

Г. В. Каныгин, В. С. Корецкая (Кишинская)

АНАЛИТИЧЕСКОЕ КОДИРОВАНИЕ

Техники кодирования, широко применяемые в социологических исследованиях с помощью компьютерных программ, представлены в статье как специфический метод инструментального согласования социологом естественно-языковых свидетельств, получаемых от информантов. Мы утверждаем, что недостатком этих техник является инструментальная слабость компьютерных инструментов кодирования. Мы предлагаем методы аналитического кодирования, которые делают отношения, широко используемые в современном программировании (модульность, наследование, визуализация, компиляция и т. д.), практически доступны для социолога для целей аналитического представления материалов, получаемых от информантов. На примере концептуализации текстовых материалов мы демонстрируем, как предлагаемые методы позволяют социологу: оперировать утверждениями на естественном языке, представленными в текстовых данных; компилировать результаты кодирования в форме семантической сети; визуализировать весь процесс работы с высказываниями информантов.

Ключевые слова: Анализ качественных данных, кодирование, социологические методы, полевое исследование, компьютерные методы социологического исследования.

Введение

Анализ качественных данных (АКД, qualitative data analysis) (Online QDA 2018) является одним из актуальных направлений развития современных методов качественного исследования. Современный АКД осуществляется с помощью компьютерных пакетов Atlas.ti, MAXQDA, NVivo, xSight, Qualrus, Ethnograph и др., составляющих отдельный класс программных разработок. Общепринятое англоязычное название этого класса компьютерных программ — qualitative data analysis software (QDAS). В качестве русскоязычного аналога мы будем использовать словосочетание «пакеты или программы АКД». Описание любого пакета АКД несложно найти в Интернете.

Специфическую функциональность пакетов АКД Келле предложил называть функциями кодирования и реконструирования данных (Kelle 1997). Льюинс и Силвер представили те же инструментальные возможности пакетов, как набор отдельных инструментов: поиска

(content searching), ассоциирования (linking), кодирования (coding), аннотирования (writing and annotating), запросов (querying), сетевого представления (mapping or networking) (Lewins, Silver 2007).

Сами разработчики указанных программ зачастую выделяют именно кодирование в качестве ключевого инструментального средства их пакетов (MaxQDA 2018; Quirkos 2018). Такое внимание к инструменту кодирования в составе пакета согласуется с мнением социологов, которые рассматривают кодирование ключевой аналитической процедурой всего анализа качественных данных (Strauss 1987; Pierre, Jackson 2014).

Сент Джон и Джонсон (St. John, Johnson 2000) суммировали доводы пользователей пакетов в пользу и против их применения в качественных исследованиях. За тридцатилетнюю историю развития, о которой можно судить по работе Вольски (Wolski 2018), пакеты АКД неоднократно подвергались сравнительному анализу своих инструментальных средств (Tesch 1990; Lewins, Silver 2007; Top 16 2018). По результатам конференции KWALON (2016) Эверс очертила направления дальнейшего развития пакетов АКД (Evers 2018).

С целью совершенствования функциональности пакетов АКД рассмотрим аргументы, которые фактически свидетельствуют об инструментальной слабости этих компьютерных программ. Джанкер пишет, что «они позволяют в компьютеризированном виде делать ту же работу, которая прежним поколением исследователей выполнялась с помощью карточек и маркеров» (Junker, 2012: 85). Предвосхитив эту критику, Ла Пелль еще в 2004 г. продемонстрировала, как функциональность программы АКД может быть воспроизведена с помощью офисного пакета (La Pelle 2004).

Мы относим инструментальную слабость пакетов АКД, в первую очередь, к средствам кодирования, лежащим в основе аналитической работы, выполняемой социологом с помощью пакета. Незрелость аппарата кодирования порождает методические затруднения, встречаемые социологами в ходе практической концептуализации, но не разрешаемые ими с помощью пакета. Говоря о результатах кодирования, Бэйзли отмечает «вариативность их интерпретаций разными исследователями, и даже одним и тем же исследователем в разные моменты времени» (Bazeley 2012). Подобное затруднение социолога, возникающее в ходе кодирования, можно также увидеть в словах Томпсона: «В качественном исследовании для социолога наиболее

трудной задачей является концептуальная часть анализа данных: идентификация смысловых единиц, соединение этих единиц в категории и, наконец, описание отношений между этими категориями» (Thompson 2002). Джанкер говорит, что «анализ качественных данных становится неизмеримо труднее, когда мы отходим от базовых операций кодирования» (Junker 2012).

Говоря об инструментальной слабости пакета АКД, мы опираемся на сравнение их ассистирующих возможностей с достижениями современного программирования, также получивших развитие в последние тридцать лет. В качестве опытного факта следует констатировать, что инструменты кодирования пакетов АКД на сегодня не предоставляют своим пользователям-социологам средств аналитической работы, которые доказали эффективность в тысячах программистских проектах. Такими «недопоставленными» для социолога возможностями мы считаем модульную организацию — ключевую особенность программистских конструкций, полиморфизм, инкапсуляцию, видимость и другие отношения, доступные программисту в повседневной аналитической работе.

В статье изложено усовершенствование традиционного аппарата кодирования, развивающее возможности аналитических преобразований именно естественно-языковых обозначений, возникающих в процессе АКД. Целью совершенствования является придание аппарату кодирования средств его модульной организации, включая полиморфизм, инкапсуляцию, видимость, типизацию и т. п. Принципиально важным пунктом такого совершенствования являются новые алгоритмы, компилирующие семантическую сеть, выражающую результаты кодирования.

Кодирование на принципах информационных технологий мы называем аналитическим кодированием (АК). Внедрение инструментальных средств такого кодирования в практику концептуального взаимодействия социолога и его информантов означало бы радикальное изменение аналитических возможностей исследователя. В частности, иначе была бы осознана его собственная роль: он перестал бы воспринимать себя как рассказчик (storyteller) (Charmaz 2000), а стал бы аналитиком, обобщающим результаты своей работы с информантами на современных принципах управления знаниями.

Практическая значимость развиваемого нами аналитического подхода к кодированию объясняется рядом причин. Во-первых,

полевая работа социолога в прикладном социологическом исследовании — это представительная модель совместного создания социального знания в условиях, при которых, с одной стороны, его соавторами являются рядовые носители естественного языка, не обладающие возможностью использовать специальные знания (в области моделирования, информатики и т. п.). С другой, — объемы и хитросплетения знания, сообщаемого информантами и обобщаемого исследователями, представляют собой принципиальную аналитическую трудность, отмечаемую самими социологами (см. выше). Во-вторых, разрешение этих теоретических трудностей на принципах современных информационных технологий применительно к «камерному», но предметно наглядному общению социолога и его информантов неминуемо означает создание оригинальных структурных механизмов. Функционирование таких механизмов не зависит от объемов знания или ситуаций их применения, что позволяет перенести предлагаемые инструментальные решения на сферу управления знаниями, бурно развивающуюся в настоящее время. В-третьих, в статье в терминах аналитического кодирования впервые полностью описан инструментальный аппарат, внедрение которого в технологии управления обществом обеспечивает «визуальную прозрачность» социальных решений, ограничивая тем самым возможности ответственных лиц использовать складывающиеся обстоятельства в своих интересах. В-четвертых, знакомясь с инструментальными средствами, описанными в данной статье, социолог получает возможность войти в круг проблемных вопросов построения социального знания, которые обсуждаются сегодня в социальной и компьютерной науке.

Первоначальная версия математических моделей, лежащих в основе аналитического кодирования, изложена в работе (Каныгин, Полтинникова 2016). Каким образом методы аналитического кодирования позволяют концептуализировать «естественно-научный» и «социологический» тексты, продемонстрировано в (Каныгин, Полтинникова, Корецкая 2017). В данной статье детально разобран и обоснован новый алгоритм, представляющий результаты кодирования в виде семантической сети. Этот алгоритм программно реализован в составе созданного нами онторедатора *Diagogue* (2018). Все материалы концептуализации выложены на нашем постоянно меняющемся сайте (ГКООМ 2018). На стадии отрисовки результатов работы алгоритма использована программа *Graphviz 2.38* (2018).

Структурные механизмы аналитического кодирования

Изложение базовых идей АК мы приводим на примере концептуализации известного текста Терри, который представлен по ссылке (Terry text 2018). Для демонстрации техник АК нам достаточно совсем небольшого текстового свидетельства. Поэтому с целью упрощения изложения мы рассматриваем только одну сюжетную линию указанного текста.

При изложении нашего подхода мы опираемся на известные техники АКД — первичное кодирование, связывание кода с фрагментом свидетельства информанта, категоризацию кодов и др. Так же как в АКД, техники АК предназначены для аналитического выражения социологом потока свидетельства информанта. Однако в отличие от известных пакетов АКД результаты концептуализации, выполняемой социологом с помощью инструментальных средств АК, представлены семантической сетью оригинального вида. Предъявим новые инструментальные возможности АК путем практического преобразования указанного свидетельства в семантическую сеть.

Двойные обозначения

Прежде всего на фазе первичного кодирования социолог в соответствии с предлагаемым подходом использует нетрадиционную технику двойных обозначений.

Аналогично коду в АКД элементарной аналитической единицей в АК является естественно-языковое обозначение (словосочетание естественного языка), называемое понятием или концептом. Однако в отличие от АКД любое прикладное применение понятий социологом осуществляется в АК не иначе как с помощью двойных обозначений. Двойное обозначение — это пара понятий или кодов в традиционном понимании. Первое понятие пары двойного обозначения называется термином, второе — контекстом.

Важно, что у двойного обозначения всегда существует смысловая интерпретация. Термин — это традиционное обозначение социологом чего-либо, выполняемое в АКД с помощью кода. Контекст — это обозначение самим социологом условия, при котором он использует термин как дескриптор чего-либо. И термин, и контекст — это понятия, которые вводит человек, обозначая отдельные аспекты интуитивно понимаемой им ситуации.

Рассмотрим, что такое двойное обозначение применительно к свидетельству Терри. Обозначим текст этого свидетельства с помощью нашего собственного понятия РАЗМЫШЛЕНИЯ ТЕРРИ¹. При этом обозначим понятием МОЯ КОНЦЕПТУАЛИЗАЦИЯ условия, при которых мы выполняем нашу работу по аналитическому воспроизведению текста. При необходимости в дальнейшем мы сможем дополнительно разъяснить, что мы имеем в виду под такими условиями, используя понятие МОЯ КОНЦЕПТУАЛИЗАЦИЯ в качестве отправного пункта. Но уже теперь мы должны констатировать, что получили первое двойное обозначение, имеющее только что предложенную естественно-языковую интерпретацию. Такое двойное обозначение будем записывать как пару <РАЗМЫШЛЕНИЯ ТЕРРИ, МОЯ КОНЦЕПТУАЛИЗАЦИЯ>².

Эта пара станет единственным корневым узлом формируемой семантической сети (см. рис. 6). Сеть мы будем формировать постепенно, собирая в нее все результаты нашей аналитической работы с текстом информанта. Согласно традиции АКД начнем с первичного кодирования свидетельства.

В данной статье, выполняя первичное кодирование, для упрощения изложения непривычных техник двойных обозначений мы будем использовать в качестве понятий непосредственно текстовые фрагменты из свидетельства. Тем самым мы отстранимся от проблемы неизбежной модификации смыслов информанта по причине их переобозначения социологом с помощью собственных кодов (АКД) или понятий (АК).

В соответствии с текстом информанта вводим первый первичный код КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ. При этом в соответствии со здравым смыслом и структурными особенностями АК мы должны указать условия, при которых мы вводим указанный первичный код. Такими условиями, на наш взгляд, являются реальные обстоятельства, в которых мы проводим концептуализацию. Чуть выше мы их уже обозначили с помощью пары понятий <РАЗМЫШЛЕНИЯ ТЕРРИ, МОЯ КОНЦЕПТУАЛИЗАЦИЯ>. Таким образом, до момента введения концепта КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ мы уже придали смысл двум понятиям.

¹ В рамках статьи будем обозначать единичное понятие с помощью словосочетания, набранного заглавными буквами.

² Одна и та же пара понятий может представлять собой результат кодирования социологом свидетельства информанта и узел формируемой семантической сети. В первом случае мы используем для обозначения пары треугольные скобки <>, во втором — фигурные {}.

Тем самым в момент введения концепта КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ мы располагаем двумя понятиями, введенными нами самими на основе слов Терри. Учтем это обстоятельство, связав вводимый концепт с одним из этих двух понятий в виде двойного обозначения <КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, РАЗМЫШЛЕНИЯ ТЕРРИ>. Таким образом, мы получили первый первичный двойной код, который связывает конкретную формулировку Терри с тем, что мы сами обозначили как его размышления. Естественно, что такая связь существует явно только в наших аналитических построениях³.

Таким образом, с помощью двойных обозначений мы стремимся оградить аналитическую работу социолога от утверждений ad hoc. Новые понятия аналитик может вводить только после того, как уже сам обозначил условия, при которых осуществляет концептуальные нововведения. Эти обозначения условий, которые мы называем контекстами, могут быть также названы областями определенности вновь вводимых понятий.

Продолжая первичное кодирование, основываясь на смыслах текста Терри, вводим концепт — СТАНОВИТСЯ ОДИНОКО (в лексике АКД — первичный код). Однако тот же текст ясно подсказывает, что «становится одиноко» при условии, что переехал в свой дом. Поэтому вводим двойное первичное обозначение в виде пары термин-контекст <СТАНОВИТСЯ ОДИНОКО, КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ>.

Следуя тексту Терри, нетрудно создать термин НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ, который передает ощущения Терри, обусловленные его вселением в новый дом. Таким образом, в рамках первичного кодирования возникает двойной код <НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ, КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ>. КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ — это контекст, в котором осмыслены, согласно свидетельству, уже два введенных нами термина — СТАНОВИТСЯ ОДИНОКО и НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ.

Первичное кодирование в виде двойных кодов — это процесс создания базовой структуры, лежащей в основе модульной организации аналитических конструкций самого социолога. Элементарным

³ Выбор в качестве контекста РАЗМЫШЛЕНИЯ ТЕРРИ, а не МОЯ КОНЦЕПТУАЛИЗАЦИЯ является нашим авторским произволом, который упрощает изложение основных идей аналитического кодирования.

аналитическим обозначением или единицей в этих конструкциях является понятие. Повторим, что в зависимости от того, на каком месте пары (двойного кода) социолог использует понятие, оно называется термином или контекстом. Существенно, что спецификация понятия по месту в паре влияет на его обработку при построении семантической сети, возникающей как итог аналитической работы социолога.

Как показывает выполненное нами кодирование, АК предполагает, что любой термин может появиться в составе пары не иначе, как в ассоциации с контекстом, относительно которого он рассматривается. Сначала мы ввели понятие РАЗМЫШЛЕНИЯ ТЕРРИ, а затем мы его использовали при указании контекста для вновь вводимого концепта КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ. Этот уже введенный концепт послужил контекстом для нового нашего обозначения СТАНОВИТСЯ ОДИНОКО. И т. д.

Ввиду того, что любые понятия, так же как коды в традиционном случае, возникают в результате действий человека, то двойное кодирование означает определенное правило, которому подчиняется концептуальная работа социолога. Правило прагматически понятно: социологу разрешено вводить собственные аналитические обозначения — понятия, — только предварительно указав условия их существования. Это условие контекстной определенности понятий существует в АК не только в виде общенаучного пожелания, а проверяется с помощью алгоритмической генерации семантической сети.

Но для того чтобы такая генерация, а с ней и контроль аналитических действий социолога стали возможными, ему необходимо выполнить первичное связывание созданных им двойных обозначений.

Первичное связывание двойных кодов

Связывание двойных обозначений в АК — это расширенный аналог привычных действий пользователя компьютера при создании папок на своем рабочем столе. Рабочий стол / Библиотеки / Видео / Документы / Музыка / Изображения и далее пользователь может неограниченно создавать нужные ему папки, следуя своим семантическим соображениям. Эти действия пользователя компьютера социолог воспроизводит в традиционном АКД при создании категорий, объединяющих первичные и вторичные коды в единые структуры графов.

Работа социолога при АК имеет два существенных инструментальных отличия от традиционного кодирования и структурирования

папок на компьютере. Во-первых, пользователь изначально вводит двойные обозначения. Поэтому и связывает он между собой пары понятий. Во-вторых, социологу разрешено связывать пары понятий только в двухуровневые графы с единственной корневой вершиной. Такие двухуровневые графы, называемые ветвлениями или локальными связями понятий, показаны применительно к уже выполненному кодированию на рис. 1–2.



Рис. 1. Начальное ветвление

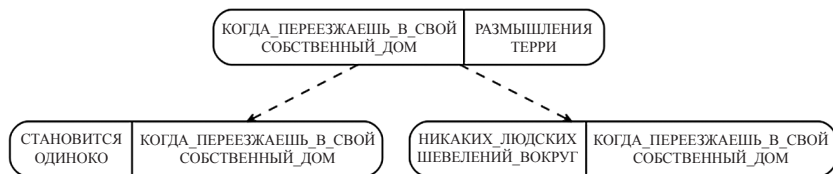


Рис. 2. Второе ветвление

Формально: начальное ветвление состоит из двух пар понятий. По смыслу: в размышлениях Терри выделен только один аспект — ситуация вселения в свой собственный дом.

Второе ветвление состоит из трех пар понятий. В ситуации вселения в свой собственный дом выделены два аспекта — одиночество и отсутствие звуков, издаваемых другими людьми.

В компьютерных терминах ограничение на уровень графов означает, что пользователь создает отношение между осмысленными парами понятий не сложнее, чем папка — ее подпапки. Более сложные структуры связей и семантическая сеть в целом создаются алгоритмически на основе множества ветвлений, появившимся благодаря концептуальным действиям социолога. Социолог создает ветвления по мере необходимости в ходе концептуализации.

Для введенных четырех пар понятий предлагаемое нами связывание выглядит так, как показано на рисунках 1 и 2. Для удобства восприятия введенные пары понятий пронумерованы от 1 до 4.

Человеко-машинное разделение «концептуального труда» социолога и компьютера имеет обоснование, возникающее благодаря двойному кодированию. С одной стороны, выполняемое социологом связывание пар понятий по аналогии с созданием папок не нагружается какой-то привходящей семантикой, в частности, не трактуется как отношение «часть-целое» или «общее-частное». Так же как в случае папок на компьютере, социолог в АК вводит все понятия в рамках конкретной задачи концептуализации. А каждую возникающую двухуровневую структуру «папок», обозначенных парами понятий, трактует по своему усмотрению.

С другой стороны, глядя на графы, можно заметить двойную «семантическую определенность» понятий КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ (рис. 1), СТАНОВИТСЯ ОДИНОКО, НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ (рис. 2). Рассмотрим, в чем она состоит на примере концепта КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ. Как объяснено выше, в составе пары 2 (см. рис. 1 и 2) это понятие имеет условием своего существования контекст РАЗМЫШЛЕНИЯ ТЕРРИ. В то же время очевидно, что существование папки является условием возникновения ее подпапок. В используемой нами двойной нотации папок это означает, что пара 1 является условием существования пары 2 (рис. 1), а пара 2, в свою очередь, является условием существования пар 3 и 4 (рис. 2). Эта зависимость, возникающая благодаря нашим привычным действиям с папками, является опытным фактом, который необходимо принять во внимание.

Таким образом, мы можем констатировать возникновение при конструировании ветвлений двух явно выраженных условий, которым должно отвечать существование вводимого нами понятия КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ. Первое условие, назовем его контекстным, задается двойным кодированием и явно выражается контекстом. Вторым неперменным условием существования понятия КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ служит наличие «над ним» «папки», «подпапкой» которой оно является⁴. Назовем это условие структурным.

⁴ Таким образом мы всегда разьясняем что-то через другое что-то.

Взгляд на рисунок 1 подсказывает, что в рассматриваемом случае обозначением контекстного и структурного условий, при которых существует понятие КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, является одно и то же понятие РАЗМЫШЛЕНИЯ ТЕРРИ.

Отметим, что структуры, во-первых, пары <КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, РАЗМЫШЛЕНИЯ ТЕРРИ> и, во-вторых, ветвления для пары <РАЗМЫШЛЕНИЯ ТЕРРИ, МОЯ КОНЦЕПТУАЛИЗАЦИЯ> мы вводим независимо друг от друга. Поэтому совпадение контекстного и структурного условий, в которых в наших построениях существует понятие КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, вовсе не является обязательным. И то, что эти обозначения совпали благодаря нашим действиям, мы трактуем как алгоритмическую проверку таких действий при введении понятий и их локальном связывании.

Графы рисунков 1 и 2 несут в себе ровно те смыслы, которые, как мы думаем, видит в рассказе информанта любой носитель языка. В размышлениях Терри одним из сюжетов является описание происходящего с ним в новом доме. Это отражено в том, что «папка» РАЗМЫШЛЕНИЯ ТЕРРИ имеет «подпапку» КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ. В свою очередь, «подпапка» КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ имеет «подпапки» СТАНОВИТСЯ ОДИНОКО и НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ. Тем самым рисунок 2, воспроизводя естественно-языковые формулировки информанта, выражает его же утверждение: оказавшись в собственном доме, Терри ощущает одиночество и отсутствие шумов, издаваемых другими людьми. В АКД выстраивание социологом отношений между первичными кодами с помощью вторичных кодов называется категоризацией.

Почему мы предлагаем социологу строить множество отдельных ветвлений, а не создавать единую структуру графа наподобие иерархии папок или кодов? Как сказано выше, социолог при концептуализации с помощью АК вводит одни понятия (термины) относительно других понятий (контекстов). При этом любой контекст должен быть введен социологом до того момента, как он использует его для определения термина. Такое требование предопределения понятия до его использования лежит в основе современных аналитических методов. Например, в компьютерных науках: прежде чем программист сможет использовать переменную в составе оператора, он должен идентифицировать ее в соответствующем блоке.

Организуя концептуальную работу социолога путем создания локальных связей понятий, мы получаем возможность алгоритмического контроля соблюдения им этого общенаучного требования в своих практических концептуальных действиях. Человек лишается возможности вводить свои обозначения *ad hoc*, он оказывается «концептуально обязанным» учитывать то, что уже введено им в качестве обозначений. При этом такая «концептуальная обязанность или ограничение» возникает не в силу чьего-то недоброжелательства, а по причине собственных, ранее сделанных «концептуальных заявлений».

При этом ключевое отслеживание, какие понятия могут играть роль контекстов для конкретного термина, осуществляется с помощью алгоритмически конструируемой семантической сети. Для того чтобы понять, каким образом семантическая сеть позволяет упорядочить концептуальные действия, осуществляемые социологом с помощью ветвлений, рассмотрим основную идею ее построения на основе уже введенных нами локальных связей понятий, показанных на рисунках 1 и 2.

Алгоритмы аналитического кодирования

Итак, мы ввели четыре пары понятий и связали их между собой с помощью двух наглядных графов, которые в зависимости от своих привязанностей читатель волен интерпретировать как папки-подпапки, категория-подкатегория, часть-целое и т. д. Как при введении пар, так и при их связывании мы руководствовались смыслами информанта и общенаучным требованием предварительного явного указания условий, при которых проводится обозначение этих смыслов. Соблюдение нами этого требования отслеживается алгоритмами АК следующим образом.

Мы выбираем пару понятий, связь которой мы хотим выстроить с помощью алгоритма. В данном случае это <РАЗМЫШЛЕНИЯ ТЕРРИ, МОЯ КОНЦЕПТУАЛИЗАЦИЯ>. Эта пара единственным образом определяет ветвление, созданное нами в результате первичного связывания (рис. 1). Алгоритм делает выбранную пару первым (корневым) узлом {РАЗМЫШЛЕНИЯ ТЕРРИ, МОЯ КОНЦЕПТУАЛИЗАЦИЯ} формируемой семантической сети. Далее он рассматривает все ее связи с другими парами, задаваемые ветвлением, созданным нами к моменту запуска алгоритма⁵.

⁵Эти пары еще не входят в состав формируемой семантической сети.

В нашем случае таких пар всего одна — <КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, РАЗМЫШЛЕНИЯ ТЕРРИ> (пара 2 на двух рисунках 2 и 3).

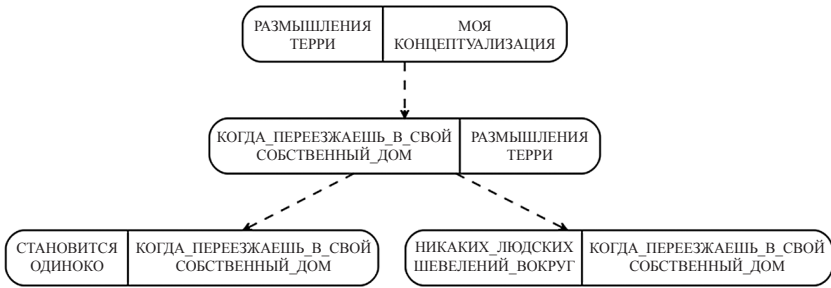


Рис. 3. Алгоритмическое связывание двух ветвлений: простейший случай генерации семантической сети

Алгоритм готов присоединить эту найденную пару к корневой паре семантической сети, но при соблюдении контекстного и структурного условий. Причем в развитие наших предыдущих рассуждений алгоритм трактует структурное условие более развернуто. Структурным условием алгоритмического включения найденной пары в состав семантической сети служит существование ее контекста в сформированной части графа этой сети.

В нашем случае это условие выполняется, так как РАЗМЫШЛЕНИЯ ТЕРРИ — это как раз одно из понятий корневого узла. Поэтому сеть автоматически обретает второй узел {КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, РАЗМЫШЛЕНИЯ ТЕРРИ} (см. рис. 3).

Теперь алгоритм проверяет связи пары <КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, РАЗМЫШЛЕНИЯ ТЕРРИ> только что присоединенного узла с другими введенными нами парами. Эти связи на данном этапе нашей концептуализации текста Терри задаются единственным ветвлением, показанном на рисунке 2. С помощью двойных обозначений в качестве «подпапок» для «папки» <КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, РАЗМЫШЛЕНИЯ ТЕРРИ> в этом ветвлении указаны две пары <СТАНОВИТСЯ ОДИНОКО, КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ> и <НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ, КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ>. Вновь перед

тем, как присоединить каждую из этих пар в качестве нового узла формируемой семантической сети, расположенного ниже только что созданной второй вершины {КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ, РАЗМЫШЛЕНИЯ ТЕРРИ}, алгоритм проверяет, присутствует ли контекст присоединяемой пары в уже построенной части семантической сети⁶.

Вновь оказывается, что один и тот же контекст КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ уже присутствует во втором построенном узле семантической сети. Поэтому ко второму узлу автоматически присоединяются два новых узла {СТАНОВИТСЯ ОДИНОКО, КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ} и {НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ, КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ}. Полученный результат показан на рисунке 3.

Обратим внимание, что алгоритм работает в соответствии с указанным общенаучным требованием: он проверяет, присутствует ли уже в результатах концептуализации тот контекст, с помощью которого пользователь определяет понятие в составе очередного узла семантической сети. Упорядоченность существования понятий во времени для пользователя воспроизводится в ходе его работы (см. выше, как мы строили понятия и их локальные связи). Алгоритм, в свою очередь, эмулирует эту упорядоченность за счет семантической сети, генерируемой им по результатам текущей работы пользователя.

В случае, если пользователь в ходе своей аналитической работы отклонится от правила включенности контекста двойного обозначения в сформированную часть семантической сети, то это скажется на структуре последней. Например, если ветвление, показанное на рисунке 2, преобразовать в ветвление рисунка 4 (в паре 4 рисунка 2 контекст КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ заменен на контекст СТАНОВИТСЯ ОДИНОКО), то после соответствующей автоматической генерации графа он будет иметь вид, показанный на рисунке 5. Пара <НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ, СТАНОВИТСЯ ОДИНОКО> не войдет в состав семантической сети, так как контекст СТАНОВИТСЯ ОДИНОКО не представлен в построенной части графа на момент присоединения этой пары.

⁶ Точнее, проверяется не вся построенная часть семантической сети, создаваемой в виде дерева, а только та ее ветка, для которой проводится алгоритмическое присоединение рассматриваемого узла.

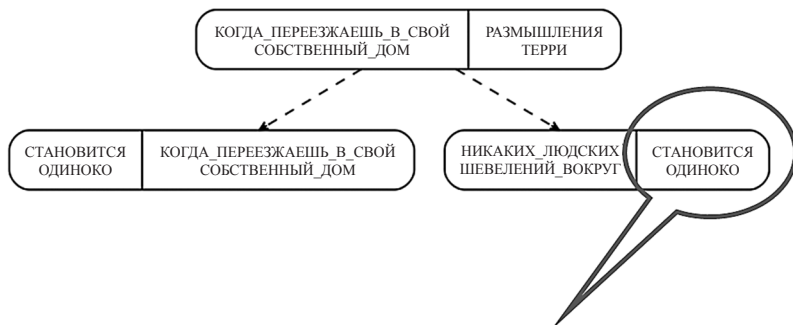


Рис. 4. Ветвление с парой понятий, контекст которой не согласуется со структурой формируемой семантической сети

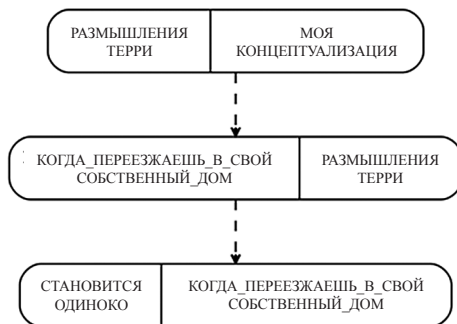


Рис. 5. Алгоритмическое связывание ветвлений: пара <НИКАКИХ ЛЮДСКИХ ШЕВЕЛЕНИЙ ВОКРУГ, СТАНОВИТСЯ ОДИНОКО> отсеяна из-за структурной несогласованности контекста (ср. рис. 3)

Как видно из наших построений, алгоритмическая проверка двойного кодирования относится только к аналитическим действиям социолога и никак не отвечает на вопрос, насколько свидетельства информанта соотносятся с реальностью, анализируемой исследователем. Фактически проверяется связность собственных аналитических конструкций социолога.

Результат кодирования в виде семантической сети

Пояснив основную идею алгоритмов АК, генерирующих семантические сети по результатам двойного кодирования, закончим

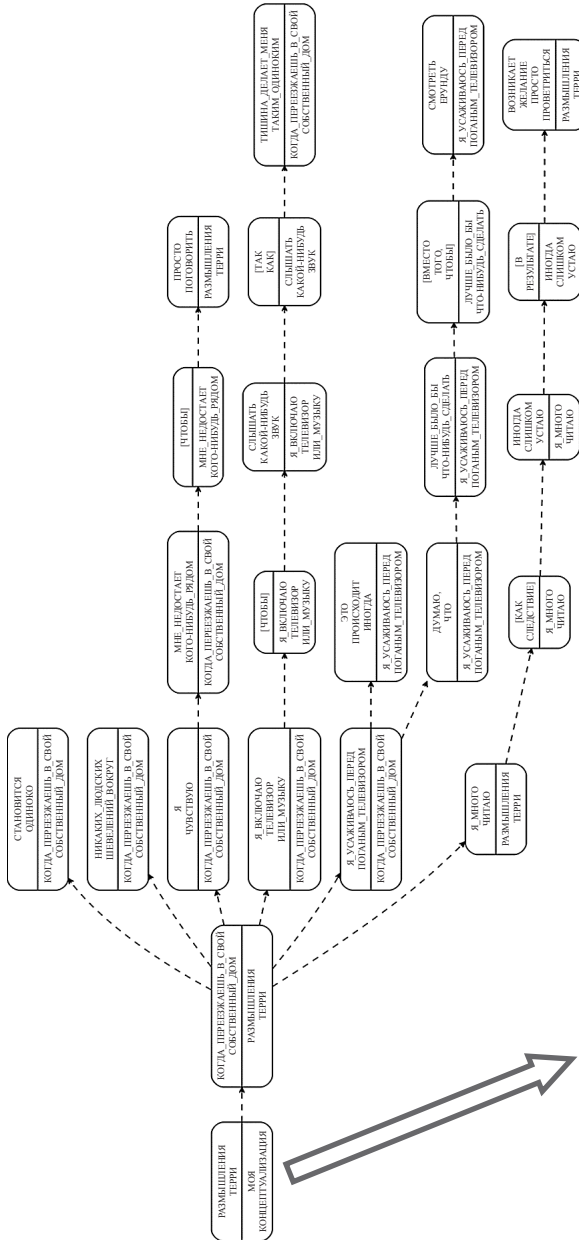
первичное двойное кодирование и локальное связывание его результатов для текста Терри. В целях экономии места статьи полученный итог предъявлен по следующим ссылкам: множество всех введенных нами понятий, называемое словарем (ГКООМ Словарь 2018), и множество всех созданных нами ветвлений, называемое тезаурусом (ГКООМ Тезаурус 2018).

Исходный текст и алгоритмически сгенерированная семантическая сеть свидетельства Терри, выполненная по результатам нашей концептуализации этого текста, показана на рис. 6. Фактически на рисунке представлена одна из многих возможных концептуализаций текста Терри. Каждая из них возникает на основе принятия автором концептуализации явных или неявных допущений. Скажем, одним из наших допущений было использование в качестве аналитических обозначений собственных формулировок информанта. Выбор допущений полностью относится к компетенции аналитика. Роль АК состоит в контроле связности результатов концептуализации, получаемых аналитиком в рамках допущений, принятых им самим.

Из рисунка видно, что граф семантической сети формируется алгоритмами АК в виде дерева, составленного из пар понятий. Такой вид наглядно демонстрирует цепочки естественно-языковых рассуждений, фиксируемых в тексте информанта. При интерпретации этих рассуждений полезно иметь в виду наше исходное толкование обозначений с помощью пар. Первое понятие в узле — это термин или традиционный первичный код. Второе понятие — это контекст, т. е. обозначение условий или областей существования термина, зафиксированных нами на основании слов информанта.

Таким образом, в терминах современного АКД построенная нами семантическая сеть может трактоваться как первичное кодирование, выполненное с помощью словесных обозначений самого информанта. Но с принципиальным отличием — указанием областей существования кодов с помощью самих кодов. Это отличие обеспечивается функционированием принципиально нового механизма связывания аналитических обозначений, отсутствующим в современном АКД.

В связи с интерпретацией построенного графа в виде цепочек рассуждений информанта поясним еще одну его особенность. Наряду с формулировками информанта мы использовали собственные обозначения, которые играют роль языковых связей (служебных слов), обеспечивающих для носителя языка «гладкость» переходов между отдельными аспектами ситуаций, которые мы подсмотрели у Терри.



Когда ты переедешь в свой собственный дом, то тебе становится одиноко. Никаких людских шевелений вокруг. Я чувствую, что мне недостает кого-нибудь рядом, чтобы просто поговорить. Я включаю телевизор или музыку, для того чтобы слышать какой-нибудь звук, тишина делает меня таким одиноким. Иногда я усаживаюсь перед поганым телевизором и думаю, что лучше было бы что-нибудь сделать, вместо того чтобы смотреть эту ерунду. Я много читаю, но иногда слишком устаю, и возникает желание просто проветриться.

Рис. 6. Семантическая сеть, сгенерированная в человеко-машинном режиме по свидетельству информанта

К числу таких понятий-связок мы отнесли ЧТОБЫ, ВМЕСТО ТОГО, ЧТОБЫ, В РЕЗУЛЬТАТЕ, КАК СЛЕДСТВИЕ, ТАК КАК. Все они показаны на графе в квадратных скобках.

Обсуждение

Главным инструментальным достоинством АК являются структурные механизмы, которые, с одной стороны, согласно традиции АКД, позволяют социологу кодировать свидетельства информантов с помощью естественно-языковых обозначений (кодов), с другой стороны, расширяют инструментарий социолога, давая в его распоряжение возможность устанавливать между кодами специфические отношения, доказавшие свою эффективность в компьютерных науках — модульность, видимость, полиморфизм и другие.

Компьютерные инструменты АК

Визуализация. Хотя АК предоставляет пользователю инструментальные возможности современного программирования, но от социолога не требуется трудоемкого обучения формальному синтаксису. Все связи между аналитическими единицами АК — понятиями — формируются пользователем непосредственно в виде графов. Результат работы человека также представлен наглядно — семантической сетью.

Модульность. Семантическая сеть, получаемая социологом как результат концептуализации свидетельства информанта, имеет модульную организацию. В качестве модуля мы рассматриваем ветвление, структура которого определяет для любого понятия его существование в семантической сети. Описанная модульная организация инструмента кодирования лежит в основе других отличительных аналитических возможностей АК, делающих его похожим на программирование.

Компиляция. Семантическая сеть представляет собой единую структуру знания, отдельные аспекты которого сообщаются информантами и аналитически фиксируются социологом. Говоря формально, социологическое знание, получаемое с помощью АК, алгоритмически компилируется в виде графа специального вида на основании множества отдельных ветвлений.

Социологические традиции АК

АКД был придуман для ассистирования социологу в его полевой работе с информантами. Существенной особенностью полевой

работы является применение как социологом, так и информантами естественно-языковых высказываний в качестве описаний социального явления. Продолжая эту традицию концептуализации человеческих мнений, высказываемых информантами в привычной словесной форме, инструментальные средства АК построены полностью на естественно-языковых обозначениях в качестве основы своего аппарата.

Алгоритмы АК не используют операций с числовыми значениями. Алгоритмы АК представляют собой средства согласования аналитиком естественно-языковых высказываний, работающими на принципах модульной организации отношений между этими высказываниями. Для алгоритмов АК не важно, кто является автором высказываний — социолог или его информанты.

Масштабируемость семантической сети

Для словесной детерминации вводимых понятий мы использовали формулировки Терри. Каждая из этих формулировок может быть подвергнута дальнейшей детализации, исходя из нашего человеческого понимания слов информанта. Например, КОГДА ПЕРЕЕЗЖАЕШЬ В СВОЙ СОБСТВЕННЫЙ ДОМ может быть уточнено за счет введения понятий ДОМ, КОГДА, ПЕРЕЕЗЖАЕШЬ, СВОЙ и других. АК позволяет выполнить такую детализацию с помощью тех же приемов, которые описаны в статье.

Оперирование сетью практически не зависит от ее объемов. При создании сети любого объема пользователь оперирует отдельными структурами не сложнее двухуровневого графа с ясной локальной смысловой интерпретацией. Для алгоритмов не имеет значения, какое число таких структур породит пользователь, выражая требуемую ему семантику.

Оценивание результатов АК

Результатом концептуализации свидетельств информантов с помощью инструментов АК является семантическая сеть специального вида. Наличие исходного текста и семантической сети, все обозначения которой представляют собой естественно-языковые высказывания, позволяет любому носителю языка оценить, насколько социолог преуспел в аналитическом воспроизведении смыслов информанта. В частности, читатель имеет такую возможность относительно результатов нашей концептуализации текста Терри.

Заключение

АК является новым оригинальным аналитическим аппаратом, способным в случае своей полноценной программной реализации в составе современных пакетов АКД принципиально расширить их ассистирующие возможности.

Развитие аналитического аппарата АК в направлении создания средств типизации, прототипирования, полиморфизма, командной работы и других инструментов, известных в программировании, создает возможности разработки новых методов управления знаниями, ориентированных на моделирование социальных процессов, описываемых их участниками на естественном языке.

Источники

ГКООМ [Электронный ресурс] // Сайт проекта ГКООМ. 2018. URL: <http://coknowledge.ru/materials/> (дата обращения: 23.09.2018).

ГКООМ Словарь [Электронный ресурс] // Сайт проекта ГКООМ. 2018. URL: <http://coknowledge.ru/wp-content/uploads/2018/06/RussianDictionary.txt> (дата обращения: 23.09.2018).

ГКООМ Тезаурус [Электронный ресурс] // Сайт проекта ГКООМ. 2018. URL: <http://coknowledge.ru/wp-content/uploads/2018/06/RussianThesaurus.doc> (дата обращения: 23.09.2018).

Каныгин Г. В., Полтинникова М. С. Контекстно-ориентированные онтологические методы в социологии // Труды СПИИРАН. 2016. Вып. 48. С. 107–124.

Каныгин Г. В., Полтинникова М. С., Корецкая В. С. Опыт построения социального знания на основе компьютерных онтологических методов // Социологический журнал. 2017. Т. 23, № 3. С. 25–41.

Онторедатор Diagogue 2018: Графовые контекстно-ориентированные онтологические методы: программная реализация [Электронный ресурс] // Сайт проекта ГКООМ. 2018. URL: http://coknowledge.ru/wp-content/uploads/2017/02/GKOOM_Program.pdf (дата обращения: 23.09.2018).

Bazeley P. Regulating qualitative coding using QDAS? // Sociological Methodology. 2012. Vol. 42 (1). P. 77–78.

Charmaz K. Grounded Theory: Objectivist and Constructive Methods // Handbook of Qualitative Research / Eds. K. Norman, S. Denzin, Y. Lincoln. 2nd edition. Thousand Oaks, CA, 2000. P. 509–535.

Evers J. C. Current issues in qualitative data analysis software (QDAS): A user and developer perspective // The Qualitative Report. 2018. Vol. 23 (13).

P. 61–74. URL: <http://nsuworks.nova.edu/tqr/vol23/iss13/5> (access date: 12.09.2018).

Graphviz — graph visualization software [Electronic resource]. 2018. URL: <https://www.graphviz.org/> (access date: 12.09.2018).

Junker A. Optimism and Caution Regarding New Tools for Analyzing Qualitative Data // *Sociological Methodology*. 2012. Vol. 42 (1). P. 85–87.

Kelle U. Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data [Electronic resource] // *Sociological Research Online*. 1997. Vol. 2 (2). URL: <http://socresonline.org.uk/2/2/1.html> (access date: 12.09.2018).

KWALON Conference 2016: Reflecting on the future of QDA Software: Chances and Challenges [Electronic resource]. URL: <http://www.kwalon.nl/about-kwalon> (access date: 12.09.2018).

La Pelle N. Simplifying Qualitative Data Analysis Using General Purpose Software Tools // *Field Methods*. 2004. Vol. 16 (1). P. 85–108.

Lewins A., Silver C. Using Qualitative Software: A Step-by-Step Guide. London: Sage, 2007.

MaxQDA [Electronic resource]. [2018]. URL: <https://www.maxqda.com/qualitative-data-analysis-software> (access date: 12.09.2018).

Online QDA [Electronic resource]. [2018]. URL: <http://onlineqda.hud.ac.uk/> (access date: 12.09.2018).

Pierre E. St., Jackson A. Y. Qualitative Data Analysis after Coding // *Qualitative Inquiry*. 2014. Vol. 20 (6). P. 715–719.

Quirkos [Electronic resource]. [2018]. URL: <https://www.quirkos.com/learn-qualitative/qualitative-analysis-software.html> (access date: 12.09.2018).

St. John W., Johnson P. The pros and cons of data analysis software for qualitative research // *Journal of Nursing Scholarship*. 2000. Vol. 32 (4). P. 393–397.

Strauss A. L. Qualitative Analysis for Social Scientists. Cambridge, MA: Cambridge University Press, 1987.

Terry text [Electronic resource]. [2018]. URL: http://onlineqda.hud.ac.uk/Intro_QDA/phpechopage_titleOnlineQDA-Examples_QDA.php (access date: 12.09.2018).

Tesch R. Qualitative Research: Analysis types and software tools. New York: The Falmer Press, 1990.

Thompson R. Reporting the Results of Computer-assisted Analysis of Qualitative Research Data [Electronic resource] // *Forum: Qualitative Social Research: [On-line Journal]*. 2002. Vol. 3, No. 2. URL: <http://www.qualitative-research.net/index.php/fqs/article/view/864/1878> (access date: 12.09.2018).

Top 16 Qualitative Data Analysis Software [Electronic resource]. [2018]. URL: <https://www.predictiveanalyticstoday.com/top-qualitative-data-analysis-software/> (access date: 21.01.2018).

Wolski U. The history of the development and propagation of QDA software [Electronic resource] // *The Qualitative Report*. 2018. Vol. 23 (13). P. 6–20. URL: <http://nsuworks.nova.edu/tqr/vol23/iss13/2> (access date: 12.09.2018).

References

Bazeley P. Regulating qualitative coding using QDAS? *Sociological Methodology*, 2012, 42 (1), pp. 77–78.

Charmaz K. Grounded Theory: Objectivist and Constructive Methods. *Handbook of Qualitative Research*. Eds. K. Norman, S. Denzin, Y. Lincoln. 2nd edition. Thousand Oaks (CA), 2000, pp. 509–535.

Evers J. C. Current issues in qualitative data analysis software (QDAS): A user and developer perspective. *The Qualitative Report*, 2018, 23 (13), pp. 61–74, URL: <http://nsuworks.nova.edu/tqr/vol23/iss13/5> (access date: 12.09.2018).

GKOOМ // *Sajt proekta GKOOМ [Site of the project GKOOМ]*, 2018, URL: <http://coknowledge.ru/materials/> (access date: 23.09.2018). (In Russian)

GKOOМ Slovar' [GKOOМ Dictionary] // *Sajt proekta GKOOМ [Site of the project GKOOМ]*, 2018, URL: <http://coknowledge.ru/wp-content/uploads/2018/06/RussianDictionary.txt> (access date: 23.09.2018). (In Russian)

GKOOМ Tezaurus [GKOOМ Thesaurus] // *Sajt proekta GKOOМ [Site of the project GKOOМ]*, 2018, URL: <http://coknowledge.ru/wp-content/uploads/2018/06/RussianThesaurus.doc> (access date: 23.09.2018). (In Russian)

Graphviz — *graph visualization software*. 2018, URL: <https://www.graphviz.org/> (access date: 12.09.2018).

Junker A. Optimism and Caution Regarding New Tools for Analyzing Qualitative Data. *Sociological Methodology*, 2012, 42 (1), pp. 85–87.

Kanygin G. V., Poltinnikova M. S. Kontekstno-orientirovannye ontologicheskie metody v sotsiologii [Context-oriented ontological methods in sociology]. *Trudy SPIIRAN [Proceedings of SPIIRAS]*, 2016, iss. 48, pp. 107–124. (In Russian)

Kanygin G. V., Poltinnikova M. S., Koretskaya V. S. Opyt postroeniya sotsial'nogo znaniya na osnove komp'yuternykh ontologicheskikh metodov [Experience in Building Social Knowledge Based on Computer Ontological Methods]. *Sotsiologicheskij zhurnal [Sociological journal]*, 2017, vol. 23, no. 3, pp. 25–41. (In Russian)

Kelle U. Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data. *Sociological Research Online*, 1997, 2 (2), URL: <http://socresonline.org.uk/2/2/1.html> (access date: 12.09.2018).

KWALON Conference 2016: *Reflecting on the future of QDA Software: Chances and Challenges*. URL: <http://www.kwalon.nl/about-kwalon> (access date: 12.09.2018).

La Pelle N. Simplifying Qualitative Data Analysis Using General Purpose Software Tools. *Field Methods*, 2004, 16 (1), pp. 85–108.

Lewins A., Silver C. *Using Qualitative Software: A Step-by-Step Guide*. London, Sage, 2007.

MaxQDA. [2018], URL: <https://www.maxqda.com/qualitative-data-analysis-software> (access date: 12.09.2018).

Online QDA. [2018], URL: <http://onlineqda.hud.ac.uk/> (access date: 12.09.2018).

Ontoredaktor Diagogue 2018: Grafovye kontekstno-orientirovannye ontologicheskie metody: programmaya realizatsiya [Ontoeditor Diagogue 2018: Graph context-oriented

ontological methods]. *Sajt projekta GKOOM [Site of the project GKOOM]*, 2018, URL: http://coknowledge.ru/wp-content/uploads/2017/02/GKOOM_Program.pdf (access date: 23.09.2018). (In Russian)

Pierre E. St., Jackson A. Y. Qualitative Data Analysis after Coding. *Qualitative Inquiry*, 2014, 20 (6), pp. 715–719.

Quirkos. [2018], URL: <https://www.quirkos.com/learn-qualitative/qualitative-analysis-software.html> (access date: 12.09.2018).

St. John W., Johnson P. The pros and cons of data analysis software for qualitative research. *Journal of Nursing Scholarship*, 2000, 32 (4), pp. 393–397.

Strauss A. L. *Qualitative Analysis for Social Scientists*. Cambridge (MA), Cambridge University Press, 1987.

Terry text. [2018], URL: http://onlineqda.hud.ac.uk/Intro_QDA/phpechopage_titleOnlineQDA-Examples_QDA.php (access date: 12.09.2018).

Tesch R. *Qualitative Research: Analysis types and software tools*. New York, The Falmer Press, 1990.

Thompson R. Reporting the Results of Computer-assisted Analysis of Qualitative Research Data. *Forum: Qualitative Social Research: [On-line Journal]*, 2002, vol. 3, no. 2, URL: <http://www.qualitative-research.net/index.php/fqs/article/view/864/1878> (access date: 12.09.2018).

Top 16 Qualitative Data Analysis Software. [2018], URL: <https://www.predictiveanalytics.today.com/top-qualitative-data-analysis-software/> (access date: 21.01.2018).

Wolski U. The history of the development and propagation of QDA software. *The Qualitative Report*, 2018, 23 (13), pp. 6–20. URL: <http://nsuworks.nova.edu/tqr/vol23/iss13/2> (access date: 12.09.2018).