В-четвертых, при генерации нового графа числовые значения автоматически разъясняются с помощью естественно-языковых выражений, то есть локально документируются (в данном случае для этой цели указана единица измерения).

Концептуализация «социологического» текста с помощью ГКОО методов

Теперь обратим внимание на достоинства ГКОО инструментария при концептуализации «социологического» текста. Начнем с решения основной задачи любой концептуализации — передачи смыслов сви-

детельства информанта. Словарь и тезаурус этой концептуализации приведены по ссылке [12]. Результаты, получившие пометку «версия 1», показаны в виде графа на рисунке 3.

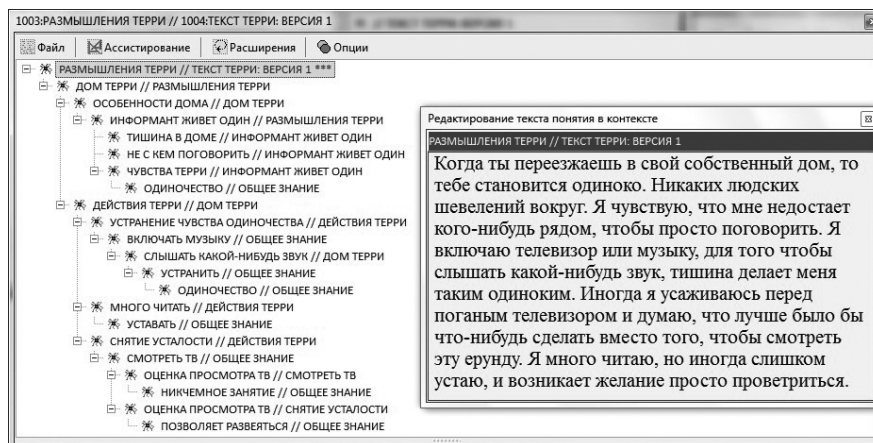


Рис. 3. Концептуализация «социологического» текста (версия 1: «инвентаризация» действий информанта)

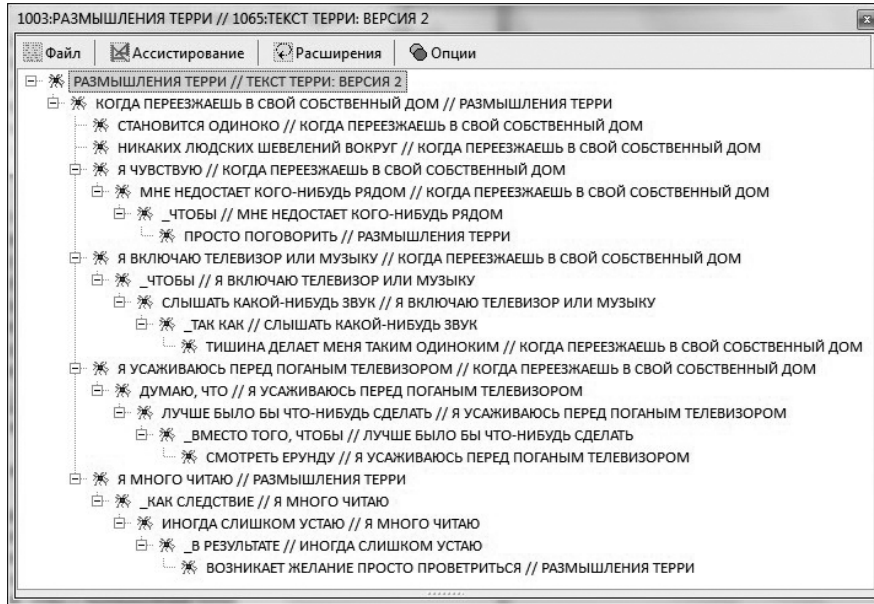
Что касается процедуры концептуализации — мы достигли своей цели: все смыслы информанта воспроизведены в модели. Однако нас не устроил ее «естественнонаучный» стиль. Концептуальная модель «инвентаризирует» смыслы информанта, но приходится затратить усилия, чтобы соотнести понятия ДОМ ТЕРРИ, ОСОБЕННОСТИ ДОМА, ОДИНОЧЕСТВО и т. п. с его текстом (см. рис. 3).

Стремясь улучшить «читаемость» результатов концептуализации, мы реорганизовали онтологию версии 1 за счет двух приемов. Во-первых, мы выразили понятия онтологии в виде собственных речевых формулировок Терри. Во-вторых, мы соединили их общеупотребительными связками (на рисунках они представлены словами ЧТОБЫ, ТАК КАК и др. со знаком «_» в начале). Все созданные понятия и связи мы добавили в уже указанные словарь и тезаурус. Так появилась версия 2, результаты которой показаны на рисунке 4.

Чего мы добились в результате второй концептуализации? Во-первых, адресуясь к навыкам речевого общения, можно утверждать, что преобразованная концептуальная модель не столь «шероховата» для восприятия носителем языка, как ее предыдущий аналог.

Во-вторых, она конкретизирует общеупотребительные конструкции естественного языка, делая их определенными строго в рамках собственных формулировок информанта. Например, введенное нами слово-связка _ЧТОБЫ в речи может означать «неопределенно многое»: причинно-следственную связь, указание на цель действия,

намерение информанта и т. п. Смыслы **_ЧТОБЫ**, соотнесенного моделью со словесными формулировками, значительно конкретнее (см. рис. 4).



**Рис. 4. Концептуализация «социологического» текста
(версия 2: словами информанта)**

В-третьих, цепочки рассуждений информанта воспроизведены на естественном языке, но структурно. Таким образом можно алгоритмически проследивать обоснования, интерпретации, умозаключения и другие логические действия человека. Более того, если информант включил в свой текст какие-либо числовые свидетельства, то становится возможным проводить вычисления на основе полученной модели.

В-четвертых, проводимая нами для примера «социологическая» концептуализация — это микроуровень глобальной практической задачи построения знания, используемого при социальной коммуникации. И хотя на этом уровне уже задействованы все ГКО механизмы, но ввиду вынужденной «камерности» примеров можно не обратить внимания, что графы генерируются, и для такой генерации совершенно неважен объем контента.

Заключение

Приверженность традиции качественного социологического исследования не позволяет нам предъявлять специальные требования к квалификации носителей знания. Наш носитель знания является

обычным человеком, квалификация которого в компьютерной сфере ограничена умением пользоваться некоторыми функциями офисного пакета, которые он сам для себя счел полезными. Тем самым, в отличие, например, от И. Девятко и соавторов, нам не нужно предполагать, что «индивиды обладают (неявным) знанием подлинного априорного распределения в данной области социально-демографических фактов» [1, с. 6].

Для малых «семантических объемов», которыми оперируют примеры в статье, предлагаемый инструментарий не выглядит «прорывным» результатом: пользователь компьютера может нарисовать показанные графы вручную. Но примеры специально созданы небольшими по объему, чтобы обеспечить обзорность выполняемых концептуальных действий. В работе [2] мы описываем концептуализацию социального объекта, являющегося одним из институтов РАН. Эта концептуализация приближает нас к реальности: словарь онтологии уже на начальном этапе концептуализации содержит сотни понятий, а тезаурус — сотни локальных определений этих понятий. При таких условиях рисовать графы знания вручную становится невозможно. Графы концептуальной модели уже не помещаются на экран, и действия с ними могут осмысленно осуществляться только с помощью компьютера. И при таких действиях их проверка и визуальное отображение с помощью ГКОО инструментария становятся жизненно необходимыми.

Другими словами, с помощью тех действий, которые мы описываем в рассматриваемых примерах, социолог, основываясь на краудсорсинговом подходе, становится способен предложить методы создания социального знания любого объема.

Ввиду ограниченности объема статьи мы не рассматривали инструментальные возможности предлагаемых методов при организации коллективной концептуальной работы создателей знания. Исследование этих возможностей требует отдельного изложения.

ЛИТЕРАТУРА

1. Девятко И.Ф., Абрамов Р.Н., Кожанов А.А. О пределах и природе дескриптивного обыденного знания о социальном мире // Социологические исследования. 2010. № 9. С. 3–17.
2. Каныгин Г.В., Полтинникова М.С. Конструирование социальных объектов с помощью контекстно-ориентированных онтологических методов // Социология и общество: социальное неравенство и социальная справедливость (Екатеринбург, 19–21 октября 2016 года) Материалы V Всероссийского социологического конгресса / Отв. ред. В.А. Мансуров. М.: Российское общество социологов, 2016. С. 7933–7945 [электронный ресурс — DVD ROM].

3. *Каныгин Г.В., Полтинникова М.С.* Контекстно-ориентированные онтологические методы в социологии. Труды СПИИРАН. 2016. Вып. 48. С. 107–124.
4. *Каныгин Г.В., Хорева Л.В.* Инновационный взгляд на сферу услуг: онтологии сервиса // Известия Санкт-Петербургского государственного экономического университета. 2016. № 5 (101). С. 78–85.
5. *Лапшин В.А.* Онтология в компьютерных системах. М.: Научный мир, 2010. — 224 с.
6. Обыденное и научное знание об обществе: взаимовлияния и реконфигурации: [Монография] / Под ред. И.Ф. Девятко, Р.Н. Абрамова, И.В. Катерного. М.: Прогресс-Традиция, 2016. — 328 с.
7. *Рубашкин В.Ш.* Онтологическая семантика. Знания. Онтологии. Онтологически ориентированные методы информационного анализа текстов. М.: Физматлит, 2012. — 348 с.
8. Теория метафоры: Сборник / Пер. с англ., фр., нем., исп., польск. яз.; Под ред. Н.Д. Арутюновой, М.А. Журиной; Вступ. ст. и сост. Н.Д. Арутюновой. М.: Прогресс, 1990. — 512 с.
9. *Фрумкина Р.М.* Психология и лингвистика как контексты социального познания // Пути России: Проблемы социального познания / Под общ. ред. Д.М. Рогозина. М.: МВШСЭН, 2006. С. 128–144.
10. Электронный документооборот. Базовые понятия и терминология. DOCFLOW [электронный ресурс]. Дата обращения 14.06.2017. URL: <<http://www.docflow.ru/edu/glossary/detail.php?ID=27946>>.
11. *Allais M.* La science économique d'aujourd'hui et les faits. Revue des Deux Mondes. Juin 1990. P. 54–74.
12. Context oriented knowledge [online]. Accessed 10.06.2017. URL: <<http://coknowledge.ru/materials-072017/>>.
13. *Guizzardi G., Wagner G., Almeida J.P.A., Guizzardi R.S.S.* Towards Ontological Foundations for Conceptual Modeling: The Unified Foundational Ontology (UFO) Story, 2015 [online]. Accessed 25.08.2017. URL: <<https://inf.ufes.br/~gguizzardi/UFO-Story.pdf>>.
14. *Howe J.* Crowdsourcing. Why the Power of the Crowd is Driving the Future of Business. New York: Crown Publishing Group, 2008. — 320 p.
15. *Kelle U.* Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data // Sociological Research Online. 1997. Vol. 2. No. 2 [online]. Accessed 05.04.2008. URL: <<http://www.socresonline.org.uk/2/2/1.html>>.
16. *Lewins A. and Silver C.* Using Qualitative Software: A Step-by-Step Guide. London: Sage publ., 2007. — 283 p. DOI: 10.4135/9780857025012
17. *Nilsen P., Nordström G., Ellström P.E.* Integrating research-based and practice-based knowledge through workplace reflection // Journal of Workplace Learning. 2012. Vol. 24. Iss. 6. P. 403–415. DOI: 10.1108/13665621211250306
18. How and what to code. ONLINE QDA. Learning Qualitative Data Analysis on the Web [online]. Accessed 01.09.2017. URL: <http://onlineqda.hud.ac.uk/Intro_QDA/how_what_to_code.php>.

19. *Wagner Ch.* Breaking the Knowledge Acquisition Bottleneck Through Conversational Knowledge Management // *Information Resources Management Journal*. January-March 2006.No. 19 (1). P. 70–83.

Дата поступления: 11.07.2017.

SOTSILOGICHESKIY ZHURNAL = SOCIOLOGICAL JOURNAL.

2017. Vol. 23. No. 3. P. 125–143. DOI: 10.19181/socjour.2017.23.3.5367

G.V. KANYGIN, M.S. POLTINNIKOVA, V.S. KORETSKAYA

Sociological Institute of FCTAS RAS;
St Petersburg, Russian Federation.

Gennady V. Kanygin — Doctor of Sociology, Leading Researcher.

Address: 25/14, 7-ya Krasnoarmeyskaya Str., 190005, St Petersburg, Russian Federation. **Phone:** +7 (921) 352-14-41. **Email:** g.kanygin@gmail.com

Maria S. Poltinnikova — Candidate of physical and mathematical sciences, Senior Researcher.

Address: 25/14, 7-ya Krasnoarmeyskaya Str., 190005, St Petersburg, Russian Federation. **Phone:** +7 (921) 648-47-64. **Email:** maria.poltinnikova@gmail.com

Victoria S. Koretskaya — junior researcher.

Address: 25/14, 7-ya Krasnoarmeyskaya Str., 190005, St Petersburg, Russian Federation. **Phone:** +7 (812) 316-24-96. **Email:** interperfection@gmail.com

**EXPERIENCE IN SOCIAL KNOWLEDGE CONSTRUCTION
BASED ON COMPUTERIZED ONTOLOGICAL METHODS**

Abstract. In this article, the problems of creating tools for the collective description of social processes are discussed. Demonstrated are the pros and cons of using natural language as a tool for generating such a description. Analyzed are the benefits of a new graphic context-oriented approach (GCO) to knowledge management, which allows creating and integrating knowledge arising within various levels of social communication. Shown are the merits of GCO methods as an instrument helping ordinary participants describe social processes in natural-language terms: they are intuitive to non-specialists; any amount of knowledge can be represented structurally in the form of graphs composed of natural language designations; it is possible to control the coherence of the analytical representation of knowledge expressed in natural language; finally, it is possible to connect the results of different conceptualizations into an algorithm. The practical advantages of GCO methods are analyzed based on examples of conceptualizing texts with varying content.

Keywords: social knowledge; knowledge management; qualitative data analysis; graph context-oriented ontological methods; metaphor.

For citation: Kanygin G.V., Poltinnikova M.S., Koretskaya, V.S. Experience in Social Knowledge Construction Based on Computerized Ontological Methods. *Sotsiologicheskii Zhurnal = Sociological Journal*. 2017. Vol. 23. No. 3. P. 125–143. DOI: 10.19181/socjour.2017.23.3.5367

REFERENCES

1. Devyatko I.F., Abramov R.N., Kozhanov A.A. On limits and nature of descriptive everyday knowledge in the social world. *Sotsiologicheskie issledovaniya*. 2010. No. 9. P. 3–17. (In Russ.)
2. Kanygin G.V., Poltinnikova M.S. Designing social objects using context-oriented ontological methods. *Sotsiologiya i obshchestvo: sotsial'noe neravenstvo i sotsial'naya spravedlivost' (Ekaterinburg, 19–21 oktyabrya 2016 goda)* [Sociology and society: social inequality and social justice (Yekaterinburg, October 19–21, 2016).] Ed. by V.A. Mansurov. Moscow: Rossiiskoe obshchestvo sotsiologov publ., 2016. P. 7933–7945. [DVD ROM]. (In Russ.)
3. Kanygin, G.V., Poltinnikova, M.S. Context-Oriented Ontological Methods in Sociology. *SPIIRAS Proceedings*. 2016. Iss. 48. P. 107–124. (In Russ.)
4. Kanygin G.V., Khoreva L.V. Innovative view on the sphere of services: ontology of service. *Proceedings of the St Petersburg State Economic University*. 2016. No. 5 (101). P. 78–85. (In Russ.)
5. Lapshin V.A. *Ontologiya v komp'yuternykh sistemakh*. [Ontologies in computer systems.] Moscow: Nauchnyi mir publ., 2010. 224 p. (In Russ.)
6. *Obydennoe i nauchnoe znanie ob obshchestve: vzaimovlianiya i rekonfiguratsii: [Monografiya]*. [Ordinary and scientific knowledge of society: Interference and reconfiguration.] Ed. by I.F. Devyatko, R.N. Abramov, I.V. Katernyi. Moscow: Progress-Traditsiya publ., 2016. 328 p.
7. Rubashkin V.Sh. *Ontologicheskaya semantika. Znaniya. Ontologii. Ontologicheski orientirovannyye metody informatsionnogo analiza tekstov*. [Ontological semantics. Knowledge. Ontologies. Ontologically oriented methods of information analysis of texts.] Moscow: Fizmatlit publ., 2012. 348 p. (In Russ.)
8. *Teoriya metafory: Sbornik*. [Theory of metaphor: Collection.] Transl. from Eng., French, Germ., Span., Polish; Ed. by N.D. Arutyunova, M.A. Zhurinskaya. Moscow: Progress publ., 1990. 512 p. (In Russ.)
9. Frumkina R.M. Psychology and linguistics as the contexts of social cognition. *Puti Rossii: Problemy sotsial'nogo poznaniya*. [The Ways of Russia: Problems of Social Cognition.] Ed. by D.M. Rogozin. Moscow: MVShSEN publ., 2006. P. 128–144. (In Russ.)
10. Electronic document management. Basic concepts and terminology. *DOCFLOW*. Accessed 14.06.2017. URL: <<http://www.docflow.ru/edu/glossary/detail.php?ID=27946>>. (In Russ.)
11. Allais M. La science économique d'aujourd'hui et les faits. *Revue des Deux Mondes*. Juin 1990. P. 54–74.
12. *Context oriented knowledge*. Accessed 10.06.2017. URL: <<http://coknowledge.ru/materials-072017/>>.
13. Guizzardi G., Wagner G., Almeida J.P.A., Guizzardi R.S.S. *Towards Ontological Foundations for Conceptual Modeling: The Unified Foundational Ontology (UFO) Story*, 2015. Accessed 25.08.2017. URL: <<https://inf.ufes.br/~gguizzardi/UFO-Story.pdf>>.
14. Howe J. *Crowdsourcing. Why the Power of the Crowd is Driving the Future of Business*. New York: Crown Publishing Group, 2008. 320 p.
15. Kelle U. Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data. *Sociological Research Online*. 1997. Vol. 2. No. 2. Accessed 05.04.2008. URL: <<http://www.socresonline.org.uk/2/2/1.html>>.
16. Lewins A. and Silver C. *Using Qualitative Software: A Step-by-Step Guide*. L.: Sage publ., 2007. 283 p. DOI: 10.4135/9780857025012

17. Nilsen P., Nordström G., Ellström P.E. Integrating research-based and practice-based knowledge through workplace reflection. *Journal of Workplace Learning*. 2012. Vol. 24. Iss. 6. P. 403–415. DOI: 10.1108/13665621211250306
18. How and what to code. ONLINE QDA. *Learning Qualitative Data Analysis on the Web*. Accessed 01.09.2017. URL: <http://onlineqda.hud.ac.uk/Intro_QDA/how_what_to_code.php>.
19. Wagner Ch. Breaking the Knowledge Acquisition Bottleneck Through Conversational Knowledge Management. *Information Resources Management Journal*. January-March 2006. No. 19 (1). P. 70–83.

Received: 11.07.2017.
