

КОНТЕКСТНО-ОРИЕНТИРОВАННАЯ КОНЦЕПТУАЛИЗАЦИЯ ПОНЯТИЙ

В статье описываются основные методологические принципы контекстно-ориентированного анализа качественных данных. Проводится сравнение двух процедур концептуализации текстовых свидетельств, получаемых в ходе прикладного социологического исследования, с одной стороны, осуществляемых с помощью пакета *Atlas.ti*, с другой — программного пакета, работающего на контекстно-ориентированных принципах. Рассмотрены основные инструментальные средства контекстно-ориентированного анализа данных, приведены примеры использования контекстно-ориентированных методов при анализе транскрипта интервью. Исследованы возможности пилотной версии программы, ассистирующей проведению контекстно-ориентированного анализа. Проанализирована роль терминологического графа как инструмента для контроля связей между понятиями.

Ключевые слова: социологическое исследование, качественные данные, методы качественного исследования, качественный анализ, контекстно-ориентированный подход.

Введение

На сегодняшний день становится очевидной возрастающая популярность качественных социологических исследований, расширяется область их применения, они становятся все более востребованными; активно развиваются прикладные программы для анализа качественных данных. В связи с этим возникают вопросы как методологического, так и практического характера, обусловленные отсутствием опыта применения таких программ исследователями. Проблемы методологического характера выражаются в том, что социологи и специалисты других гуманитарных дисциплин, долгое время оперировавшие количественными методами анализа данных, рассматривают нечисловой формат представления данных при анализе как недостаток: такие данные трудоемки для анализа, и специализированное программное обеспечение для анализа таких данных появилось позже, чем программы для анализа количественных данных. Сегодня существует большое количество программных продуктов для поддержки анализа качественных данных, например *Atlas.ti*, *MAXQDA*, *NVivo* и др. Программы этого класса широко используются такими компаниями, как высокотехнологичная *Nokia* и известная западная консалтинговая компания

Flack Consulting Ltd. Один из известных пакетов анализа качественных данных (АКД) *Atlas.ti* используется в Британском Парламенте для поддержки принятия управленческих решений.

Существует проблема использования методов анализа качественных данных в социологических исследованиях. Программное обеспечение, распространяемое сегодня, не обладает инструментами для решения таких задач, как высокотехнологичное обеспечение коллективной работы исследователей, построение объективной модели изучаемого объекта, возможность проверки построенной модели. Проблема использования анализа качественных данных в целях обеспечения поддержки принятия решений почти не разработана. Безусловно, можно обратиться к западным разработкам пакетов АКД, широко применяемым зарубежными исследователями и малоизвестными российскими учеными. С другой стороны, существуют отечественные проекты, посвященные анализу качественных данных, в том числе разработке программных продуктов для АКД. На сегодняшний день с сожалением приходится констатировать, что эти работы слабо освещены в печати и известны только узкому кругу специалистов. Одним из примеров может служить проект, посвященный контекстно-ориентированным методам анализа качественных данных.

В статье рассматриваются основные методологические предпосылки контекстно-ориентированной методологии, также представлен практический пример, иллюстрирующий конструктивные отличия контекстно-ориентированных методов от традиционных методов анализа качественных данных.

Общие методологические требования к построению контекстно-ориентированных структур

Анализируя документ, исследователь сталкивается с проблемой концептуализации исходных данных, т. е. приведения их к виду, позволяющему повысить эффективность их дальнейшего изучения. В анализе качественных данных концептуализация выполняется с помощью функций кодирования. Применяя эти функции, исследователь руководствуется принятыми им субъективными методологическими ориентациями, что затрудняет дальнейшую корректировку и развитие научных взглядов, выражаемых в процессе концептуализации. Вместе с тем в современном АКД отсутствуют формальные критерии, по которым было бы возможно сопоставление выводов исследователей, работающих с качественными данными.

Другая, не менее важная проблема, с которой сталкивается исследователь при анализе качественных данных, это «ловушка естественного языка» или неоднозначность понятий, используемых при кодировании и построении теорий. Подобная проблема свойственна не только социологии, но и другим гуманитарным наукам, где в процессе работы над данными аналитик использует естественный язык. Конечно, в большинстве случаев существует установленный терминологический аппарат, принятый научным или профессиональным сообществом в определенной предметной области, но это не решает проблему по нескольким причинам: во-первых, термины и определения описаны опять же с помощью естественного языка, во-вторых, отсутствует четкий механизм формирования новых терминов. В связи с этим возникает необходимость в инструменте, позволяющем строить понятийные структуры на основе терминов естественного языка при анализе качественных данных. Такой инструмент предлагает контекстно-ориентированный подход к анализу качественных данных (КО АКД), помогающий исследователю при решении проблем, связанных с неоднозначностью понятий естественного языка при работе с документами.

Рассмотрим подробнее ключевые термины данного подхода — абстрактное понятие и контекстно-фиксированное разъяснение — которые служат базовыми инструментами, обуславливающими методологические преимущества КО АКД.

Абстрактное понятие — «это простейший инструмент теоретизирования, который выполняет три главные функции при решении проблемы многосмысленности и контекстной обусловленности теоретических терминов, которое осуществляется исследователем с помощью действий с понятиями». Иными словами, абстрактное понятие выступает своего рода гарантом идентичности в разных контекстах использования, очерчивает границы набора свойств понятия, а также осуществляет связь между самим понятием и действующим лицом (актером), которое оперирует этим понятием. Абстрактное понятие, таким образом, означает что-то, что остается неизменным независимо от контекста рассмотрения.

Контекстно-ориентированный подход основан на утверждении, что, работая с текстом, исследователь фактически объясняет одни понятия через другие, обязательно связывая разъяснения с определенным контекстом. Таким образом, появляется структура, состоящая из понятия, раскрытого через другие понятия, и контекста, в котором данные понятия существуют. Эта структура получила название

«контекстно-фиксированное разъяснение». Разъяснение понятий продолжается до тех пор, пока понятия не будут разъяснены в так называемые «общепринятые понятия» — понятия, которые исследователь считает общими для всей системы разъяснений, в частности, понятными для каждого ее контекста.

Итогом работы исследователя, использующего данный подход АКД, будет структурно упорядоченная система разъяснений, включающая в себя все понятия и контекстно-фиксированные разъяснения, которые, на взгляд исследователя, отражают точку зрения информанта на исследуемую проблему.

Проведем ряд сравнений между традиционным АКД и его контекстно-ориентированным расширением. Отправной структурной единицей в обеих версиях служит дескриптор, известный в традиционном подходе к АКД как *код*. В контекстно-ориентированном подходе дескриптор называется «понятие». Итак, понятие — простейший структурный элемент, используемый в контекстно-ориентированном анализе, своего рода «атом», из множества которых исследователь создает «молекулы», описывающие интересующую его предметную область.

Отвечая на вопрос, каким образом из атомов-понятий складывается целостная система, описывающая предметную область, мы переходим ко второму базовому элементу контекстно-ориентированного подхода: к так называемому *контекстно-фиксированному разъяснению*.

Контекстно-фиксированное разъяснение (КФР) можно определить как простейшее отношение, связывающее между собой понятия. КФР является структурным элементом, который конструируется из простейших понятий с помощью отношения «родитель—дети», пришедшего из теории графов. Схематично описанное отношение можно представить следующим образом:

$$t^{(c)}:t^{(p)}\{t_i\},$$

где $t^{(p)}$ — родительский или разъясняемый термин, $\{t_i\}$ — разъяснение или дочерний термин. Постепенно мы подошли к третьему ключевому понятию в контекстно-ориентированной методологии — к самому *контексту*. Контекстом в нашем случае будет являться $t^{(c)}$; иными словами, контекст — некое условие, своего рода контейнер, в который исследователь помещает разъяснение; контекст играет уточняющую роль в процессе разъяснения, это своего рода идентификатор, добавляющий к имеющемуся разъяснению еще одно измерение: у нас может быть

сколько угодно одинаковых отношений «родитель—дети», но в разных контекстах они будут иметь разное значение, определяемое контекстом.

Теперь рассмотрим, какие конструктивные преимущества для построения теорий и описания предметной области имеют описанные выше структуры, в частности, контекстно-фиксированное разъяснение (КФР).

Прежде всего, необходимо отметить, что КФР является достаточно гибкой структурой, не претендующей на задание единственно правильного отношения между термином-родителем и термином-потомком. Это свойство оправдано не только с точки зрения контекстно-ориентированной методологии, но и с точки зрения классических функций кодирования: ведь известно, что один и тот же поток данных можно закодировать по-разному в зависимости от множества факторов, начиная с целей и задач анализа и заканчивая субъективным взглядом и профессиональными компетенциями аналитика.

Следующий важный момент заключается в том, что КФР позволяет проиллюстрировать сам факт разъяснения с помощью одного термина (контекста), что является интуитивно понятным в качестве обозначения *обстоятельств*, в которых данное разъяснение имело место и *смысл*.

Коллекция понятий, заведенных аналитиком в процессе работы и объединенных им в контекстно-фиксированные разъяснения, называется *тезаурус*. Следует отметить, что не каждое понятие, встречающееся в тезаурусе, должно быть включено в какое-либо КФР; понятия могут существовать в тезаурусе отдельно. Такой термин игнорируется контекстно-ориентированным алгоритмом при построении результирующего *терминологического графа* (об этом будет рассказано ниже). Второй момент заключается в том, что связи между понятиями существуют только внутри конкретного КФР, тезаурус не содержит объектов, напоминающих графы или деревья.

Итак, тезаурус содержит все разъяснения понятий, составляющие терминологию исследователя, которой он пользовался в процессе кодирования. И понятия, и их разъяснения верифицируются (определяются, разъясняются) в тезаурусе аналитиком, производящим кодирование. Исходя из идеи разъяснения, тезаурус может быть определен как концептуальный инструмент, связывающий между собой термины, введенные аналитиком, и общие понятия (*commonplace*).

Завершая описание основных объектов контекстно-ориентированных процедур, можно сформулировать общие правила, регулирующие их применение в процессе концептуализации:

— понятие, представленное в тезаурусе, является *контекстно-определенным*, если оно является разъясненным хотя бы в одном КФР;

— одно и то же понятие может быть определено в нескольких контекстах: в таком случае понятие будет *поликонтекстуальным*. Соответственно, понятие, разъясненное только в одном КФР, будет *моноконтекстуальным*;

— контекстное определение понятия, представленного в КФР, не зависит от набора контекстов, используемых в данном КФР как разъяснения; набор разъяснений может быть и пустым.

Концептуализация как набор разъясняемых понятий

Проиллюстрируем на примере прикладной концептуализации свидетельства информанта отличия контекстно-ориентированных методов от традиционных функций кодирования. Для наглядности приведем два примера кодирования одного и того же из отрывка интервью молодого человека о жизни в пригороде Санкт-Петербурга:

— *Какие недостатки в работе транспорта, по вашему мнению, существуют в нашем городе?*

— *Ну, как для студента в первую очередь для меня важна стоимость проезда. Все-таки 40 рублей за маршрутку до метро — это ощутимое бремя... Да ладно 40 рублей, они бы хоть ходили нормально — а то стоишь на холоде полчаса, очередь с километр длиной, и опаздываешь вечно...*

— *Удовлетворены ли Вы работой медицинских учреждений нашего города?*

— *Мне, к счастью, редко приходится там бывать — вот был за справкой в военкомат. Больше ни за что туда не сунуть — очереди страшные, номерков никогда нет, и еще врач принимает раз в неделю... В общем, хорошего мало.*

— *Удовлетворены ли Вы возможностями для проведения свободного времени в нашем городе?*

— *Нет, конечно! На весь городишко один тренажерный зал — и то со старыми тренажерами, не то что тебе инструктора, даже вахтера не дозовешься, чтобы раздевалку открыл. Было хоть поле футбольное около 2-й школы — и то снесли, строят там спортивный комплекс. Открытых площадок тоже нет — ни кольца баскетбольного, ни волейбольной сетки. С друзьями негде собраться, в кафе полно курящих — а мы не курим, между прочим... А в «Фортуну» не особо хочется идти — музыка не нравится, да и там одни взрослые.*

— Как вы рассматриваете возможность дальнейшего трудоустройства в нашем городе?

— Да никак не рассматриваю. Здесь по моей специальности — я юрист — мне делать нечего. К тому же, все мои друзья в городе, хочется перебраться к ним поближе. Да и платят здесь, куда ни посмотри, копейки. Одним словом — бесперспективняк. Нет, здесь я работать не собираюсь.

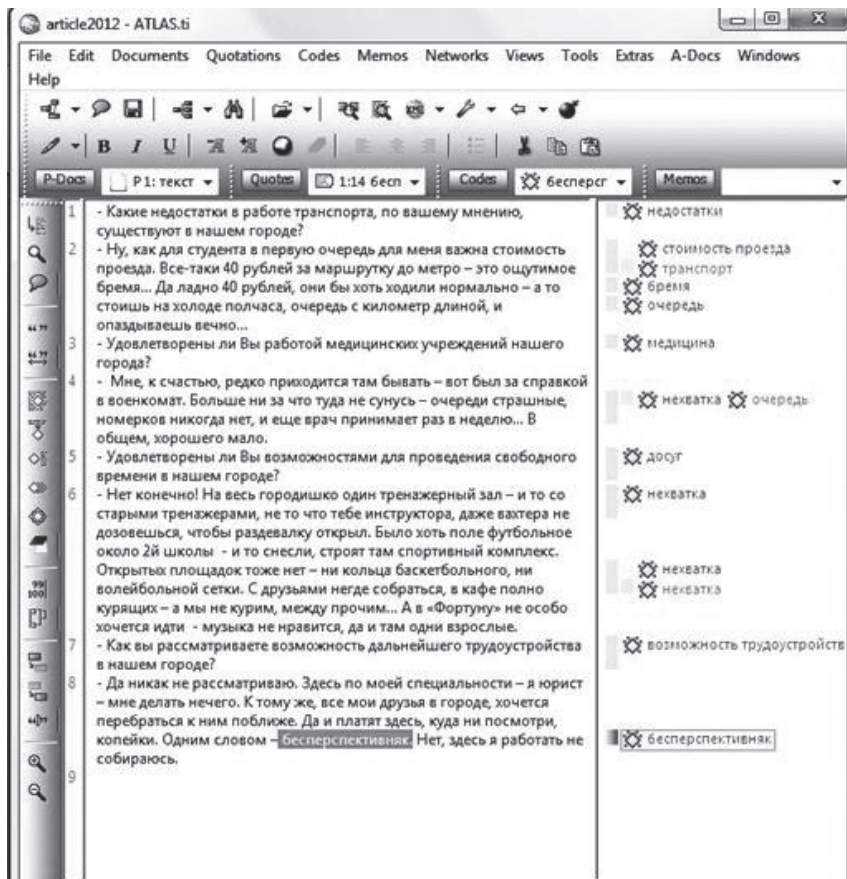
Данный текст взят из пилотажного исследования, посвященного проблемам молодежи Ленинградской области и проведенного автором статьи в рамках производственной практики. В результате анализа предполагается получить упорядоченную информацию о проблемах молодежи Ленинградской области с точки зрения интервьюируемого.

Для анализа свидетельств информанта используется пакет *Atlas.ti*. В процессе анализа мы проведем концептуализацию свидетельств информанта, то есть перейдем от естественного языка к новым структурам (кодам, понятиям). Такой переход необходим по ряду причин: во-первых, естественный язык неоднозначен, его трудно формализовать; во-вторых, переход к концептуальным структурам позволит отделить факты, перечисленные информантом, от его оценок этих фактов, приведенных в этом же интервью.

Первым шагом анализа является кодирование. В *Atlas.ti* можно использовать как открытое кодирование, так и кодирование *in vivo*. Полученные коды отображаются в правой вспомогательной области (см. рис. 1).

На следующем шаге анализа устанавливаем связи между кодами. В пакете существует фиксированный набор связей, которые можно использовать для связывания кодов: часть—целое, причинно-следственная связь, ассоциативная и др. Однако как бы ни был универсален с точки зрения разработчиков набор имеющихся связей, он не отвечает всем требованиям исследователя и не позволяет создавать новые виды связей, что довольно неудобно. Например, в процессе анализа исследователю понадобится новый вид связи, или необходимо будет скомбинировать одну связь из нескольких (что часто бывает при работе с данными на естественном языке).

Представим себе, что исследователю необходимо отразить, что какое-либо событие является следствием нескольких последовательно произошедших до этого событий. Стандартными средствами существующих пакетов АКД такие требования реализовать трудно: придется каждый раз при использовании такой связи ссылаться

Рис. 1. Кодирование в *Atlas.ti*

на цепь произошедших до этого событий. Это приведет к увеличению количества кодов и связей между ними.

Процесс установления связей между кодами показан на рис. 2.

Результат анализа может быть представлен в виде древовидной структуры (см. рис. 3), отражающей связи между кодами.

Исходя из данных, представленных в виде древовидной структуры, аналитик может сделать выводы о структуре изучаемых данных и детально изучить полученные зависимости и связи между кодами.

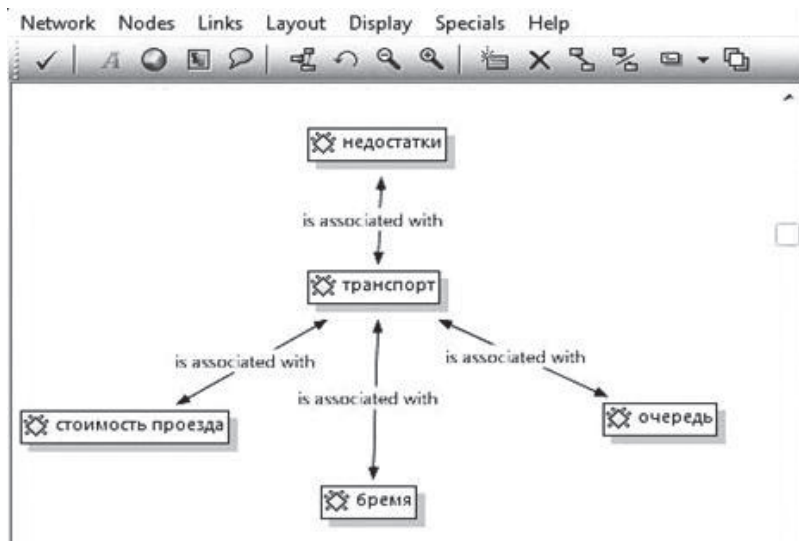


Рис. 2. Установление связей между кодами

Например, в нашем случае налицо следующая закономерность: говоря о каждом из аспектов проживания в пригороде Санкт-Петербурга, респондент описывает недостатки, причем такие недостатки, как очередь и нехватка ресурсов, упоминаются им неоднократно применительно к разным областям (транспорту, медицине и др.). Скажем, очередь как недостаток присутствует в речи респондента во всех затронутых им областях жизни в пригороде: это и очередь на транспорт, и очередь к врачу в поликлинике. Нехватка, недостаток ресурсов тоже упоминается респондентом неоднократно в разных ситуациях: это и нехватка врачей в поликлинике, и нехватка спортивных площадок, мест для проведения досуга молодежи, и, в заключение, нехватка рабочих мест в пригороде.

Далее проведем процедуру анализа данного транскрипта интервью с использованием контекстно-ориентированных методов.

На первом шаге анализа исследователь производит ровно те же действия, какие производил при использовании пакета *Atlas.ti*, а именно вводит в систему понятия, которые понадобятся ему в дальнейшем для анализа. Понятия, заведенные аналитиком, хранятся в базе данных и представлены в виде списка, как это показано на рис. 4.

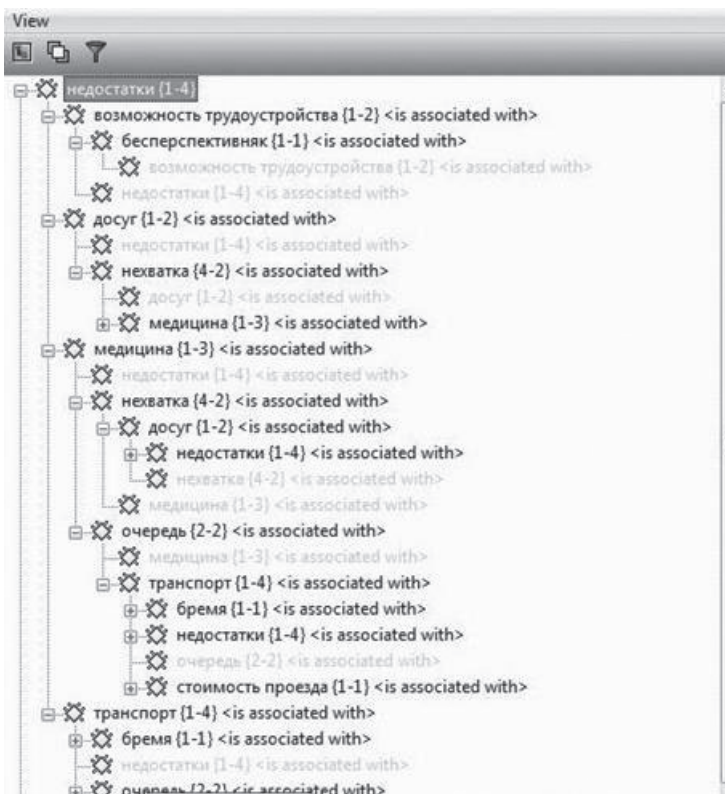


Рис. 3. Древовидная структура кодов

Необходимо отметить некоторые особенности реализованного программного прототипа, наверняка привлечшие внимание читателя на рис. 4. Каждому заведенному в систему понятию (термину, контексту) присваивается уникальный номер (идентификатор), который можно видеть слева напротив каждого термина. По умолчанию понятиям тезауруса присваивается уникальный идентификатор. Сортировка понятий тезауруса настроена по возрастанию значений уникального идентификатора, поэтому каждое вновь созданное понятие оказывается после создания сверху (это очень удобно при работе с большими объемами информации, когда тезаурус может насчитывать сотни понятий). Далее при работе по концептуализации значение идентификатора может

№	Мнемоника
-3066	ПЛОХО
-3065	ОЧЕРЕДЬ
-3064	ОЧЕРЕДЬ К ВРАЧУ
-3063	ДОСУГ
-3062	ПРИГОРОД САНКТ-ПЕТЕРБУРГА
-3061	БЕСПЕРСПЕКТИВНЯК
-3060	БЕЗРАЗЛИЧИЕ К ТРУДОУСТРОЙСТВУ
-3059	ВОЗМОЖНОСТЬ ТРУДОУСТРОЙСТВА
-3058	ОЧЕРЕДЬ НА ТРАНСПОРТ
-3057	БРЕМЯ
-3056	НЕХВАТКА
-3055	НЕДОСТАТКИ
-3054	МЕДИЦИНА
-3053	ТРАНСПОРТ
0	MISSING A DEFINITE TERM
1	INTEGER NUMBER
2	TEXTUAL DATA
50	PREDEFINED TERMS
51	CONTEXT RAMIFICATION
52	THIS IS
100	ACTOR
101	DEFINITIVE CONCEPT
121	ACTORS
122	ACTIONS//UNSPECIFIED ACTOR

Рис. 4. Понятия, заведенные исследователем для контекстно-ориентированного анализа

быть изменено на положительное для удобства пользователя (при соблюдении требований уникальности). Англоязычные термины, указанные в нижней части рабочей области, являются системными и уже существуют в тезаурусе.

После того как исследователь ввел в систему необходимые понятия, можно приступить к процедуре разъяснения. В процессе этой процедуры исследователь создает контекстно-фиксированные разъяснения, т. е. разъясняет одни понятия через другие в различных контекстах.

Разъяснять можно любое важное с точки зрения исследователя понятие. В нашем случае было выбрано понятие «бесперспективняк», воспроизводящее лексику информанта и внесенное в систему в процессе анализа. Это понятие было выбрано по двум причинам: во-первых, оно

привнесено информантом, во-вторых, на взгляд исследователя, оно отражает отношение информанта к обсуждаемому вопросу, в-третьих, это понятие может оказаться довольно емким (т. е. может оказаться, что с помощью этого понятия информант объясняет свое отношение не только к возможностям трудоустройства по специальности, но и к ситуации в пригороде в целом).

Примечательно, что процесс ввода в систему понятий для анализа является итерационным, т. е. по мере углубления в данные аналитик вводит новые понятия, позволяющие полнее раскрыть суть изучаемого. Например, приступая к анализу приведенного выше отрывка из интервью, мы ввели в систему только те термины, которые являлись кодами при обработке транскрипта с помощью *Atlas.ti*. Однако при дальнейшей работе с текстом мы осознали необходимость введения в систему контекстов (например, «пригород Санкт-Петербурга») и оценочных понятий (респондент в тексте дает оценку описываемым событиям, поэтому в систему было введено оценочное понятие «плохо»). Так, шаг за шагом, исходя из целей исследования, аналитик вводит в систему новые термины, устанавливая связи между ними с помощью контекстно-фиксированных разъяснений.

Анализируя представленный отрывок интервью, можно привести следующий вариант контекстно-фиксированного разъяснения: термин «бесперспективняк», разъясненный в контексте «пригород Санкт-Петербурга», определяется через понятия «проблемы с транспортом», «плохая медицина», «отсутствие возможностей проведения досуга», «отсутствие возможностей трудоустройства». Это подтверждает гипотезу исследователя о том, что термин «бесперспективняк» описывает отношение информанта не только к возможностям трудоустройства, но и к ситуации в его городе в целом.

Следует пояснить соотношения между понятиями и контекстами, в которых они определены. Как было сказано выше, под контекстом в системе понимается некое условие, в которое, как в контейнер, помещаются разъясняемые понятия. Контекст, таким образом, способствует однозначному определению понятия, вписанного в него. По умолчанию в системе задан контекст высшего уровня — так называемый «актуальный тезаурус». Этот контекст означает, что все разъясненные в нем понятия определены в рамках конкретной процедуры анализа. Появление этого термина объясняется рядом причин: во-первых, логика работы системы предусматривает необходимость наличия хотя бы одного контекста при построении терминологического графа; во-

вторых, отвлеченные понятия (например, оценочные суждения) могут быть использованы неоднократно при анализе и в разных контекстах, поэтому удобнее разъяснить их в контексте высшего уровня, чтобы корректно использовать в дальнейшем при работе с другими контекстами.

Далее при работе пользователь может создавать свои контексты, в рамках которых будут разъяснены его термины. В нашем примере события, о которых рассказывает информант, происходят в конкретном месте — в пригороде Санкт-Петербурга, в связи с чем целесообразно ввести в систему контекст «пригород Санкт-Петербурга». Вводя такой контекст, аналитик получает возможность в ходе своих формализованных рассуждений явно указывать место совершения событий, привязывать события к определенным условиям, в которых они происходили. Таким образом, понятия «недостатки медицины», «очередь на транспорт» и др. будут однозначно определены в рамках контекста «пригород Санкт-Петербурга».

После того, как исследователь ввел в систему все необходимые для работы понятия и создал КФР, можно приступать к так называемой «финализации термина» или к построению терминологического графа. Смысл процедуры финализации состоит в конструктивном прослеживании связей между всеми терминами, введенными при концептуализации; в частности, в получении ответа на вопрос, каким образом понятия, привнесенные аналитиком для переосмысления текстового свидетельства, соотношены с собственными суждениями информанта. Такое конструктивное прослеживание осуществляется в виде терминологического графа.

Терминологический граф — объект контекстно-ориентированной методологии, представляющий собой древовидную структуру связей между понятиями с обозначением контекстов, в которых они разъяснены. Терминологический граф отображает структуру анализируемых данных и позволяет делать выводы о качестве связей между понятиями. Пример терминологического графа для анализируемого в данной статье транскрипта интервью представлен на рис. 5.

Важно отметить, что процесс анализа данных с использованием контекстно-ориентированных методов является итерационным процессом. Анализируя данные, практически невозможно с первого раза точно определить, какое количество понятий необходимо ввести, чтобы максимально точно описать структуру исследуемого объекта. Создавая КФР, исследователь по мере работы может добавлять вспомогательные контексты и понятия, создавать с их использованием новые КФР.

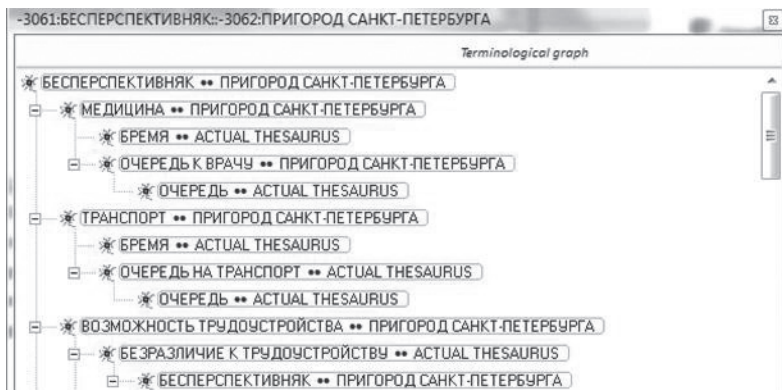


Рис. 5. Терминологический граф

Терминологический граф, таким образом, является инструментом для проверки связей между понятиями; например, если возникает ситуация, когда понятия разъясняются друг через друга, терминологический граф оказывается «зацикленным». Глядя на такой граф, исследователь понимает, что в разъяснениях допущена ошибка, исправляет ее и снова проверяет себя с помощью построения терминологического графа.

Заключение

На сегодняшний день существует проблема использования методов анализа качественных данных в социологических исследованиях. Программное обеспечение, существующее сегодня, не обладает инструментами для решения таких задач, как построение объективной модели изучаемого объекта, возможность проверки построенной модели.

Описанные в статье инструменты контекстно-ориентированного анализа качественных данных (КО АКД) позволяют проводить анализ свидетельств информанта, используя контексты (определяя временную и/или пространственную принадлежность событий к конкретному времени/месту) и специальный инструмент для проверки конструируемой модели — терминологический граф.

Программное обеспечение для проведения контекстно-ориентированного анализа существует сегодня в виде программного прототипа, ведется разработка пилотной версии программы, предназначенной для практического применения.