# Каныгин Г.В.

# Методы управления знаниями в анализе качественных данных

Техники кодирования, широко используемые в качественном исследовании, рассмотрена в статье как специфический метод концептуализации нечисловой информации, осуществляемый с помощью пакета анализа качественных данных (АКД). Отмечены недостатки современных инструментальных средств кодирования. Предложено в качестве инструмента кодирования в пакетах АКД использовать графовые контекстно-ориентированные онтологические методы (ГКООМ) управления знаниями, сочетающие возможности современного программирования с доступностью для социолога. Рассмотрены приемы и особенности применения ГКООМ для целей кодирования качественных данных. На конкретных примерах концептуализации текстового свидетельства информанта продемонстрированы инструментальные преимущества ГКООМ. Обсуждены основные особенности ГКООМ – кодирование с учетом контекста, структурная декомпозиция свидетельства информанта, генерация концептуальной модели на основе множества смысловых аспектов свидетельства и др.

Анализ качественных данных, кодирование, социологические методы, графовые контекстно-ориентированные онтологические методы

## Введение

Анализ качественных данных (АКД, англоязычные аббревиатуры QDAS и CAQDAS [23]) является одним из актуальных направлений развития современных методов качественного исследования. Методы АКД существуют в виде программных пакетов Atlas.ti, MAXQDA, NVivo, xSight, Qualrus, Ethnograph и др., описания которых нетрудно найти в Интернете. Инструментальные средства пакетов представляют собой аппарат концептуализации нечисловых данных, выражающих свидетельства информантов о социальных процессах, интересующих социолога, и могут быть подразделены на отдельные инструменты – поиска (content searching), ассоциирования (linking), кодирования (coding), аннотирования (writing and annotating), запросов (querying), сетевого представления (mapping or networking) [21].

Специфическую функциональность методов АКД выражает инструмент кодирования [ibid.], называемый также функциями кодирования и реконструирования данных [19].

За время более чем тридцатилетнего развития в адрес пакетов АКД неоднократно высказывались критические замечания, основанные на прикладном применении пакетов. Как справедливо указывает R. Thompson, резюмируя работы своих предшественников, в случае, если данные собраны неудачно, то ни мощь программных средств пакета АКД, ни мастерство применяющего их социолога не могут сделать результаты всего исследования приемлемыми [26].

Однако следует заметить, что чуть ниже тот же автор высказывает ключевое для АКД мнение, которое не устарело за прошедшие годы: «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.» [ibid.].

А в таком случае оказывается, что вне зависимости от качества анализируемых данных важнейшей функцией инструментов АКД, в частности, кодирования, является помощь социологу при концептуализации собранных материалов. В этой связи показательны сомнения социологов в ассистирующих возможностях инструментальных средств пакетов АКД. Так, K. MacMillan and S. McLachlan указывают на существенное несоответствие между тем, что социолог понимает под термином «теория» в качестве результата своего исследования, и тем, что он может получить из собранных данных с помощью инструментов пакетов АКД [22].

Такие сомнения находят свою подтверждение в критике инструментальных недостатков пакетов АКД. Одним из последних проявлений этой критики являются суждения, высказанные в работах R. Franzosi, G. DeFazio, S. Vicari , G. Gibbs, A. Junker, M. White, M. Judd, S. Poliandri [16, 17, 18, 27].

Приведем слова A. Junker о пакетах АКД: «они позволяют в компьютеризированном виде делать ту же работу, которая прежним поколением исследователей выполнялась с помощью карточек и маркеров» [18, с. 85]. В подтверждение этих слов вспомним, что еще в 2004 году N. LaPelle продемонстрировала, как функциональность пакета АКД может быть воспроизведена с помощью текстового редактора [20].

Резюмируя высказанные замечания, можно сказать, что авторы «инструментальной» критики, во-первых, отдают должное широкому применению этих пакетов в социологических исследованиях. Во-вторых, разочарованы тем, что инструментальные средства концептуализации данных, представленные в пакетах АКД, фактически не выходят за пределы возможностей операций с текстом, доступных пользователям офисных программ.

Для расширения ассистирующих возможностей сегодняшних пакетов АКД мы предлагаем в их составе в качестве функций кодирования использовать методы управления знаниями специального вида, а именно предложенные нами ранее графовые контекстно-ориентированные онтологические методы (ГКООМ) [5, 25].

В данной работе показано, во-первых, каким образом ГКООМ могут быть применены в качестве инструмента кодирования текста информанта. Во-вторых, в чем состоят принципиальные достоинства ГКООМ по сравнению с функциями кодирования, реализованными в современных пакетах АКД.

## Основания концептуализации с помощью ГКООМ

Подчеркнем общую «двухуровневую» идею АКД, которая лежит в основе перехода от текста[[1]](#footnote-1) информанта к аналитическим построениям социолога. Во-первых, аналитик переобозначает что-либо (факты, события, переживания и т.п.), увиденное им в тексте информанта, с помощью собственных естественно-языковых обозначений. В случае традиционного кодирования с помощью инструментов пакета АКД такие обозначения называются кодами. При использовании ГКООМ с целью концептуализации свидетельства информанта указанные обозначения называются понятиями.

Во-вторых, аналитик устанавливает отношения между кодами или понятиями не с помощью естественного языка, а в явном виде (ассоциаций, графов, списков и др.). Последнее принципиально важно: таким образом в результате концептуализации удается устранить характерные недостатки естественно-языковых описаний. В их числе: метафоры, умолчания и другие [6].

Переобозначение в виде кодов или понятий и введение отношений между ними – это два аналитических действия, которые неизбежно деформируют свидетельство информанта. Для нивелирования таких деформаций мы хотели бы дать в распоряжение социолога инструментальные средства концептуализации, которые успешно зарекомендовали себя в компьютерных науках.

Примерами этих средств служат языки объектно-ориентированного [7] и аспектно-ориентированного [8] программирования. Достоинством современного программирования является мощный структурный аппарат, реализующий общенаучную идею контекстной обусловленности фактов, событий и других явлений, подвергаемых концептуализации.

Применение такого аппарата социологом для целей концептуализации свидетельства информанта назовем кодированием в стиле современного программирования. В качестве синонимов будем употреблять словосочетания аналитическое или концептуальное кодирование. Инструментальные средства, позволяющие пользователю осуществлять аналитическое кодирование при решении задач АКД, назовем инструментами аналитического кодирования.

Каким образом выглядит и функционирует аналитическое кодирование, осуществляемое с помощью ГКООМ, показано в данной статье. Наши методические предложения состоят в том, чтобы в качестве первого шага анализа текста аналитик переопределил бы его структурные связи с помощью концептуального кодирования. Такое переопределение, с одной стороны, приводит к устранению метафор и других двусмысленностей языка. С другой, - представляет текст и стоящий за ним «предмет» в явном структурном виде, что позволяет в последующем ввести аналитические методы анализа.

Предлагаемое переопределение или переосмысление высказываний респондента не следует лингвистическим методам анализа текста, построенным на алгоритмическом разборе его состава на отдельные части (словоформы, существительные, действия и т.п.) [9]. «Когнитивным механизмом» переосмысления оказывается сам исследователь. Ориентируясь на словесные выражения свидетельства, аналитик должен передать структуры смыслов информанта на основе неформального анализа отдельных аспектов свидетельства.

Возможность реализации наших предложений зависит от того, насколько концептуальное кодирование оказывается доступным для аналитика. Эта проблема доступности инструментальных средств современного программирования для неподготовленного пользователя решена в [5, 25] в виде ГКООМ. В частности, все аналитические построения социолога строятся на основе поименований, т.е. обозначений на естественном языке [11].

Само собой разумеется, что в процессе переосмысления собранных материалов с помощью ГКООМ социолог вправе обращаться к многочисленным методическим руководствам по концептуализации в социологических проектах [10, 12, 13, 15 и др.].

Нам представляется, что внедрение концептуального кодирования в арсенал социологических методов дает импульс их развитию в целом. Применительно к задачам АКД предлагаемые методы позволяет поставить под контроль аналитические построения самого социолога, отслеживая связность обозначений, вводимых им в процессе концептуализации.

### Идентификация смыслов информанта: введение объектов

Сначала представим несложный текст, на основании которого будем разворачивать пример концептуализации с помощью ГКООМ.

«Мой дом расположен на горе. В доме две комнаты. Одна из них - столовая, в которой находятся мойка, обеденный стол, шкаф и четыре стула. Другая – спальня. В ней стоят шкаф, кровать, письменный стол, печка и два стула. Первая комната больше второй. В большой комнате три окна, а в спальне – одно. Еще в доме есть прихожая, из которой лестница ведет на второй этаж, а окно выходит во двор. Второй этаж – это мансарда с двумя окнами.»

Стремясь воссоздать смыслы, распознаваемые нами у информанта, будем говорить, что они передаются в тексте в трех видах - объекта, его свойств и отношений. Отношения могут выражать соотнесенность как объектов между собой, так и объектов со своими свойствами. Будем определять, что такое объект, свойство и отношение строго на основе примера.

Начнем с выделения объектов. Ориентируясь на наше понимание всего текста, сделаем утверждение, что в нем описаны объекты, которые могут быть обозначены следующими понятиями (здесь и далее заглавными буквами обозначены элементарные единицы проводимой концептуализации - понятия):

ТЕКСТ ИНФОРМАНТА, ДОМ, СТОЛОВАЯ, ГОРА, мойка, обеденный стол, шкаф, стул1, стул2, стул3, стул4, СПАЛЬНЯ, шкаф, кровать, письменный стол, печка, стул1, Стул2, ОКНО1, ОКНО2, ОКНО3, ОКНО4, ОКНО5, ОКНО6, ОКНО7, ПРИХОЖАЯ, ЛЕСТНИЦА, ЭТАЖ1, ЭТАЖ2.

В качестве первого обоснования скажем, что понятия введены в соответствии с тем, что мы увидели в свидетельстве. Понятие ТЕКСТ ИНФОРМАНТА обозначает свидетельство в целом.

Теперь перейдем к объяснению принципиальной особенности ГКООМ, которая позволяет существенно уточнить представления об объекте. Обратим внимание, что нарратив информанта позволяет на основе здравого смысла (иначе говоря, нашего опыта или интуитивных оснований) указать условия, при которых согласно свидетельства существуют уже обозначенные нами объекты. Например, читая текст информанта, можно понять, что объект СТОЛОВАЯ существует в границах объекта ДОМ; один из стульев, обозначенных СТУЛ2, стоит в СПАЛЬНЯ и т.д.

Структурные модели ГКООМ [5, 25] устроены таким образом, чтобы помочь аналитику при введении формальных конструкций выявлять контекстную обусловленность любых утверждений информанта. Опираясь на текст свидетеля, аналитик побуждается структурными моделями ГКООМ фиксировать условия, при которых осмыслены те или иные утверждения информанта. Так, обозначение СТОЛОВАЯ//ДОМ указывает на то, что объект СТОЛОВАЯ существует только при условии существования объекта ДОМ. Знак // - это способ сократить рутинное естественно-языковое описание (читается «в контексте»).

Формально соотнесение объекта и условий его существования выполняется путем составления пар понятий. Первое понятие указывает объект, второе обозначает область его существования. Вводимые понятия, конечно, вызваны к жизни свидетельством информанта. Однако для выявления смыслов свидетельства социолог вправе выходить за непосредственные границы анализируемого свидетельства. Например, он может предположить существование некоего общего знания, разделяемого им и информантом. И концептуализировать это предположение с помощью понятий (см. далее введение понятия ГОРА). Такого сорта подход возлагает на аналитика отслеживание контекстов, в которых существуют объекты, засвидетельствованные информантом.

Обратим внимание, что именно пара понятий является описанием любого реального объекта. Скажем, не существует шкафа вообще. Тем самым, опираясь на конкретность здравого смысла и текст информанта, в котором указаны два места расположения шкафа, мы пишем ШКАФ// СТОЛОВАЯ и шкаф//СПАЛЬНЯ. При этом, несмотря на то, что использовано одно и то же понятие ШКАФ, указанные пары обозначают разные объекты!

Социолог может не сразу оценить, зачем нужны непривычные двойные обозначения. Двойное обозначение – это метод учета локальных смыслов, понимаемых при кодировании на основе текстовых утверждений информанта.

Локальны эти смыслы из-за того, что мы не знаем, в какие дальнейшие концептуальные построения будут включены фиксируемые нами объекты. Но любое последующее упоминание вводимых нами предметов должно учитывать, что они имеют границы своего существования, которые мы наблюдаем у информанта как исходный факт. И эти границы мы указываем в момент смыслового анализа текста в виде части самого обозначения объекта.

Именно возможность алгоритмически отслеживать контекстную обусловленность реальных объектов является принципиальным преимуществом методов, построенных на идеях управления знаниями. Таким образом существенно ограничиваются смысловые искажения, в изобилии возникающие в силу трудностей мысленного отслеживания семантических хитросплетений текста.

По ссылке [2] указан полный список пар, описывающих все объекты, обнаруженные нами в тексте информанта. Каждая из введенных пар – это обозначение чего-либо, воспринимаемого как объект согласно нашей интерпретации свидетельства. Ввиду новизны предлагаемого подхода рассмотрим подробнее, на каких основаниях осуществляется двойное кодирование ГКООМ.

### Объекты и здравый смысл

Текстовые обозначения понятий (слова, написанные заглавными буквами) имеют в своей основе словесные выражения информанта. Однако, ориентируясь на семантические связи текста, мы не воссоздаем буквально существительные, другие части речи или члены предложений, которые можно распознать в тексте. Скажем, «комната» и «мансарда» не воспроизведены в концептуальном кодировании, т.к. они уже введены как объекты СТОЛОВАЯ и ЭТАЖ2 в контексте ДОМ. Двор, упомянутый информантом, оставлен без внимания, т.к. целью нашего описания является именно дом. Однако в случае необходимости ничто не мешает дополнить наши обозначения соответствующим понятием, скажем, ДВОР.

Объекты появляются в концептуальной модели на основе здравого смысла, и гипотезы о том, что он одинаков для нас с информантом. Скажем, мы используем встреченные числительные для того, чтобы воспроизвести соответствующий объект указанное число раз. Поэтому формулировка информанта «четыре стула» превращается в четырехэлементное множество объектов, соотнесенных со столовой стул1//СТОЛОВАЯ, стул2//СТОЛОВАЯ, стул3//СТОЛОВАЯ, стул4//СТОЛОВАЯ. А утверждение «в ней стоят… два стула» применительно к спальне создает двухэлементное множество стул1//СПАЛЬНЯ, Стул2//СПАЛЬНЯ.

Еще один пример опоры на здравый смысл. Мы используем предложение «Первая комната больше второй» только в качестве указания на то, что в следующем предложении речь идет о столовой, в которой имеется три окна. Сопоставление размеров комнат никак иначе не отражено в концептуальной модели, т.к. и в тексте это количественное утверждение не используется.

Сообщение о числе окон мы преобразуем в аналитические обозначения таким образом, чтобы зафиксировать факт существования указываемых объектов. Основываясь на свидетельстве, мы не имеем никаких дополнительных сведений о том, какие это окна – выходящие на север или на юг, пластиковые или деревянные и т.д. Поэтому мы просто резервируем на будущее возможность таких уточнений, вводя семь понятий, с различающимися между собой названиями – ОКНО1..ОКНО7.

Аналогично утверждение «лестница ведет на второй этаж» позволяет нам засвидетельствовать в аналитическом виде существование двух этажей, каждый из которых является самостоятельным объектом. Таким образом появляются ЭТАЖ1 и ЭТАЖ2. Разумеется, может быть и третий этаж, но об этом у информанта ничего не сказано.

### Выбор контекста

В принципе в качестве контекста может использоваться любое понятие, лишь бы возникающая пара передавала конкретный смысл, который аналитик распознал благодаря информанту. Например, обозначив с помощью понятия ТЕКСТ ИНФОРМАНТА концептуализируемый нами текст, мы делаем вполне допустимой пару ДОМ//ТЕКСТ ИНФОРМАНТА. Она означает, что понятие ДОМ рассматривается как существующее применительно к конкретному тексту примера.

Обратим внимание, что «двойное кодирование» является операцией структурирования, которая возлагает на аналитика ответ на вопрос, в каких границах он будет рассматривать реально существующий объект. Например, один из стульев, засвидетельствованных информантом в столовой, мы обозначили как СТУЛ1//СТОЛОВАЯ. Однако наш опыт подсказывает, что СТУЛ1, существуя в столовой, существует также и в доме. Та и другая пары СТУЛ1//СТОЛОВАЯ и СТУЛ1//ДОМ не противоречат нашему опыту.

Вместе с тем этот опыт говорит, что реальный объект всегда существует лишь в одном экземпляре. Если мы не стремимся сконструировать какие-то специальные отношения, такая реальная уникальность побуждает нас выбрать только один из контекстов. На стадии введения объектов выбор контекста часто не имеет единственного решения. Дело в том, что критерием такого выбора является соответствие концептуальной модели свидетельству информанта в целом. А концептуальная модель на рассматриваемой фазе кодирования еще не построена.

Как быть социологу в условиях такой неопределенности? Он должен иметь в виду итерационный процесс построения концептуальной модели. На данной стадии он выбирает один из двух контекстов, затем алгоритмически строит концептуальную модель (см. ниже). И по результатам собственного анализа соответствия модели тексту социолог либо выбирает другой контекст, либо оставляет выбранный, либо выполняет какие-то иные действия, влияющие на концептуальную модель.

## Свойства объектов

В свидетельстве информанта наряду с объектами можно увидеть и другие описания. В частности, он рассматривает дом, состоящим из ряда помещений. Выразим этот наблюдаемый в тексте факт путем введения понятия СОСТОИТ ИЗ. Понятие СОСТОИТ ИЗ выражает собой связь между объектом ДОМ и содержащимися в нем объектами. В то же время СОСТОИТ ИЗ может трактоваться как свойство объекта ДОМ.

Контекстом для СОСТОИТ ИЗ выбрано понятие ДОМ, т.к. именно дом подразделяется в концептуальной модели на ряд помещений в соответствии с утверждениями информанта.

В свою очередь, свойство СОСТОИТ ИЗ предполагает указание частей, из которых что-либо состоит, например, ДОМ. Но это уже другая связь, устанавливаемая между свойствами и объектами, которая рассмотрена ниже.

Аналогично для объектов СПАЛЬНЯ и СТОЛОВАЯ мы вводим свойства СОДЕРЖИТ ВНУТРИ СЕБЯ, которое предполагает так же как СОСТОИТ ИЗ последующее перечисление «объектов», входящих в состав помещений, обозначенных как СПАЛЬНЯ и СТОЛОВАЯ.

Обратим внимание, что СОСТОИТ ИЗ и СОДЕРЖИТ ВНУТРИ СЕБЯ для носителя языка означают примерно одно и то же. Делая эти словосочетания понятиями, мы игнорируем это интуитивно ощущаемое сходство, основанное на владении естественным языком. Такое игнорирование возникает из-за нашего исходного обязательства прослеживать связи понятий явно, а не с помощью интуиции, опирающейся на неясные механизмы словотворчества.

Утверждение информанта «лестница ведет на второй этаж» побуждает нас ввести свойство ВЕДЕТ НА, которым обладает лестница. Аналогично появляются свойства РАСПОЛОЖЕН НА и ИМЕЕТ ОКНА, относимые далее к объектам, определенным на основе свидетельства.

Нельзя забывать, что свойства, так же как объекты, имеют контекстную определенность. Следуя тексту информанта, можно сказать, что, например, ИМЕЕТ ОКНА определены в следующих контекстах: ЭТАЖ1, ЭТАЖ2, СТОЛОВАЯ, СПАЛЬНЯ, ПРИХОЖАЯ.

Полный список контекстно определенных свойств приведен по ссылке [3].

Обратим внимание, что свойства могут быть разными. Например, свойством этажей в доме может быть их объединение в одну категорию. Ничто не мешает обозначить эту категорию как новое отдельное понятие ЭТАЖИ.

## Связывание объектов между собой через их свойства

К этому моменту мы выделили отдельные аспекты свидетельства в виде множеств объектов и их свойств. Теперь установим отношения между объектами и их свойствами, вновь обращаясь к тексту информанта.

Краткая запись или формула одного из таких отношений имеет вид:

ДОМ//ТЕКСТ ИНФОРМАНТА (1)  
РАСПОЛОЖЕН НА//ДОМ   
СОСТОИТ ИЗ//ДОМ.

Что означает приведенная запись? Только то, что согласно текста информанта мы можем утверждать, что его дом обладает двумя свойствами: одно описывает его расположение, другое – его состав. Обратим внимание, что понятия, использованные в границах одного отношения, имеют разные контексты своего существования.

В свою очередь свойство РАСПОЛОЖЕН НА предполагает указание какого-то определенного места. Согласно текста и ранее введенных понятий мы можем написать:

РАСПОЛОЖЕН НА//ДОМ (2)  
ГОРА//ОБЩЕЕ ЗНАНИЕ.

Таким образом фразу «дом расположен на горе» мы представили в виде двух отношений: 1) объектов и их свойств, 2) свойств, описываемых через объекты. И объекты и свойства зафиксированы нами ранее.

Аналогично второе свойство дома мы свяжем с уже введенными объектами. Подчеркнем, что проводимое связывание осуществляется строго по тексту.

СОСТОИТ ИЗ//ДОМ (3)  
СТОЛОВАЯ//ДОМ  
СПАЛЬНЯ//ДОМ  
ПРИХОЖАЯ//ДОМ  
ЭТАЖИ//ДОМ

Весь процесс создания отношений между объектами и их свойствами (и наоборот), встреченными в свидетельстве, можно представить как череду ответов на простые вопросы. Что такое ДОМ//ТЕКСТ ИНФОРМАНТА? Это два его свойства РАСПОЛОЖЕН НА и СОСТОИТ ИЗ. В каких контекстах определены эти свойства? РАСПОЛОЖЕН НА//ДОМ и СОСТОИТ ИЗ//ДОМ.

На чем расположен дом? ГОРА. Это понятие никак не разъяснено в тексте информанта. Значит, мы можем предположить существование некого общего знания, в контексте которого осуществимо взаимопонимание с информантом. Таким образом можем написать ГОРА//ОБЩЕЕ ЗНАНИЕ.

Как информант отвечает на вопрос из чего состоит дом? Смотри формулу (3).

Теперь свяжем категорию ЭТАЖИ с зафиксированными нами объектами ЭТАЖ1 и ЭТАЖ2 в соответствующих контекстах. Таким образом возникает отношение:

ЭТАЖИ//ДОМ (4)  
ЭТАЖ1//ДОМ  
ЭТАЖ2//ДОМ

Обратим внимание, что ЭТАЖ2 использовано в приведенных аналитических построениях неоднократно: как объект, входящий в состав ДОМ и как объект, конкретизирующий свойство ИМЕЕТ ОКНА. Такое использование отражает структуру самого свидетельства.

Можем ли мы разъяснить ДОМ//ТЕКСТ ИНФОРМАНТА как-то иначе? Да, можем. Например, написать:

ДОМ//ТЕКСТ ИНФОРМАНТА (5)  
СТОЛОВАЯ//ДОМ  
СПАЛЬНЯ//ДОМ  
ПРИХОЖАЯ//ДОМ  
ЭТАЖИ//ДОМ.

Но в этом случае не все объекты и их свойства, введенные ранее по тексту информанта, найдут выражение в едином знании о предмете свидетельства. В частности, окажется не учтенным свойство СОСТОИТ ИЗ//ДОМ (см. далее генерация единого знания). Имея в виду именно полноту концептуального моделирования, мы разъяснили ДОМ через два его свойства.

Будем говорить, что объекты, их свойства и отношения «объект-свойства» и «свойство-объекты» являются отдельными аспектами знания, выраженного информантом с помощью текста. Создав эти аспекты согласно текста информанта, мы фактически выполнили то, что называется в терминах АКД первичным кодированием.

В ГКООМ все аспекты знания, выражаемого текстом информанта и зафиксированные социологом в аналитическом виде, собраны в единую коллекцию, называемое тезаурусом. Как устроен тезаурус для данного примера можно увидеть, перейдя по ссылке [4].

Практика концептуализации, знакомая каждому социологу по полевой работе, показывает, что кодирование, с одной стороны, весьма трудоемко. С другой, - не может быть выполнено сразу, а требует постоянных смысловых сверок с текстом информанта и модификаций аналитических обозначений на этой основе. Именно для организации итерационного процесса выявления и согласования смыслов, существующих в свидетельстве информанта, ГКООМ предоставляют аналитику средства «смысловой сборки» выявленных аспектов знания в единое целое. Визуальная картина формируемого знания оказывается инструментом его проверки.

## Генерация знания на основе его аспектов, зафиксированных в свидетельстве информанта

Дальнейшие концептуальное переформулирование текстового свидетельства средствами ГКООМ осуществляются с помощью алгоритмической генерации общей структуры знания на основе его аспектов, обнаруженных у информанта. Эта генерация становится возможной благодаря соответствующим алгоритмам. Часть этих алгоритмов описана в [5, 25]. Процесс разработки ГКООМ продолжается, поэтому для ознакомления с актуальной версией ГКООМ следует обратиться к автору.

Для социолога генерация знания, создаваемого из приведенного свидетельства информанта, приводит к графу, показанному на рис. 1. Этот граф получен в результате многих итераций формирования тезауруса. Граф построен с помощью тестовой программной реализации ГКООМ [1] и назван терминологическим.

### Что такое терминологический граф?

Применительно к задаче кодирования в анализе качественных данных терминологический граф - это визуализированное выражение семантики свидетельства информанта, выполненное социологом под определенную задачу. В нашем примере такой задачей являлось возможно полное концептуальное воспроизведение дома, описанного информантом.

Терминологический граф может рассматриваться как результат агрегирования свидетельства информанта. Причем таким результатом являются не отдельные понятия, часто называемые категориями и рассматриваемые аналитиком как обобщающие на основании эвристических критериев [14, 26], а семантическая структура высказываний информанта в целом.

Чрезвычайно важно, что терминологический граф не создается социологом вручную подобно структуре папок на компьютере, а генерируется алгоритмами по текущей массе концептуальных обозначений, выполненных социологом. Тем самым любое концептуальное действие аналитика, локально выражающее тот или иной аспект свидетельства информанта, находит ясное визуальное выражение в генерируемой семантической структуре.

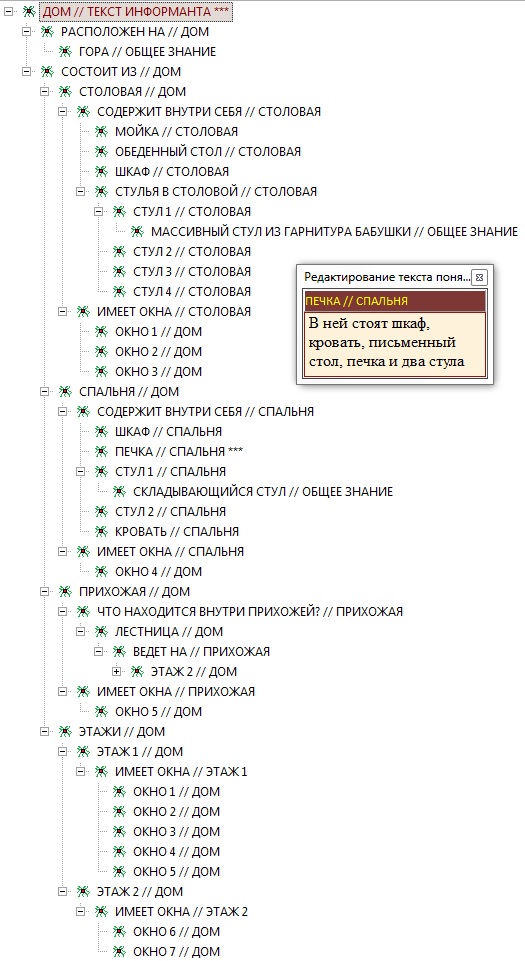


Рис. 1. Терминологический граф   
(\*\*\*: пары понятий, ассоциированные с фрагментом текста информанта).

По результатам концептуализации в свидетельстве информанта аналитик может выделить не одну семантическую структуру. Например, опираясь на текст нашего примера можно увидеть описание комнат в доме, а не всего дома. Какая именно структура окажется представленной в виде терминологического графа, определяется аналитиком.

### Применение терминологического графа при оценке концептуального кодирования свидетельства информанта

Нашей целью являлось наиболее полное выражение семантики текста. Выделяя отдельные смысловые аспекты, мы старались собрать все сведения, сообщаемые информантом. Будучи погруженными в отдельные пункты свидетельства, мы не могли видеть всю семантику, передаваемую информантом с помощью текста.

Выражая аспекты, мы следовали интуитивно предполагаемой связности свидетельства. Поэтому мы рассчитываем, что эта связность окажется присущей концептуальной модели, получаемой на основе материалов информанта.

Насколько связной оказывается концептуальная модель, показывает терминологический граф. А насколько полно терминологический граф выражает семантику, показывает его неформальное сравнение с исходным текстом. По нашему мнению, которое читатель может проверить самостоятельно, интерпретируемость обозначений и связей графа на рис. 1 не уступает возможностям интерпретации текста, лежащего в их основе.

Вместе с тем семантические связи текста значительно нагляднее представлены терминологическим графом, чем самим текстом. Кроме того, на графе становится возможным осуществлять различные операции, имеющие целью анализ семантики. Например, визуализацию структуры свойств или взаимосвязей, относящихся к отдельным объектам.

Аналитическое кодирование не может быть «точнее», чем свидетельство информанта. Скажем, начальное утверждение информанта «Мой дом» на основе здравого смысла может восприниматься двояко – как его недвижимость и как дом, где он живет. В обоих случаях могут быть многочисленные нюансы. В свидетельстве далее никак не разрешается эта коллизия. Поэтому мы проигнорировали это сведение информанта.

Побочный, но принципиально важный эффект ассистирования с помощью ГКООМ состоит в том, что, находясь под постоянным аналитическим контролем, аналитик побуждается максимально адекватно воспроизводить смыслы информанта.

### Уникальность обозначений, выполняемых человеком

Важно, что генерация графа действительно собирает отдельные аспекты, выраженные двумя участниками концептуализации. В этом можно убедиться, если обратить внимание, что ЭТАЖ2, ОКНО1 и ряд других понятий фигурируют в разных частях терминологического графа. Это происходит из-за того, что указанные нами отношения используют одни и те же аспекты неоднократно. Как именно выглядит такое использование, показывает терминологический граф.

При этом человек создает многократно используемые аспекты только в одном экземпляре. Тем самым он принципиально избавляется от необходимости постоянно отслеживать актуальность одной и той же информации, представленной в разных частях знания.

## Дискуссия

Отметим, что нашей целью было начальное ознакомление отечественного социологического сообщества с возможностями аппарата управления знанием, развиваемым в рамках качественной социологии и расширяющим социологический инструментарий. Мы не ставили своей целью обучение разнообразным техникам, существующим в методах управления знаниями, в частности, в ГКООМ. Важно также иметь в виду, что пример в силу своего предназначения не является полноценным применением ГКООМ.

### Кодирование в стиле программирования

Во введении мы обещали предложить инструмент кодирования, дающий в распоряжение неподготовленного пользователя структурный аппарата современного программирования. Отметим только три характерных отношения, на которых строилась выполненная концептуализация. Во-первых, «двойное кодирование» - это модульная организация множества понятий. Модульность является отличительной чертой всех современных средств программирования.

Во-вторых, по ходу изложения мы отмечали, что одни и те же понятия в разных контекстах имеют разные последующие разъяснения, т.е. смыслы. Например, одно и то же понятие СТУЛ1 в разных контекстах СТОЛОВАЯ и СПАЛЬНЯ (см. рис. 1) по-разному представлено в терминологическом графе. В первом случае СТУЛ1 – это МАССИВНЫЙ СТУЛ ИЗ ГАРНИТУРА БАБУШКИ, а во втором - СКЛАДЫВАЮЩИЙСЯ СТУЛ. Такое поведение понятия СТУЛ1 называется в программировании полиморфизмом.

В-третьих, генерация терминологического графа на основе конкретных аспектов «текстового знания» информанта является аналогом компиляции единой программы из отдельных модулей.

### Сравнение с традиционным кодированием

**Связь со свидетельством информанта.** В АКД обычно выделяется фаза открытого, первичного или описательного (дескриптивного) кодирования. Как пишут G. Gibbs и C. Taylor, «это когда кодирование используется для описания того, что содержится в данных.» [24]

При таком кодировании устанавливается явная связь первичного кода с фрагментом(ами) текста информанта. Такая связь как раз является конструктивным выражением идеи обоснования теоретических построений социолога через факты, в том или другом виде сообщаемые ему информантом.

В ГКООМ понятие вводится аналитиком также на основе текста информанта. Однако обоснованием введения понятий является не их явная ассоциация с фрагментом текста, а связность всего множества аналитических обозначений, введенных социологом.

Ассоциирование с текстом, конечно, возможно и в ГКООМ. Например, ПЕЧКА//СПАЛЬНЯ ассоциирован с анализируемым текстом (см. рис. 1). Но такая связь является факультативной, она не рассматривается как фактологическое обоснование концептуальных действий аналитика.

**Обобщение аналитических обозначений.** В АКД существует вторая фаза кодирования, при которой появляются «коды, основанные на аналитических размышлениях исследователя, почему то, что описано в данных, привело к тому, что произошло» [ibid.].

Другими словами, существует фаза вторичного или теоретического кодирования, т.е. построение новых кодов сначала на основании первичных аналитических обозначений, а затем с помощью различных преобразований всего множества кодов, возникающего в результате аналитического переформулирования свидетельства информанта. Такие преобразования могут существовать под именами категоризации [14], осевого кодирования [10] и т.п.

С помощью компьютерных программ эти преобразования выполняются социологом в человеко-машинном режиме с использованием известных компьютерных инструментов – списков, деревьев, сетей и др. Однако в пакетах АКД не существует какого-либо собственного специфического аппарата, который позволял бы обобщить всю массу кодов, созданных социологом на основе сведений информанта.

Иная ситуация при концептуальном кодировании с помощью ГКООМ. Конечно, адепт ГКООМ также должен многократно переформулировать свидетельство в виде значительной массы понятий и отношений между ними. Но ГКООМ обладают собственными алгоритмическими средствами оперирования возникающей «концептуальной массой», построенными на идеях структурного выявления контекстных ограничений тех или иных фактов, сообщаемых информантом.

Все операции с понятиями ГКООМ осуществляются аналитиком в виде конструирования общей структуры знания, постепенно распознаваемого им в свидетельстве информанта на основе здравого смысла. В любой момент построения аналитик может видеть, насколько он преуспел в создании такой структуры, за счет ее генерации алгоритмами ГКООМ. Генерация является оригинальным средством обобщения отдельных сведений информанта в единое знание. В то же время сгенерированная структура знания показывает, насколько связными оказались концептуальные действия самого аналитика.

## Заключение

В статье впервые представлено применение оригинального аппарата управления знаниями - ГКООМ - в задаче АКД. ГКООМ принципиально расширяют возможности инструментальных средств АКД за счет структурного учета контекстов, в которых существуют отдельные высказывания информанта.

Дополняя существующие пакеты АКД, ГКООМ дает в распоряжение социолога современный структурный аппарат программирования в виде отношений полиморфизма, модульности, компиляции.

Концептуальная модель, получаемая в результате применения ГКООМ, не строится автоматически по анализируемому тексту. Она появляется как результат многочисленных итерационных усилий аналитика, с помощью которых он уточняет смыслы, распознаваемые им в свидетельстве информанта.

Применение ГКООМ для целей АКД позволяет осуществлять аналитическую проверку концептуальных построений социолога в области качественной социологии.

Направления дальнейшего развития методов АКД средствами ГКООМ могут быть многообразны: внедрение в пакет АКД, разработка методик применения, создание оригинального онторедактора, ориентированного на создание социологического знания и т.п.

## Библиографический список

1. Графовые контекстно-ориентированные онтологические методы: программная реализация // Сайт проекта ГКООМ [электронный ресурс]. Дата обращения 21.01.2017. URL: <<http://coknowledge.ru/wp-content/uploads/2017/02/GKOOM_Program.pdf>>.
2. ГКООМ: объекты, выделенные в тексте информанта // Сайт проекта ГКООМ [электронный ресурс]. Дата обращения 21.01.2017. URL: <<http://coknowledge.ru/wp-content/uploads/2017/02/Objects.pdf>>.
3. ГКООМ: свойства объектов, выделенные в тексте информанта // Сайт проекта ГКООМ [электронный ресурс]. Дата обращения 21.01.2017. URL: <<http://coknowledge.ru/wp-content/uploads/2017/02/Properties.pdf> >.
4. ГКООМ: тезаурус // Сайт проекта ГКООМ [электронный ресурс]. Дата обращения 21.01.2017. URL: < <http://coknowledge.ru/wp-content/uploads/2017/02/Thezaurus.pdf>>.
5. Каныгин Г.В., Полтинникова М.С. Контекстно-ориентированные онтологические методы в социологии // Труды СПИИРАН. 2016. №5(48). С. 107-124.
6. Каныгин Г.В. Текст как инструмент социального теоретизирования // Пути России: Проблемы социального познания / Под общ. ред. Д.М. Рогозина. – М.: МВШСЭН, 2006. С. 144-160.
7. Объектно-ориентированное программирование: <https://en.wikipedia.org/wiki/Object-oriented_programming>
8. Аспектно-ориентированное программирование: <https://en.wikipedia.org/wiki/Aspect-oriented_programming>
9. Рубашкин В.Ш. Онтологическая семантика. Знания. Онтологии. Онтологически ориентированные методы информационного анализа текстов. М: Физматлит, 2012.
10. Страусc А., Корбин Дж. Основы качественного исследования: обоснованная теория, процедуры и техники / Пер. с англ. и послесловие Т. С. Васильевой. — М: Эдиториал УРСС, 2001.
11. Чесноков, С.В. «Мне интересен человек как человек…» // Социологический журнал, №2, 2001. С. 63-122.
12. Barbour R. Quality of Data Analysis, in: Flick U. (ed.) The Sage Handbook of Qualitative Data Analysis. Sage, 2014, pp. 496—509.
13. Bazeley P. A. and O'Rourke B. Getting started in research: A series of reflections for the beginning researcher. University of Western Sydney: Macarthur, 1996.
14. Charmaz, K. Grounded Theory: Objectivist and Constructive Methods. In Norman K. Denzin & Yvonna S. Lincoln (Eds.), Handbook of Qualitative Research, 2nd edition. 2000. Thousand Oaks, Ca.: Sage (pp.509-535).
15. Denzin N., Lincoln Y. (eds.) Handbook of Qualitative Research 2nd ed. Thousand Oaks, CA: Sage, 2000.
16. Franzosi R., Gianluca De Fazio and Vicari S. Ways of Measuring Agency: An Application of Quantitative Narrative Analysis to Lynchings in Georgia (1875–1930)// Sociological Methodology, 2012, 42:1–42.
17. Gibbs G. R. Different Approaches to Coding // Sociological Methodology, 2012. 42(1):82–4.
18. Junker A. Optimism and Caution Regarding New Tools for Analyzing Qualitative Data// Sociological Methodology, 2012, 42(1):85-7.
19. Kelle U. Theory Building in Qualitative Research and Computer Programs for the Management of Textual Data, Sociological Research Online, 1997, 2(2)  Accessed 25.11.2015. URL:<<http://socresonline.org.uk/2/2/1.html> >.
20. LaPelle N. Simplifying Qualitative Data Analysis Using General Purpose Software Tools, Field Methods, 2004, 16(1), pp. 85—108.
21. Lewins A., Silver C. Using Qualitative Software: A Step-by-Step Guide. London: Sage, 2007.
22. MacMillan K. and McLachlan S. Theory-Building with Nud\*Ist: Using Computer Assisted Qualitative Analysis in a Media Case Study // Sociological Research Online, vol. 4, no. 2. Accessed 01.06.2006. URL<<http://www.socresonline.org.uk/4/2/macmillan_mclachlan.html>>.
23. Online QDA. Accessed 12.01.2017. URL:< <http://onlineqda.hud.ac.uk/index.php>>.
24. Online QDA. Accessed 16.01.2017. URL:< <http://onlineqda.hud.ac.uk/Intro_QDA/how_what_to_code.php>>
25. Poltinnikova M. and Kanygin G. Graphic Ontological Methods to Compile Collective Knowledge of Social Processes //Proceedings EGOSE 2016 'Electronic Governance and Open Society: Challenges in Eurasia' 2016. P. 223-228.
26. Thompson R. Reporting the Results of Computer-assisted Analysis of Qualitative Research Data [42 paragraphs]. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-line Journal], 2002, May, 3(2). Accessed 15.01.2017. URL:<<http://www.qualitative-research.net/fqs-texte/2-02/2-02thompson-e.htm>.>
27. White, M., M. D. Judd, and S. Poliandri. Illumination with a Dim Bulb? What Do Social Scientists Learn by Employing Qualitative Data Analysis Software in the Service of Multi-Method Designs? // Sociological Methodology, 2012, 42(1):43-76.

1. Качественные данные могут существовать во многих форматах – текста, видео, аудио и др. Но функции кодирования – это инструментальный аппарат, который одинаков для всех форматов. Поэтому, анализируя этот инструмент, мы без потери общности говорим только об одном формате свидетельства информанта – обычном (плоском) тексте. [↑](#footnote-ref-1)