

КОНТЕКСТЫ ЭМПИРИЧЕСКОГО ПОЗНАНИЯ В СОЦИОЛОГИИ И ВОЗМОЖНОСТИ ПРОГРАММЫ ВЕГА¹

Галина Саганенко, доктор социол. наук, проф., в.н.с. СИ РАН
Евгений Каневский, к. ф.-м. наук, в.н.с. Экономико-математического института РАН
Кирилл Боярский, к. ф.-м. н., доцент СПб-го государственного университета информационных технологий, механики и оптики²

I. О системах производства эмпирического знания в социологии

Воспользуемся возможностью в рамках данной статьи показать наши представления о характере идей, определяющих наши пути поиска эмпирического знания в социологии и место непосредственно компьютерной технологии в общей познавательной системе. Наши идеи выглядят достаточно просто, хотя в общем виде затрагивают глобальные аспекты: мы говорим — учитывайте эпоху на дворе, уровень развития социологии, проблему, за которую беретесь (по зубам ли она вам), каков у вас как исследователя потенциал, каковы ваши средства, что вы хотите от респондента и пр., и пр. Поставив себе задачу понять, от каких идей мы стартуем в наших амбициях, узнать коллизии социальной реальности в наших исследованиях, я пошла за примерами к работам своих коллег и своим работам периода наивного методологического энтузиазма. Получился для начала набросок такой картины.

Отечественная социология пытаясь утвердиться в лоне серьезных наук, потратила много усилий, чтобы доказать научный характер и соответствующее качество социологического знания. Многие мои коллеги, включая и нас самих, пытались подобраться к смыслу получаемого социологического знания преимущественно на пути логико-формального обоснования, социальный контекст получаемого знания даже не упоминался. Спустя многие годы можно посмотреть, какие доказательные конструкции мы использовали и от каких исходных позиций мы начинали двигаться. Например, у Г.С. Батыгина получалось так: «Обоснование является специфическим познавательным процессом, отражающим взаимопереход и взаимосвязь идеальных объектов научного исследования» (Батыгин 1986, с. 11); или еще так: «Будем трактовать обоснование достаточно широко ...как комплекс логико-теоретических процедур, правил и эвристик (неявных допущений), которые регламентируют переход от фактологической информации к научному знанию и практическим решениям» (там же, с. 12).

Фактически только в конце книги Геннадий Семенович упоминает социолога в качестве фактора познавательной деятельности, собираясь «осмыслить роль творческого компонента познавательной деятельности применительно к повседневной работе социолога, проблемам формирования подлинной культуры научного мышления» (там же, с. 202).

В своей методологической монографии (Г.И. Саганенко, 1983; научный редактор В.А. Ядов) мною с первых

страниц заявляется такой пассаж: «Надежный результат складывается из многих элементов: разработки четкой концептуальной схемы, включающей обоснование социальных переменных; применение добротного инструментария, который призван обеспечить надежные исходные данные по выбранным переменным, обоснования выборки и соблюдения ее принципов; правильной методической организации самой процедуры обследования; тщательной первичной обработки данных для последующего их анализа на ЭВМ; применения соответствующих характеру данных математических методов и адекватной интерпретации его результатов; рассмотрения выявленных фактов в одной общей системе понятий и, в конце концов, в рамках теории» (Саганенко 1983, с. 4). Как видим, перечислены основные элементы процесса социологического исследования. В принципе далее 190 страниц монографии можно было бы и не писать, поскольку «исчерпывающие» установки уже были сформулированы.

Еще в первой монографии 1977 года по изучению качества первичных социологических данных, автор этой статьи была в мучительных раздумьях на тему — как это так получается, что у нас в руках уже есть какие-то данные, к изучению качества которых мы вот сейчас приступаем. Совет от Владимира Александровича был получен и раздел о первичных данных в монографии 1983 г. был начат с пассажа: «Мы начинаем анализ проблемы с того момента, когда содержание исходных признаков, по которым будет исследоваться явление, в достаточной степени установлено и тем самым открываются возможности обсуждать вопросы на более строгом языке...» (там же, с. 29). Выходит, что проигнорировали важнейшие моменты, связанные со «стартом» и выбором и начали фактически с середины.

Представим, что говорил ведущий отечественный методолог Владимир Александрович Ядов по поводу начальных интенций взаимодействия социолога с социальной реальностью и исследованием в книге «Социологическое исследование: методология, программа, методы» (1995). Читаем: «Говоря о методологии и методах социологического исследования, мы должны, конечно, уяснить, каков предмет социологии как науки... Его движение зависит от двух решающих факторов: прогресса самого научного знания, с одной стороны, и меняющихся запросов общества, социального запроса, с другой» (с. 5-6). Тут, пожалуй, можно задать уточняющий вопрос — а куда делось само меняющееся общество, в котором масса чего исчезла и масса чего появилась.

Владимир Александрович объявляет: «*Ключевой единицей анализа становится то, что называется «событием», действием социальных агентов»* (с. 13; подчеркивания наши — Г.С.) и еще через несколько страниц: «*Понятие социальной общности представляется нам ключевым в определении предмета социологии, потому что содержит решающее качество самодвижения, развития социального целого»* (с. 18). Далее говорится — нам следует разобраться с взаимосвязью теории и эмпирического знания, с проблемой расчленения (так!) социологии на фундаментальную и прикладную, осознать необходимость в развитии специальных социологических теорий, уяснить, что же представляет собой фактуальная основа социологического знания и др.

1 Статья подготовлена в рамках выполнения проекта Российского фонда фундаментальных исследований «Статус социальной реальности в социологии. Разработка методологических, методических и технологических систем социологического изучения социальных объектов» (№06-06-80441), а также плановой темы СИ РАН «Многообразие принципиальных субъектов социологического исследования, характер их влияния на процесс и результаты исследования».

2 Автором методологических рассуждений (части I и II) является Г.И. Саганенко. Разработки и осмысление возможностей ДИСКАНТА и ВЕГИ (часть III) принадлежат всем трем авторам.

На этом пути мы почерпнули много полезного, но убедительного методологического знания у нас не получилось.

Как видим, во всех наших построениях, в наших научных книжках (а к нашей компании можно присоединить еще и ряд других основательных методологов, например, Б.З. Докторов, В.И. Паниотто, И.Ф. Девятко и др.) представлена масса логически выстроенных, формально убедительных конструкций, однако достаточно абстрактных и по-человечески не вполне понятных... В этих как бы по научному строгих системах обоснованности абсолютно не фигурировали такие «мелочи», как «эпоха на дворе», «социология» (в ее актуальном состоянии), «исследователь», «респондент», «деньги» и т.п. Кстати, идея «денег» не существовала вообще. Мы были бюджетной организацией, и как Владимир Александрович добывал нам право осуществлять очередной глобальный проект и соответственно иметь на это разные ресурсы — ставки, зарплаты, годы на реализацию проекта, грузик анкет (действительно такое было — везли из Москвы тираж методик для исследования «Ценностные ориентации») и пр., и пр. — такие вопросы нас вообще не затрагивали. Так что идея «презренных» ресурсов никак не входила в наши представления о методологии эмпирических исследований и о зависимости качества получаемых нами результатов от таковых.

Итак, посвятив методологическим бдениям более 25 лет, постепенно убеждалась, что мы мало что формально доказали, и в конце концов перестала играть в наукообразную социологию. Стала жить проще, радоваться чаще, объяснять социологию, двигаясь от нас и от нашей жизни, а не от абстрактной науки, ортодоксальных правил и «навороченных» текстов. Студент стал больше понимать, лучше осознавать себя как в социологии, так и в своей жизни, получать удовольствие и удовлетворение и, между прочим, создавать симпатичные социологические «продукты» и даже «нетленку».

Тем не менее, начиная очередное свое эссе о социологии, мне нужно кратко обрисовать «систему координат», в которой будет прописана моя локальная «единица» познания — компьютерная система ВЕГА. Каждая познавательная «единица» в моем подходе не существует сама по себе, а является производной *от других* и одновременно выступает производящим фактором *для других* элементов выстраиваемой мною познавательной системы.

О макро-субъектах социологического исследования

Эффективность исследовательского социологического предприятия при изучении социальных явлений зависит от ряда существенных факторов/составляющих/условий. В этот значимый ряд факторов мною включаются следующие основные макро-элементы — «эпоха на дворе» (в том числе социально-политическая ситуация в обществе), «социология», «социально-познавательная проблема», «исследователь», «респондент», «метод», «ресурсы» (в том числе деньги, время...), «заказчик» (Саганенко 2002). Поскольку все названные категории, так или иначе, оказывают влияние на масштаб исследования, получаемые в исследовании результаты и их качество, мы считаем их «субъектами» социологического исследования.

О «методе». Что касается «метода» исследования, то эту категорию можно представить интегральным понятием «тип социологического исследования». Типы исследований можно выделять по множеству оснований. Одним из оснований типологии исследований выступает метод актуализации/сбора данных. Метод сбора данных определяет специфику, трудоемкость, времяемкость реализации основных блоков исследования, нагрузку на исследователя на разных его этапах. Особенности первичных данных определяют в значительной степени характер и содержание финальных результатов исследования, путь следования к ним.

Идею *эмпирического результата* социологического исследования нужно рассматривать дифференцированно через три группы методов: (а) методы актуализации, фиксирования, измерения первичной социальной/социологической информации (среди них, в частности, находятся вопросники и условия получения по ним первичной информации); (б) методы формирования системной совокупности исходных данных (среди них методы обоснования системы признаков и системы объектов, методы сбора информации, процедуры реализации полевого этапа); (в) методы обработки и анализа массивов первичной/вторичной информации (качественные и количественные, текстовые и визуальные, статистические и сингулярные, описательные и объяснительные) (частично в: Саганенко 1983).

С макро-категорией «метод» находятся в наиболее активном взаимодействии такие «макро-единицы», как «исследователь», «респондент», «средства». Представим наши пояснения по названным «макро-субъектам».

Об «исследователе». «Исследователь» для нас — комплексная методологическая категория, которая предполагает понимание множества значимых познавательных аспектов, определяющих в значительной степени характер получаемых в исследовании первичных данных и итоговых результатов. Так отдельные реальные исследователи изначально имеют разный жизненный, профессиональный и непосредственно экспериментально-исследовательский опыт, владеют разными возможностями, имеют разную доступность к ресурсам разного рода, в том числе к информационным ресурсам и разным категориям населения, выработали определенные исследовательские предпочтения, базируются на определенных исследовательских парадигмах.

Исследователь, всю жизнь практикующий в рамках стандартизированной парадигмы (именно к нему чаще всего относится образ «социолог с анкетой») — тем самым, получает/имеет значительный опыт разработок стандартизированных вопросников, обоснования/организации статистических выборов, применения статистического компьютерного пакета для анализа социальных данных (например, SPSS), использования вероятностно-статистических методов, оценки устойчивости показателей и значимости их различий, получения специфических количественных результатов, их описания и интерпретации — и будет преимущественно выбирать и обосновывать такого рода исследования в своей исследовательской практике.

Исследователь-акционист будет искать объекты для полевой работы, ему будет интересно наблюдать социальные ситуации и процессы «с близкого расстояния», идти по исследованию в определенном смысле одновременно с социальным процессом — с развитием процесса и развитием представлений о самом процессе и его познании, ему будут скучны только числовые баллы, проценты, таблицы (то, что составляет итог работы классического «социолога с анкетой»). Среди акционистов, для примера, я бы назвала двух исследователей — Андрея Николаевича Алексеева с его известным в стране и мире экспериментом социолога — рабочего, но не только — во многом вся его научная жизнь состоит из «кейсов» акционистской науки/деятельности (См. Алексеев 2003-2005). В качестве второго примера я не побоюсь предьявить себя. У меня три основные акционистские сферы — наркомания в российском обществе (См. Саганенко. Лицом к лицу с наркоманией, 2005; www.azaria.rusorg.ru/). Сейчас готовлю книгу о наркозависимых, направленную на гуманизацию общества; гуманизация социологии — гуманизация образования (частично эта сфера представлена в моих книгах «Многообразие возможностей социологического изучения сферы образования», но больше делаю практи-

чески, в процессе преподавания); борьба за открытую и достойную науку в отдельно взятой организационной научной структуре (материалы на Интернетe)³.

Петербургский социолог В.И. Ильин, работающий в парадигматике качественного подхода, видит позицию «исследователя» в следующих вариантах: «исследователь как игрок», «исследователь как разведчик», «исследователь как советник», «исследователь как политик», «исследователь как путешественник» (Ильин с. 5-7). Попробовав себя в разных жанрах, сам Ильин делает свой персональный выбор в пользу «путешественника» — «просто путешествует в ходе своих полевых исследований из одного социокультурного мира в другой и рассказывает в своих публикациях и лекциях, что он там увидел» (с.7).

Самарский социолог качественной ориентации Анна Готлиб использует образ и идею «рефлектирующий социолог» (Готлиб, 2006). Отметим принципиальные для нас некоторые из ее высказываний о специфике взаимоотношения исследователя, метода, социальной реальности. «Метод означает способ видения и выговаривания действительности в такой же мере, в какой он конкретизирует техники и процедуры» (Готлиб, с. 5). Анна Готлиб отмечает, что разговор о качественном социологическом исследовании ведется в режиме постоянной переключки с классическим (количественным) исследованием, которое выступает как бы фоном (с. 5). Нам представляется, что да, действительно, эта методологическая коллизия присутствует у всех социологов, работающих с качественными методами.

Можно отметить, что только немногие отечественные исследователи владеют в достаточно продвинутой степени познавательным арсеналом разных парадигм.

Об «итоговом эмпирическом результате». Затрагивая идею получения эмпирического знания, считаем важным отметить, что полезно иметь представления о том, что же такое «итоговый эмпирический результат» и как он выглядит в исследованиях разного типа (проблема, в принципе не затрагиваемая в рассуждениях о подходах и методах). После презентации эмпирических результатов затем уже идут интерпретации и, наконец, словесно артикулированные выводы и заключения с минимумом эмпирического материала.

Итак, в стандартизированных, «количественных» исследованиях конечные результаты — это одномерные и двумерные таблицы, различные описывающие распределения параметры, показатели тесноты связи признаков, графики, гистограммы, структуры интегральных/вторичных показателей и веса в них исходных признаков (как, например, в факторном анализе), разного рода регрессионные и причинные уравнения, фиксирующие группы сходных объектов кластеры и таксоны и др.

Следует подчеркнуть, что в количественно-статистической парадигме с первого шага изучения данные о целостном человеке «рассыпают» на автономные признаки, и собственно далее исследователи занимаются только отдельными единичными признаками, парами признаков и сочетаниями. Конкретный человек далее выступает только как добавление 1 в массу, например, мужчин или женщин — если берется распределение по полу, или 1 в одну из градаций по возрасту — если рассматривается распределение по возрасту и т.п. Два признака одного респондента могут соединиться только в соответствующей двухмерке. Но определенно, что целостный или частично целостный человек исчезает в этого рода исследованиях. Таксономию, где берется множество признаков респондента и соответственно в каком-то смысле учитывается

«целостный» человек (ее редко кто использует — большинство исследователей достаточно парных зависимостей), да и тут есть условная мера «расстояния», еще надо справиться с отбором в процедуру признаков и их «разнокалиберностью».

Надо также осознавать, что применение любого статистического метода не дает единственного решения — на одном и том же исходном эмпирическом материале можно построить сотни причинных или регрессионных моделей, провести массу факторных анализов и таксономий, получить фактически бесконечное количество двумерных и многомерных таблиц. Фактически любой конкретный итоговый статистический анализ — лишь один из бесконечного множества вариантов, разъясняющих/просвещающих некоторую суть полученных данных. Подобные эмпирические результаты далее интерпретируются, обобщаются, возможно, соотносятся с определенной теорией/концепцией и приводят к содержательным, социально значимым выводам и заключениям.

В качественных исследованиях (построенных на больших текстах — интервью, биографиях, дневниках, ауторефлексиях и т.п.) эмпирические результаты представляются в форме множественного цитирования значительных фрагментов текстов, в приложении обычно прилагаются примеры полномасштабных интервью, биографий. Я предложила студентам использовать две сопряженные идеи — «вертикальный» (сплошной, интегральный по отдельным текстам) и «горизонтальный» (сквозной по отдельным аспектам через разные тексты) анализы. Примером массивной презентации текстовых материалов является книга Г.И. Саганенко «Лицом к лицу с наркоманией», в которой примерно 250 страниц из 326 составляют материалы четырех типов текстов — краткие актуальные истории с наркозависимостью из телефонных обращений родственников, беседы волонтеров с родителями, обсуждение продуктивных практик преодоления наркозависимости в формате фокус-группы, семейные биографии, написанные матерями. Они составляют отдельные главы и фактически; частично или полностью отдельные материалы проанализированы при поддержке программы ДИСКАНТ.

Исследования, построенные на системе открытых вопросов — это отличающаяся от других, качественно-количественная технология, ее результаты носят принципиально иной характер. Здесь поначалу респонденты при ответах на открытые вопросы делают «частичную концептуализацию» множественных сторон исследуемого социального объекта/ социальной коллизии. Затем исследователь изучает/выявляет синдром «социальной коллизии через респондентов и респонденты через социальную коллизию» и получает содержательные составляющие объекта, сопоставительный их вес и значимость в общей картине объекта (разработка концепциотехнологии в: Г.И. Саганенко, 2004).

Другие примеры презентации «результатов» см. у А. Готлиб («Интервью с социологом», с. 420 — 439), у В. Ильина («Анализ текстовых данных», с. 206 — 221).

Исследователи, работающие в качественной парадигме — тоже, так или иначе, прибегают к обоснованию частотами, процентами, реже зависимостями и др. (Об опыте сочетания качественной и классической методологий..., см. Готлиб, с. 303-359).

О «респонденте». В социологическом исследовании особая миссия у «респондентов» («информантов» и др.), если таковые включены в исследовательский процесс. Если респондента привлекают к заполнению стандартизированного вопросника, тем более, если сбор данных проводится в рамках процедуры стандартизированного интервью (при которой респондент даже не видит целостного вопросника, не осознает контекст задаваемых ему

3 <http://www.unlv.edu/centers/cdclv/archives/Supplements/siran.html>; «Международная биографическая инициатива» - http://www.unlv.edu/centers/cdclv/archives/Interviews/saganenko_golofast.html

вопросов и пр.), то респондент имеет минимальные возможности предъявиться как индивид, оказаться партнером исследовательского процесса, он только «инструмент», производящий данные, к тому же в специфическом формате и слабо отрефлексированной ситуации.

Если же респонденты участвуют в исследовании, строясь на интервью, биографиях, дневниках, открытых вопросах, то они в определенном смысле становятся партнерами исследователя, делятся с исследователем своими временными, эмоциональными, интеллектуальными и информационными ресурсами.

Можно также отметить, что индивидуальные и групповые способы *опроса* имеют принципиально разные познавательные характеристики, разную автономию респондента, разные условия анонимности, разный кумулятивный результат. Вообще *групповые опросы* формируют/предъявляют особый тип респондента — *респондента группового*. Работа отдельного респондента в группе влияет на группу и работа/отношение группы оказывает влияние на отдельного респондента. Оттого, удалось ли сделать аудиторию своим союзником, увлечь перспективной полезностью/привлекательностью отвлечения по проблеме, зависит качество полученного первичного материала. Анкеты на открытых вопросах в итоге обеспечивают глубокие, полноценные суждения, позволяющие респонденту «открыть» самого себя для себя и исследователю «открыть» аудиторию через респондентов.

Групповые опросы в «*стационарных*» аудиториях имеют и особый акционистско-познавательный ресурс (который, кстати, не отмечен в методологических работах) — такие опросы содержат принципиальную возможность «обратной связи» с данной аудиторией респондентов, проведения повторных обращений, сравнительных исследований, уточнения и расширения исследовательских, коммуникационных, гуманистических задач с изучаемыми группами людей.

Надо отметить, что любое общение с аудиторией в рамках социологического опроса/беседы/дискуссии — ограничено по времени, чаще всего это 1 или 2 академических часа. Ни одна социально-значимая сфера/тема/социальная коллизия (самореализация, здоровье, образование, повседневность, взаимоотношения поколений...) не может быть достаточно основательно прояснена/освоена человеком в рамках такого короткого промежутка времени, хотя, несомненно, может дать достаточное количество данных даже для докторской диссертации. Социальная коллизия «*полезности*» социологических «*рейдов*» для респондентов и аудиторий нигде не поднимается — миссия респондента, вроде, только в том, чтобы быть полезным социологу. Развивая и практикуя гуманистическую социологию, мы поэтому предпочитаем работу/коммуникацию со «стационарными» аудиториями, позволяющую студенту-респонденту или учащемуся в классе общеобразовательной школы глубже освоить отдельную социально-значимую сферу при систематической поддержке социологии.

II. Методология открытых вопросов

Исследования, построенные на открытых вопросах, имеют несколько гуманистических эффектов. Прежде всего, предлагают респонденту формат и условия для осмысления коллизий собственной жизни и окружающего общества, для приобретения определенного опыта социальной рефлексии, систематического вербализованного оформления своих собственных представлений.

Исследователь в таких исследованиях также получает нового типа исследовательский материал — имеет вербализованные свидетельства о социальной реальности, индивидуализированные презентации отдельных респондентов, получает множество первичных концептуализаций

социальных процессов/объектов со стороны множества респондентов/экспертов, в конечном итоге выстраивает устойчивую картину социального процесса/социальной коллизии на актуализациях «коллективного эксперта» — совокупности респондентов. Исследователь имеет базу для получения содержательно и статистически доказательных результатов и выводов, в своих публикациях имеет широкую возможность иллюстрировать свои аргументы, предоставляя возможность звучать «голосам молчащих» — респондентам.

Методология открытых вопросов принципиально отличается от качественного исследования по своим познавательным, гуманистическим, коммуникационным аспектам. В качественном исследовании основным героем социологической «драмы» выступает «исследователь» — он-де путешествует, наблюдает и рассказывает (как у Ильина), это способ видения и выговаривания действительности (соответственно исследователем), это идея диалогичности (активность познающего и активность открывающегося), это сложность двухстороннего акта познания-проникновения (как у Готлиба).

Что касается нашего подхода с «эксплуатацией» открытых вопросов и включением респондентов в исследовательский проект, то наши познавательные задачи артикулируются существенно иным способом. Презентации и напряжения в таком исследовании двух типов участников — исследователя и респондентов — прописаны в разных местах и в разных видах познавательной и коммуникационной деятельности. В нашем случае исследователь перекладывает проблему концептуализации некоторой значимой социальной коллизии/объекта на респондентов. С этой целью он действует следующим образом: (1) выбрав интересующую его социальную коллизию, создает привлекательный для рефлексии инструмент — вопросник с открытыми (и иными) вопросами; (2) «договаривается» с аудиторией — общается с аудиторией респондентов, создает у них мотивацию участвовать в обсуждении (поиске/записи ответов на вопросы анкеты) и тем самым потратить на исследователя свои временные, интеллектуальные, информационные и эмоциональные ресурсы; (3) респонденты, во-первых, формулируя и записывая текстовые суждения, сами разбираются по сути вопросов/формируют свои представления об объекте и, во-вторых, осуществляют основную для исследователя работу — дают частичные описания многообразных сторон изучаемого(ой) объекта/ проблемы/своей жизни; (4) исследователь проводит трудоемкую работу — изучает вариацию высказываний респондентов по каждому вопросу, ищет фокусы внимания/акценты респондентов, концептуализирует обнаруженную систему акцентов, создает обобщенную картину объекта через представления респондентов.

В принципе исследователь должен/может осуществить переход к социально значимым обобщениям и выводам за счет многократных итераций, должен описать процесс и результаты исследования привлекательным и убедительным текстом. Тем самым такая социология должна быть привлекательна «*читателю*», знакомящемуся с нашими опусами. Читателем может быть и наш коллега по научному цеху, и эксперт, и обыватель. Мы полагаем (может быть наивно), что наш текст «Представления молодежи о стране эпохи СССР» на материалах рефлексивного исследования окажется читабельным и для коллеги-социолога, и читателя-неспециалиста (Саганенко, Воронцова, 2008).

Итак, мы, надеюсь, убедили читателя, что наши проблемы как исследователя, базирующегося на открытых вопросах, в корне отличаются от проблем и нагрузок специалиста, работающего с другим типом исследований.

Категория «средства» — объединяет многообразные составляющие (включая методические, финансовые, технологические и др. средства, а также человеческие ресурсы). В данной статье мы отметим только особые возможности для качественных (в нашем случае базированных на текстах) исследований, которые обеспечивают специально разработанные для их поддержки и повышения их эффективности две *компьютерные технологии* ДИСКАНТ и ВЕГА. Они созданы в Санкт-Петербургском экономико-математическом институте РАН и предоставляют социальному исследователю разнообразные средства хранения информационных массивов исследований, их наращивания, сравнения, слияния, средства получения многообразных словарей, инструменты доказательной классификации массивов текстовых единиц, концептуального анализа сплошных текстов, структурного сравнения текстовых массивов и многое другое. Обеспечивают возможность оперативного изучения текстовых суждений респондентов, что позволяет приходить повторно в аудиторию с «адресным» аналитическим материалом, делиться полученными выводами с участниками опроса, организовать дискуссию по изучаемой социальной проблематике. Эти программы обеспечивают многообразную поддержку качественных (с текстовой и смешанной первичной информацией) исследований, предоставляют возможность работать с цифровой (альтернативной и многоальтернативной, единичной и табличной) информацией. Система ВЕГА — это функционирующая с 2007 г. современная (разработанная в среде WINDOWS) версия пионерной отечественной разработки ДИСКАНТ (в системе DOS).

III. Компьютерная технология ВЕГА как познавательный ресурс социологии⁴

Считаю принципиальным дать систематическое представление конкретной компьютерной программы поддержки качественных исследований вот по таким ТРЕМ основаниям.

(а) Факт отсутствия важности упоминания компьютерной поддержки в публикациях. Дело в том, что, сколько бы ни знакомились с отечественными описаниями качественных исследований, фактически почти нигде не встречаешь каких-либо упоминаний, что работа сделана при поддержке релевантной компьютерной программы. Есть лишь весьма редкие исключения.

Но ведь качественные текстовые исследования — это по большей части огромная масса (часто неструктурированного) материала, и проводить анализ только в текстовом редакторе типа Ворда явно недостаточно. Ведь каждый «сеанс» анализа без соответствующей программы в принципе дает значительные случайные колебания, заворачивает исследователя на случайные исследовательские «тропинки», так что в конечном итоге исследователь «волевым» образом фиксирует одну определенную версию, прописывает некий вариант интерпретационного текста.

4 Программа ДИСКАНТ в среде DOS была разработана в течение 1991-1996 гг. в Санкт-Петербургском экономико-математическом институте РАН (разработчики Е.А. Каневский, Г.И. Саганенко, Л.М. Гайдукова, Е.Н. Клименко). Получала поддержку фонда Сороса (Института Открытое общество) и дважды РФФИ. Демонстрировалась на 13-м и 14-м Международных социологических конгрессах (Билефельд, Монреаль), на 4-й и 5-й Международной конференциях по методологии социальных наук (Колчестер, Кельн). Новая программа в среде Windows разрабатывается с 2004 г. там же — в СПб ЭМИ — разработчики Е.А. Каневский, К.К. Боярский, Г.И. Саганенко, поддержка РФФИ (грант 06-06-80441).

(б) Влияния технологии на результаты исследования. На самом деле аналитический программный ресурс выступает активным «субъектом» исследования, обеспечивая исследованию динамичность, глубину, широту, доказательность, надежность исследовательского поиска и др. Когда же у исследователя отсутствуют такие ресурсы и соответственно указанные представления, тогда упоминание характера аналитической компьютерной поддержки в презентациях/публикациях просто отсутствует/опускается.

(в) Считаю, что в отечественной социологии есть потребность в демонстрации конкретной эффективной компьютерной системы в поддержку качественных исследований социологов.

Проблема *компьютерных технологий* в социальных науках — специфический методологический раздел, он либо вообще не называется исследователями-социологами, либо отдан на презентацию специалистам другого профессионального цеха. Мы знаем относительно немного работ наших коллег, которые систематически рассматривали и/или описывали компьютерные технологии (среди них А.О. Крыштановский, О.Б. Божков, С. Чесноков, А.Н. Чураков и немногие другие).

Методы анализа и компьютерной поддержки можно разделить на два широких класса — применительно к стандартизированным/количественным исследованиям и применительно к качественным исследованиям. Тему компьютерного сопровождения стандартизированных исследований не будем затрагивать. В принципе это хорошо разработанная сфера, масса публикаций, популярный пакет SPSS (правда, здесь недостаточно прояснены такие проблемы, как множественность и неопределенность результатов, устойчивость процедур для разных условий, сопоставимость результатов разных методов и др.).

Специализированная компьютерная поддержка качественных исследований используется только некоторыми исследователями: Елена Омельченко из Ульяновска упоминала об использовании в их исследованиях одного из известных зарубежных пакетов, в коллективе И.И. Травина из Санкт-Петербурга используется пакет, предоставленный финскими партнерами; ФОМ использует некую собственную компьютерную разработку и ряд других.

О кодировании единиц текстовых массивов. Как правило, при использовании компьютерных средств в качественных исследованиях основная коллизия, которая является серьезной и обсуждается методологами и эмпириками, — проблема кодирования информации. Долгое время мне не удавалось разобраться с описаниями, касающимися кодирования у «качественных» исследователей, как у зарубежных, так и отечественных. Поскольку мы, при использовании ДИСКАНТА и ВЕГИ, имеем непрерывное взаимодействие как с текстами, так и с классами/группами, к которым они приписываются, а также с их соответствиями, то у меня и мысли не было, что может быть что-то другое.

Начнем с того, что наше первичное понимание предьявленных респондентами коллизий не является окончательным, а только начальной итерацией работы с тестовыми данными. Разработка классификатора (системы позиций) и классификация (приписывания суждений к классам) идут итерациями, параллельно «рука об руку», уточняя друг друга. Мы имеем возможности менять наши первоначальные представления — менять названия классов/групп, их «внутреннюю» логику (то есть менять их последовательность и группирование), по мере прояснения материала и своих представлений изменять «прописку» суждений — отнесение таковых к конкретным классам/группам, в конечном итоге мы приходим к какому-то оптимальному результату.

Спустя 10 лет я, наконец, поняла специфику ситуации кодирования. Представляется, что те, кто пользуется компьютерной поддержкой текстовых исследований, построенных на процедурах кодирования, во-первых, решают задачу введения кодов в самом начале процесса, во-вторых, решают ее сразу и окончательно,

в-третьих, разрывают органическую связь текстов и кодов. Тексты могут остаться существовать только на бумаге, в заполненных анкетах, в отдельных текстовых файлах, а коды — отдельно, например, как база данных в SPSS. Так вот, оказывается, почему в аналитических публикациях встречаешь только какие-то иллюстрации текстовыми пассажами (в большем или меньшем количестве, в более или менее просторном виде), а для выделенных рубрик (кодов) выдаются только проценты!

Так что моя настоятельная забота, связанная с разработкой этой статьи — систематически показать значительный познавательный ресурс, который обеспечивает комфортная и эффективная компьютерная система поддержки качественных исследований, которую представляет собой отечественная система ВЕГА и которая в принципе доступна для пользователей в нашей стране. Люди либо не хотят знать, либо, не имея опыта, считают, что тут вряд ли что будет полезным. Мне представляется, что из-за отсутствия доступных программ слабо используются качественные исследования, медленно идет развитие их методологии. Мною, в том числе, двигает своего рода гуманизм, поскольку понимаю насколько тяжело исследователю работать с тестовыми массивами, насколько это мало эффективно.

3.1. Познавательная эффективность ДИСКАНТА-ВЕГИ

Здесь можно отдать должное системе ДИСКАНТ, которая за 15 лет своего существования позволила освоить разнообразие перспектив данного подхода, создать новые интеллектуальные «продукты», разработать новые познавательные ресурсы, предложить принципиально новые методологические задачи и их решения (Многообразие... Часть 1. С. 22). Эти разработки касаются осмысления методов и методологии социологических исследований, специфики социальной реальности и массового сознания, принципов преподавания, назначения социологии и ряд других.

Один из существенных результатов нашей связанной с ДИСКАНТом исследовательской активности — обоснование самостоятельного типа исследований — *«исследования, базированного на системе открытых вопросов»*. Этот тип исследования обладает особым потенциалом, позволяющим изучать и сравнивать различные сложные и изменяющиеся социальные объекты, доказательно выявлять новые социальные координаты (dimensions), понимать особенности сознания людей, добиваться привлекательной и полезной коммуникации с аудиториями, обеспечивать определенную пользу от социальной рефлексии для аудитории, респондента, самого исследователя, читающей публики.

Этот метод позволяет поддерживать интенсивным и привлекательным учебный процесс. Более того, благодаря этому подходу стало понятно, что такое гуманистическая социология, каковы должны быть отношения между изучающим и изучаемым, между обучающим и обучаемым.

Для данного типа исследований вместе с программистами СПб ЭМИ РАН удалось создать комфортный (user-friendly) и эффективный программный продукт — компьютерную систему ДИСКАНТ, к настоящему времени мы уже переключились на выросшую в новых условиях ее «модернистскую дочь» — программу ВЕГА. Уже многие годы мы, вместе с коллегами и массой студентов, смело беремся за изучение разнообразных социальных сфер,

доверяя нашим респондентам описывать сферы окружающей жизни. Наши рефлексивные методики адекватны разным категориям людей, они обеспечивают нас разнообразным материалом, релевантным эпохе, социальной сфере и социальной группе людей.

Среди наших респондентов были научные работники 30 ленинградских академических учреждений, академики и члены-корреспонденты АН СССР в 1992 г., читатели залов для научных работников Государственной публичной библиотеки им. Салтыкова-Щедрина (затем РНБ) в течение 1990-2000 г. многократно, научные сотрудники Дании и Швеции, участники международной конференции на Аландских островах, рабочие промышленных предприятий, огромная масса школьников, студентов, учащихся ПТУ, масса наркозависимых, масса вынужденных принять на свои плечи беду наркомании родителей и др.

Мы убеждены, что респонденты вместе всегда знают больше, чем один исследователь — наша задача лишь предложить им привлекательную/ значимую тему и комфортный инструмент для их персональных рефлексий и концептуализаций, а затем уже накопленный нами исследовательский опыт и программа ВЕГА (ДИСКАНТ) помогут выявить новые интересные факты и координаты в любой социально-познавательной коллизии.

Опыт, который имеется в нашем сообществе — это огромный объем (свыше 50) разнообразных качественных исследований, эффективное освоение и продвижение компьютерных технологий, базирование результатов исследований на серьезной доказательной базе, широкое внедрение технологии в образовательную подготовку социологов.

Методологические, методические, компьютерные и практические разработки, выполненные в рамках методологии открытых вопросов, позволяют нам заявить, что мы создали самостоятельную, полномасштабную, целостную исследовательскую технологию для изучения актуальных социальных явлений. При поддержке и активном участии ДИСКАНТА было создано многообразие различных познавательных продуктов. Сейчас исследовательская система/технология развивается в рамках системы ВЕГА. Все ее познавательные элементы имеют самостоятельное значение и одновременно обеспечивают целостность и устойчивость данной познавательной технологии/системы.

Систему составляют следующие макро-единицы:

- 1) Разработана методологическая концепция, систематически описывающая познавательный и гуманистический потенциал данного подхода (Саганенко. Многообразие возможностей... Часть 1, с. 19-45);
- 2) Разработана система открытых вопросов, описан их познавательный потенциал (Часть 1, с. 46-72);
- 3) Разработано множество опросных методик, построенных на системном использовании открытых вопросов (15 методик на примере изучения сферы образования опубликованы в Части 2; Саганенко 2008, с. 184-186). Методики представляют некую органическую целостность, а не просто совокупность отдельных открытых вопросов. Данного типа методики разработаны для изучения проблематики из сферы образования, наркомании, антинаркотической политики, здоровья, семьи, экологии, для мониторинга изменяющихся параметров трансформирующегося российского общества (ситуация отслеживалась регулярными панелями на интервале 1989-2000 гг.), для сравнительного изучения актуального состояния общества в восприятии жителей разных стран (России, Дании, Швеции) и др.
- 4) Проведены свыше 50 исследований разного типа, построенных на принципиальном использовании текстовой информации и методе открытых вопросов, доказательной классификации и анализе множественных

массивов текстов. Для примера, упомянем лонгитюдное исследование 1989-2000 гг. с ежегодными панелями, которые каждый раз приносили нам по 7-10 тысяч высказываний примерно по 30-40 аспектам публичной и индивидуальной жизни.

5) Разработана компьютерная система ДИСКАНТ/ВЕГА (первая — в среде DOS, вторая — в среде Windows) в поддержку исследований, базированных на текстах. Особо развитые возможности предусмотрены для исследований, построенных на структурированной текстовой информации, а именно — на вопросниках с системным использованием открытых вопросов, обеспечивающих значительные массы текстовых суждений. Программа позволяет оперировать системой из множества автономных текстовых массивов, которую, по сути, представляет первичная информация такого рода исследований.

6) Упакована и хранится в системе ДИСКАНТ/ВЕГА первичная информация свыше 50 исследований. Первичная информация исследований проанализирована частично или достаточно полно. Доступна для многократного обращения, используется при обучении студентов-социологов. Используется для проведения новых «корреспондирующих» исследований, расширяя круг включенных в исследование категорий респондентов, модифицируя базовые исследования с учетом происходящих перемен и новых познавательных задач, сохраняя при этом значительный потенциал сравнения.

7) ДИСКАНТ/ВЕГА обеспечивает возможности эффективной поддержки и анализа текстовых исследований разного типа: (а) исследования на базе открытых вопросов (выборки до 3000 респондентов с 20-60 и даже со 100 открытыми вопросами); (б) исследования на системе объемных интервью (см. пример такого исследования «Освоение социальных пространств иногородними студентами», Часть 2, с. 62-79); (в) исследования на базе неструктурированных коротких эссе (см. пример такого исследования «Учащиеся профтехучилищ о недавней школе»// Саганенко, часть 2, с. 125-134); (г) исследования на базе больших биографических текстов (см. книгу «Лицом к лицу с наркоманией» — свыше 350 биографических повествований родителей, столкнувшихся в наркозависимостью своих детей); (д) анализ целостных концепций; (е) лонгитюдное качественное исследование на длительном (в 12 лет) интервале — о переменных в мире, стране, личной жизни; (ж) международное сравнительное текстовое исследование (на три страны) — о восприятии перемен в стране и мире;

8) В рамках ВЕГИ (ДИСКАНТА) разработан особый (отсутствующий в других программах) аналитический инструментарий для проведения систематической, доказательной, оптимальной *классификации массивов текстовых суждений*. Инструментарий классификации позволяет находить и фиксировать фокусные темы/рубрики/разделы/параметры в ответах респондентов (суждениях экспертов, материалах статей, текстовых описаниях характеристик объектов) и, в конце концов, выявлять «тематические» структуры в текстовых массивах и их вес. Система классификации позволяет двигаться итерациями — начинать с выявления простых или популярных элементов в описании/характеристиках изучаемой социальной коллизии (объекта, процесса), содержащихся в массиве текста/текстов, и в конечном итоге достигать некоторого ее/его оптимального описания.

Система содержит одновременно исходные тексты, классификаторы, результаты классификации, выводит соответствия классов и текстовых единиц по любому заданному условию.

В ВЕГЕ есть также определенные возможности проводить классификацию применительно к большим произвольным текстам.

9) В итоге обоснован самостоятельный тип эмпирического социологического исследования, построенного на открытых вопросах. Проведено также систематическое сравнение специфики данного типа исследований со спецификой классического стандартизированного исследования применительно к основным этапам эмпирического исследования.

10) Имеется широкий спектр вариантов применения данной технологии. В частности, во многих сферах можно использовать своего рода анкеты в виде перечня позиций/вопросов, предполагающих текстовые описания, а также номинальные кодовые значения и количественные оценки, вести развивающиеся во времени, пополняющиеся «картотеки». Например, проводить таким образом анализ публикаций, фиксировать нормативную и творческую историю участников некоторого лонгитюдного наблюдения (например, студентов в процессе обучения), отслеживать параметры изменяющегося процесса, состояние его участников и др.

11) Важно, что есть опубликованное описание — технология открытых вопросов последовательно описана в двух монографиях (Многообразие возможностей... Части 1 и 2: 2004, 2005) и, тем самым, имеются релевантные возможности понять идеологию исследовательской системы, узнать и освоить основные ее модули.

3.2. Эмпирико-аналитические блоки системы «ВЕГА»

Компьютерная система ВЕГА представляет собой систему хранения, обработки, классификации и анализа текстовых массивов. Перечислим основные варианты «результатов», которые обеспечивает ВЕГА.

(1) Систематическое хранение многокомпонентной эмпирической информации. Это возможность, создав своего рода картотечный ящик с соответствующими рубриками (систему «*полей*», в частности для вопросов), упаковать туда первичную эмпирическую информацию по объектам («*затисям*», в частности, по респондентам). При этом разрешается использовать множество разного типа признаков — текстовых, числовых, балльных, табличных, смешанных («*двойные поля*» — с цифровой и текстовой частями). Текстовые поля могут представлять разное по форме содержание: длинный гладкий текст; развернутый многоставный ответ/фрагмент/суждение; разрезанный на тематические фразы полный ответ, ключевые слова. ВЕГА обеспечивает комфортный доступ к информационному массиву, возможность его коррекции, внесения дополнений и изменений, наращивания этого массива, его оптимизацию.

Пример исходной информации. Для исследования «**Ценности и цели получения высшего образования**» разработана анкета, которая имеет 50 вопросов, из них 40 открытых, при этом часть из них имеет весьма сложную конструкцию. В исследовании 2007-2008 гг. приняли участие 150 студентов и в сумме они дали 15 тысяч суждений.

Пример ответа на вопрос: «Каждый человек стремится к самореализации. Что для Вас означает «самореализация в полной мере» и что Вам для этого нужно?» — респондент №1 написал следующие рассуждения: «Для меня — это *затрата всех своих сил на достижение ЦЕЛИ или ИДЕИ, а сил у меня, слава богу, хватает.* Самореализация в полной мере — *стремление ДОСТИГНУТЬ РЕЗУЛЬТАТА, не только в определенной своей области, но и что-то новое все время пробовать.* Для самореализации мне *хватает сил, интереса и обыкновенной ручки*. И такого рода ответы на первый вопрос у 150 человек.

Как видим, этим респондентом был дан достаточно пространственный ответ из 50 слов, в котором представлено три самостоятельные идеи: наличие цели и идеи

и готовность приложить значительные усилия; идея достижения некоего обобщенного результата; характер требуемых для этого ресурсов. В программе полный ответ делится на независимые «фразы» знаком $\langle \rangle$; совокупность всех «фраз» по каждому полю и подлечит систематической классификации и последующему анализу.

(2) Исследователь часто имеет потребность работать с частичными массивами. Предусмотрена возможность выводить исходные данные по отдельным «записям» и любому их набору («записи» — это в опросах респонденты, в других же исследованиях соответствующие объекты — единицы наблюдения), по отдельным «полям» и любому их сочетанию с учетом любого условия формирования выборки («поля» — это в опросах отдельные вопросы, в других исследованиях — это введенные пункты описания объектов).

Пример. Вывод ответов на два корреспондирующих вопроса

1
 *01** Наплевательское, равнодушное отношение к окружающей природе || -
 *02** проживание в неблагоприятных экологических условиях | болеет вся семья || -
 # 2
 *01** курят | плюют и мусорят на улицах || -
 *02** дышим и пьем неизвестно что | едим некачественную еду || -
 # 3
 *01** большинство относится наплевательски || -
 *02** стараемся не вредить, но никак не способствуем исправлению || -

Просмотр корреспондирующих вопросов позволяет нащупывать гипотезы для уточнения, обнаружить «координаты» для анализа. Как в этом примере — мы встречаем массу формулировок, затрагивающих ситуацию со здоровьем как характеристику взаимоотношения с экологией.

(3) Имеется возможность получать разнообразные словари: с учетом любого условия отбора полей и записей; получать словари слов, «ключевых» слов (то, что выделяется в БД прописными буквами), «фраз» (полный ответ разделяется на «фразы», представляющие самостоятельные суждения, фиксирующие одну идею), частотные (с подсчетом числа упоминаний текстовых единиц) и «адресные» (с указанием номеров записей и номеров полей, в которых каждый текстовый фрагмент присутствует) словари. Есть возможность создавать словари для любого одного текстового поля, любого сочетания текстовых полей и всех полей одновременно.

(4) Разработанная развитая подсистема проведения классификации. Это возможность: (а) выявлять «классы» (повторяющиеся или значимые темы) в совокупностях суждений, создавать и закреплять «классификатор» — каркасную систему позиций (идей/тематик), представленных в анализируемом текстовом массиве; (б) создавать «классификационную модель» — «классификатор» с набором «образцовых фраз», которые представляют примеры суждений, адекватных для каждую позицию классификатора; (в) проводить «идентификацию» — процедуру соотнесения каждой исходной «фразы» из массива ответов с определенной «базовой фразой» (являющейся представителем определенного класса/группы), тем самым происходит опосредованное приписывание каждой исходной «фразы» с конкретным классом/группой. В классификаторах ВЕГИ предусмотрено два уровня позиций в классификаторе — «классы» (более крупная рубрика) и «группы» (как деления внутри отдельных «классов»).

3.3. Эмпирические результаты процедуры классификации.

Покажем, что выступает как результат процесса классификации текстового массива:

(а) Сам «классификатор», который представляет последовательный перечень классов и групп, и тем самым является своего рода концептуальной системой структурного описания анализируемого массива.

Принципиально иметь в виду, что в любом таком исследовании мы имеем множество открытых вопросов или блоков открытых вопросов, соответственно множество текстовых массивов и, соответственно, множество классификаторов и тем самым множество концептуальных систем.

(б) Статистические таблицы распределения частот (количества фраз в классах и группах) для данного расклассифицированного массива, любого отдельного его подмассива.

(в) «Текстовые распределения» — система совокупностей конкретных фраз, отнесенных к каждому классу/группе классификатора из данного анализируемого массива.

(г) Индивидуальные сочетания классов в ответах респондентов конкретной социальной группы с указанием частотности конкретных сочетаний.

(д) Сравнение итогов классификации для заданных групп респондентов (например, сравнивать структуры массивов текстовых ответов для мужчин и женщин, для студентов разных вузов или разных курсов, для местных и приезжих студентов).

(е) Есть интересные возможности классификации и последующего анализа/ сопоставления тематики/структуры «корреспондирующих вопросов». «Корреспондирующие вопросы» — это такие вопросы, которые имеют некий общий стержень или общий контекст в содержании своих апелляций.

Как правило, в большинстве наших анкет содержатся блоки «корреспондирующих» вопросов — например:

- «Почему сейчас люди получают высшее образование?» и «Почему Вы лично получаете высшее образование?»;
- «Что бы Вы убрали — из жизни страны? — ... из жизни семьи? — ... из своей жизни?»;
- «На что бы Вы потратили — 5 тысяч рублей?, ... 5 тысяч долларов?»;
- «Чтобы вы отметили как самое существенное в состоянии экологии — в мире? ...
- в Европе? ... — в России? ... — в Санкт-Петербурге?»;
- «В получаемом Вами образовании отметьте плюсы; ... отметьте минусы...»
- (См. многообразие примеров в Книге 1: с. 46-72; см. Книга 2: методики 3, 4, 11, 13).

Если ответы на такие вопросы расклассифицировать, создав единый общий классификатор, то затем можно легко и убедительно сравнивать структуры массивов ответов, выявлять и интерпретировать обнаруженные совпадения и различия по «корреспондирующим объектам». Имеется возможность получать результаты классификации в указанных выше формах для каждого из корреспондирующих признаков, любого их сочетания, для всех вместе.

Как правило, мы стараемся проводить классификацию суждений на «корреспондирующие» вопросы по одному общему классификатору.

Например, можно сравнить характер запросов респондентов, которые они бы реализовали «на пять тысяч рублей» (как бы текущих запросов) и «на пять тысяч долларов» (как бы стратегических запросов).

Можно отметить, что идея «*корреспондирующих открытых вопросов*», осмысление их методологической специфики и познавательных возможностей вообще не

представлены в социологии — ни в отечественной, ни в зарубежной.

(е) Поскольку каждая анкета состоит из множества вопросов, направленных на различные актуальные сферы, соответственно мы получаем множество совокупностей высказываний по определенной тематике и в итоге самостоятельные текстовые подмассивы, которые требуют проведения самостоятельных классификаций. Максимальное количество требуемых классификаций в отдельном исследовании равно количеству отдельных открытых вопросов. Но, как мы показывали в предыдущем пункте — как правило, любая анкета состоит из некоторого количества групп корреспондирующих вопросов (блоков вопросов) и отдельных единичных вопросов. И тогда получается количество в принципе возможных систем классификаций равно количеству блоков и отдельных «независимых» вопросов.

Как правило, не в каждом нашем исследовании удается систематически проанализировать всю информацию и соответственно сделать полновесную классификацию по каждому вопросу/блоку вопросов анкеты. Основные усилия обычно мы направляем на основные блоки анкеты, а другие открытые вопросы анализируем/представляем в итоговом описании либо иллюстрациями, либо вариациями примеров, либо примерным выявлением тематического набора ответов за счет анализа словарей.

(6) Для цифровых признаков в ВЕГЕ предусмотрены определенные возможности анализа. Можно проводить собственное шкалирование — использовать разные варианты квантования признаков, например, исходную запись возраста или дохода респондентов можно делить на 2,3,7, (в принципе на любое количество) групп, искать оптимальное (под какой-то критерий) их разбиение на интервалы или использовать некое формально заданное разбиение признака.

Для многоальтернативных признаков предусмотрено выявление всех наличных сочетаний ответов (выборов), а также выдаются все сопряженности по всем двум, трем и любым другим сочетаниям ответов.

Можно получать одномерки, двухмерки, в том числе с учетом любых условий формирования подмассивов респондентов.

В принципе можно из базы данных ВЕГИ выводить подмассив цифровых признаков и анализировать его с помощью более развитых статистических средств, импортируя таковые в SPSS, Excel, другие аналитические пакеты.

(7) Есть возможность более продуктивно работать с цифровыми «табличными» признаками. Примером такого «табличного признака» может выступать блок «55 × 6» — оценки 55 профессий по 6 критериям в 10-балльной шкале (Книга 2: с. 111).

(8) Есть новые возможности эффективно работать с «*корреспондирующими исследованиями*» — это когда мы имеем исследования либо на базе полностью идентичных анкет, проведенные лишь в разных условиях (на разных группах, в разное время и др.), либо исследования, имеющие лишь частично совпадающие совокупности исходных признаков/ вопросов. Система ВЕГА позволяет объединять в единый массив «корреспондирующие» базы данных — единый массив в любом случае облегчает работу, обеспечивает в рамках ВЕГИ эффективные возможности анализа и сравнений. Соответственно это позволяет проводить систематические доказательные сравнения текстовых описаний, полученных в разных средах. Так, например, мы создали единую базу данных из трех исследований, проведенных на тему экологии среди научных работников в Санкт-Петербурге, Ольборге (Дания) и Маарилахамине (Аландские острова), хотя вопросы имели определенные различия. Исследования

содержали описания ситуаций в мире, стране, собственном городе, в том числе был значительный блок вопросов по экологии.

ВЕГА позволяет готовить из некоей «базовой» структуры обновленную структуру базы данных, которая в новом исследовании будет отличаться добавлением/убиранием некоторого количества новых/старых признаков.

(8) ВЕГА обеспечивает возможность переноса системы классификации, разработанной на одном «корреспондирующем» исследовании, для использования в других, «корреспондирующих» с ним, исследованиях (речь идет только применительно к одним и тем же отдельным вопросам). И главное — позволяет систематически сравнивать результаты текстовых рефлексий — и тем самым находить особенности различных социальных сред, социальных групп, социальных коллизий через анализ артикулированных суждений.

3.4. Основные функциональные модули системы ВЕГИ

Основная задача этого раздела — не столько аналитическая, сколько демонстрационная — показать разные аналитические модули и возможности — как метода открытых вопросов, так и разработанной аналитической программы ВЕГА.

Опишем основные функциональные блоки систем ВЕГА:

1. Система создания Базы данных реляционного типа: ее структуры и ее заполнения.
2. Система получения различных словарей. Возможно получение словарей полных ответов, «фраз», слов, ключевых слов. Возможен пермутационный вывод содержания любого текстового массива. Словари выступают первичным ресурсом анализа текстов.
3. Средства классификации: включая создание формата классификации (классификатора), проведение классификации текстовых суждений, приведение классификатора/ классификации к оптимальному виду;
4. Средства анализа текстовых массивов, сравнения структур корреспондирующих текстовых массивов.
5. Средства шкалирования цифровых признаков, средства статистической обработки цифровых и количественных признаков.

База данных в системе ВЕГА

Создание **Базы данных** (здесь и далее термины и категориальные понятия системы ВЕГА будут выделены жирным шрифтом) состоит из двух этапов:

- (а) разработка **структуры БД** и
- (б) заполнение структуры первичными данными.

Структура БД в ВЕГЕ представляет систему **полей**, каждое поле имеет **название**, зафиксированный **тип поля**, запланированную его **длину**. База данных в системе ВЕГА представляет совокупность **записей** — описания изучаемых объектов согласно разработанной структуре.

Типы полей в БД

В системе ВЕГА предусмотрены 8 типов полей, под основные из них разработаны специальные средства последующего анализа. Опишем их:

Поле «Символьное» — может содержать любую последовательность символов, которые понимаются ВЕГОЙ как текст. Полный текст поля может быть «разрезан» на отдельные самостоятельные **фразы** (на суждения, выражающие одну идею ответа). Каждое поле может содержать до 32767 символов (фактически это около 30 стандартных страниц текста).

«Символьное» поле позволяет при обработке получать словари разного типа, проводить классификацию,

совместный анализ с другими полями, выбирать для этого любые подмассивы.

Поле «Целое» предназначено для записи целых чисел (одного или нескольких). Это может быть один цифровой ответ на альтернативный вопрос (например, поле «Ваш возраст» — 37) или несколько ответов на многоальтернативный вопрос (например, поле «Ваши любимые занятия» — 1,7,12,18 — записываются номера отмеченных респондентом занятий).

Поле данного типа допускает шкалирование цифровой информации, его имеет смысл делать, если признак означает количество (как в случае возраста — можно разбить информацию по возрасту на любые интересные нас интервалы), а не номер (как в случае ответа про любимые занятия). Можно получать статистические распределения, выявлять все многомерные сочетания.

Разработаны два поля *смешанного* типа:

Поле «Смешанное 1» — его левая часть предназначена для записи целых чисел, правая — для записи текстов.

Поле «Смешанное 2» — левая часть поля предназначена для записи текстов, правая — для цифровой информации.

Каждая часть поля обрабатывается в рамках ресурсов, разработанных для поля соответствующего типа. Любая часть поля может быть пустой. Две части этих полей могут быть связаны друг с другом, а могут быть автономными. Например, в правой части были записаны названия вузов, а в левой введены их цифровые коды. Цифровые коды могут быть разработаны и введены позже.

Поле «Таблица» — в названии этого поля указывается заголовок таблицы (текст до двоеточия) и ее размер: количество столбцов и строк. При записи ответов в поле «Таблица» должно следовать столько полей с данными, сколько строк указано в заголовке этого поля; количество данных в каждом поле таблицы должно соответствовать количеству столбцов, указанных в заголовке.

Пример базы данных «Болонский процесс в России»

№	Название поля	Тип поля
1	Самореализация	Смеш. 1
2	Жизненные ценности	Смеш. 1
3	Почему люди получают ВО	Смеш. 1
4	Почему респондент получает ВО	Смеш. 1
5	Почему оказались в этом вузе, на этом факультете, специальности	Целое
6	Требования для поступления в этот вуз	Смеш. 1
7	Привлекательные моменты в получаемом образовании	Смеш. 1
8	5 наиболее важных дисциплин	Смеш. 1
9	5 наименее важных дисциплин	Смеш. 1
...

В данном случае при вводе данных опроса в систему ВЕГА было заведено 82 поля, для большинства полей выбран тип поля «смешанное 1». В последнее время мы стали предпочитать использовать в БД этот тип поля как для текстовых, так и для целочисленных признаков — тем самым мы как бы резервируем дополнительные поля, что позволяет затем вводить новые признаки и приписывать коды без каких бы то ни было затруднений, не переписывая базу данных.

Ввод первичных данных

Возможны два варианта получения БД:

а) Информационный массив анкета за анкетой (запись за записью) вводится непосредственно в Базу данных «внутри» системы ВЕГА по разработанной форме («структуре»).

б) В подготовленную структуру БД можно также импортировать «внешний» текстовый массив, набранный за рамками ВЕГИ в текстовом формате.

Текст при этом может быть структурирован — и тогда каждая единица текста (строка, абзац, фрагмент) должна соответствовать одному полю базы данных.

Текст может быть также импортирован в базу данных ВЕГИ в неструктурированном виде (и уже затем разностыся по отдельным содержательным/смысловым полям).

Пример: мы имели 50 значительных по объему биографических текстов, написанных матерями об их семейной ситуации в связи с наркозависимостью детей, набранные в формате Word. Все-таки это только тексты, но и значительная статистика — 50 историй и огромный объем текста, так что «носиться взад вперед» по огромному массиву в Ворде, не имея возможности «закрепить» найденные идеи — мало эффективное занятие. Поэтому эти биографические тексты мы закачали полностью как 50 записей в структуру ДИСКАНТА, предусмотрев массу запасных полей, и затем уже в рамках ДИСКАНТА выявляли в каждом полном тексте фрагменты текста по интересующим нас идеям и помещали/перемещали их в соответствующие поля. Например, нас интересовало, была ли в биографиях информация о ситуации у детей в школе — если находили высказывания по этой теме, то заношили их на определенное поле у соответствующих респондентов.

Еще один пример БД. Например, при анализе средствами ВЕГИ проекта «Концепции государственной политики профилактики наркомании и правонарушений, связанных с незаконным оборотом наркотических средств и психотропных веществ в РФ» была создана База Данных всего из одного поля и программа импортировала в нее полный текст Концепции как систему «записей» из отдельных абзацев текста. Далее мы не искали систематические рубрики, а только пытались решить такую проблему — «зерна отделить от плевел» — ОТДЕЛИТЬ «пустые», риторические, не несущие никакой содержательной или прагматической нагрузки куски текста ОТ существенных высказываний. При анализе характера предлагаемой обществу Антинаркотической концепции мы обошлись только получением пермутационного словаря — содержательных «зерен» в конечном итоге оказалось немного.

3.5. Работа со словарями

Работа со словарями представляет начальный ресурс для анализа текстовых полей. Иногда при отсутствии времени на систематическую классификацию суждений, мы ограничиваемся только анализом словарей.

В ВЕге можно получать разнообразные словари, при этом в качестве текстовых единиц могут выступать слова, ключевые слова, «фразы» и полные ответы на один вопрос.

Ключевые слова вводятся в текстовые высказывания заглавными буквами и позволяют быстрее ориентироваться в главной мысли суждения, ключевые слова имеют больший вес при поиске в ВЕГе оптимального класса для отдельных суждений,

Словари могут быть получены по одному текстовому полю, по нескольким или по всем текстовым полям исследования. При этом словари могут быть получены на весь массив анкет или на сформированный по одному или нескольким значениям признаков конкретный подмассив (например, студентки РГПУ, 4 курс).

В конечном итоге анализируемым текстовым массивом может быть: вся совокупность фраз (слов, ключевых слов) на один вопрос, вся совокупность фраз на систему «корреспондирующих» вопросов, все текстовые ответы одного респондента, тексты определенных подмассивов, текстовые ответы всего массива одного исследования, текстовые массивы «корреспондирующих» исследований.

Анализ текста при поддержке разного рода словарей, позволяет выявить фокусы авторского внимания, определить вариацию значимых высказываний — так в антинаркотической концепции до 80% суждений использовались как чисто декларативные формулы, не дающие содержательного приращения ни в плюс, ни в минус.

Поскольку далее мы собираемся объяснять инструменты ВЕГИ для классификации текстовых суждений на примерах одного конкретного исследования «Высшее образование и Болонского процесс в России», рассмотрим возможности словарей.

Большие возможности для анализа предоставляет, так называемый, **квазичастотный словарь**. Он подобен словарю по словам, однако частоты встречаемости всех близких однокоренных слов суммируются и в итоге приписываются первой встреченной форме каждого слова. Порядок слов ранжирован по частоте встречаемости, как это имеет место в обычном частотном словаре. Фрагмент такого словаря по второму полю приведен ниже:

*Квазичастотный словарь
«Почему Вы получаете высшее образование»*

64 работы	19 высшее	11 специальности
47 знания	18 диплом	10 будущем
47 получить	15 найти	10 интереса
38 хорошо	14 желание	10 карьера
36 образования	14 жизнь	10 профессию
32 хочу	14 иметь	10 статус
26 быть	14 людьми	9 город
26 возможных	14 общении	9 дает
24 родители	12 самореализация	9 повышения
22 новые	11 деньги	8 армии
21 интересная	11 общества	8 личная
20 престижно	11 развитие	8 учиться

Пояснение: представлены суждения 150 респондентов. Всего в словаре слов 1864, из них «уникальных» 550. В примере представлено 45 первых по частоте встречаемости слов, суммарная численность их частот составляет 682.

При **пермутационном выводе текста** дается полный перебор всех фраз из заданного текстового массива в алфавитном порядке по каждому слову, содержащемуся в соответствующем словаре. Например, фраза «учиться заставляют родители» попадет в таком варианте вывода в трех местах: сначала на букву «З»: «учиться ЗАСТАВЛЯЮТ родители»; затем на «Р»: «учиться заставляют РОДИТЕЛИ», последний раз на букву «У»: «УЧИТЬСЯ заставляют родители». В каждом случае рядом будут идти «стройные» ряды высказываний разных респондентов, в которых имеются вариации ВЫДЕЛЕННОГО слова.

Пермутационный словарь очень полезен для получения исследователем начальных представлений, о чем же толкуют респонденты в своих ответах, при поиске тематических классов для процедуры классификации массивов ответов на открытые вопросы. Очень эффективен анализ пермутационного словаря для изучения большого слитного текста, с тем чтобы понять, вокруг каких идей выстраивается этот текст и др. Ниже приведен фрагмент пермутационного вывода массива ответов.

Пример. Фрагмент **ПЕРМУТАЦИОННОГО** словаря,
 Ответы на вопрос
 «Почему Вы лично получаете Высшее образование»
 научиться АНАЛИЗИРОВАТЬ
 отсрочка от службы в АРМИИ
 некоторые косят от АРМИИ
 возможность продлить БЕЗЗАБОТНУЮ ЮНОСТЬ
 УСТРОИТЬСЯ на хорошо ОПЛАЧИВАЕМУЮ РАБОТУ
 получить ОПЫТ от преподавателей
 приобретение ОПЫТА в жизни
 желание ОСУЩЕСТВЛЯТЬ мечту
 страх перед ОТВЕТСТВЕННОСТЬЮ
 ПЕРСПЕКТИВА
 ПОВЗРОСЛЕТЬ
 ПОВЫШЕНИЕ своего статуса
 ПОВЫШЕНИЕ квалификации
 для ПОВЫШЕНИЯ своих знаний
 ПОЗНАКОМИТЬСЯ с новыми
 людьми
 ПОЗНАНИЕ в обществе
 это дает много новых и ПОЛЕЗНЫХ знакомств

Освоить ИНТЕРЕСНЫЕ и полезные если честно, не
 ПОСЛУЖИТЬ своей стране
 ПРЕДМЕТЫ
 ПРЕДСТАВЛЯЮ себя без высшего
 ПРЕСТИЖ
 ПРЕСТИЖНАЯ специальность
 ПРЕСТИЖНОСТЬ
 ЖЕЛАНИЕ УСТРОИТЬСЯ на ПРЕСТИЖНУЮ работу
 ПРИЗВАНИЕ — план Божий, судьба
 так ПРИНЯТО в семье
 стремление ПРОДОЛЖИТЬ образование
 Без этого не ПРОЖИТЬ
 обретение ПРОФЕССИИ

Полученные словари можно просматривать на экране, сохранять в виде файла, выводить на печать. При просмотре словаря на экране индуцируется словарная единица и все ее «адреса» (номера респондентов и полей, в которых встречается эта словарная единица). Можно сохранять словари словарных единиц с их частотами или с их адресами.

3.6 Средства классификации текстовых массивов (пример)

Представим идеологию, логику и технологию классификации на примере данных исследования о высшем образовании в России. В анкету включены два вопроса:

1. Почему люди в нашей стране учатся в вузах или получают высшее образование? Назовите 10 суждений/идей:
2. Почему Вы лично сейчас учитесь в вузе, получаете высшее образование? Назовите 10 причин/идей, связанных у Вас с этим выбором:

Данный опрос проводился в 2007-2008 г. в рамках исследования, посвященного внедряемому в отечественное образование Болонскому процессу. В опрос были включены студенты СПб госуниверситета культуры и искусств (место преподавания Г.И. Саганенко) и еще 4 престижных вузов Санкт-Петербурга (СПб ГУ ИТМО, РГПУ им. А.И. Герцена, СПб ГУ ФИНЭК, СПб ИНЖЭКОН)⁵.

Использование результатов предшествующих исследований позволило составить первую версию классификатора. Проведение классификации по указанным двум вопросам и изучение содержания и объема отдельных градаций, сокращение/перевод слабо наполненные градации в другие более или менее релевантные классы — дало в итоге следующий классификатор.

КЛАССИФИКАТОР: «Почему люди получают высшее образование»

1. **ДЕНЬГИ, МАТЕРИАЛЬНАЯ ОБЕСПЕЧЕННОСТЬ, ДОСТАТОК**
2. **РАБОТА:** 1 ИНТЕРЕСНАЯ
2 ПЕРСПЕКТИВНАЯ
3 РАЗНОЕ
3. **КАРЬЕРА**
4. **ПРОФЕССИЯ, СПЕЦИАЛЬНОСТЬ**
5. **ДИПЛОМ:** 1 ДОЛЖЕН, НАДО, НОРМА
2 ВСЕ УЧАТСЯ
3 СЕМЕЙНЫЕ ТРАДИЦИИ
4 КОРОЧКИ, ДИПЛОМ
6. **ВОЗМОЖНОСТИ, БУДУЩЕЕ**
7. **ЗНАНИЯ, УЧЕБА, ОБРАЗОВАНИЕ, НАУКА**
8. **САМОРАЗВИТИЕ, САМОРЕАЛИЗАЦИЯ, ДОБИТЬСЯ В ЖИЗНИ**
9. **СТАТУС, УВАЖЕНИЕ:** 1 ОБЩЕСТВЕННЫЙ
2 ЛИЧНОСТНЫЙ, В СВОИХ ГЛАЗАХ

⁵ Эти вопросы уже несколько раз были использованы в других исследованиях, в частности в дипломном проекте «Выбор высшего образования» (см. Книга 2. С. 24-39) 2003 года.

10. ПСИХОЛОГИЧЕСКАЯ УСТОЙЧИВОСТЬ
11. ЗНАКОМСТВА, ОБЩЕНИЕ С ЛЮДЬМИ, СВЯЗИ, ВСТРЕЧИ
12. АРМИЯ
13. УЧИТЬСЯ В БОЛЬШОМ ГОРОДЕ
14. СТУДЕНЧЕСКАЯ ЖИЗНЬ:
 - 1 СТУДЕНЧЕСКАЯ ЖИЗНЬ
 - 2 ПРОДЛИТЬ ДЕТСТВО, ПОБЫТЬ НА ШЕЕ РОДИТЕЛЕЙ
15. УХОД ОТ НЕГАТИВА, НЕ ДЕГРАДИРОВАТЬ, НЕ БЫТЬ СЕРОСТЬЮ
16. ЛИЧНЫЙ ВЫБОР
17. РАЗНОЕ

Как видим, для классификации ответов построена в основном линейная последовательность классов, только в четырех классах введен «второй этаж» — внутренние деления на «группы». Ниже представим частотное распределение совокупного текстового массива двух вопросов по позициям (классам/ группам) классификатора (группы в каждом классе свои и потому разные).

Таблица 1.
«ПОЧЕМУ ЛЮДИ ПОЛУЧАЮТ ВЫСШЕЕ ОБРАЗОВАНИЕ»
Распределение фраз по классам/группам классификатора

КЛАССЫ	ГРУППЫ					Сумма ответов	%%
	0	1	2	3	4		
1 ДИПЛОМ	5	38	44	72	39	198	16,3
2 РАБОТА	60	22	33	50	-	165	13,5
3 САМОРАЗВИТИЕ	152	-	-	-	-	152	12,5
4 ЗНАНИЯ	150	-	-	-	-	150	12,3
5 СТАТУС, УВАЖЕНИЕ	1	106	11	-	-	118	9,7
6 ДЕНЬГИ	92	-	-	-	-	92	7,6
7 ПРОФЕССИЯ	63	-	-	-	-	63	5,2
8 АРМИЯ	52	-	-	-	-	52	4,3
9 ЗНАКОМСТВА	48	-	-	-	-	48	3,9
10 КАРЬЕРА	35	-	-	-	-	35	2,9
11 СТУДЕНЧЕСКАЯ ЖИЗНЬ	-	20	15	-	-	35	2,9
12 РАЗНОЕ	32	-	-	-	-	32	2,6
13 ВОЗМОЖНОСТИ	28	-	-	-	-	28	2,3
14 УЧИТЬСЯ В ГОРОДЕ СПБ	18	-	-	-	-	18	1,5
15 УХОД ОТ НЕГАТИВА	13	-	-	-	-	13	1,1
16 ЛИЧНЫЙ ВЫБОР	11	-	-	-	-	11	0,9
17 ПСИХОЛОГИЧЕСКАЯ УСТОЙЧИВОСТЬ	8	-	-	-	-	8	0,7
18 Разное	32	-	-	-	-	32	2,6
ИТОГО						1218	100

Примечание. Ответы 150 респондентов по двум полям вместе — всего фраз 1218.

Таблица 2. «ПОЧЕМУ ЛЮДИ ПОЛУЧАЮТ ВЫСШЕЕ ОБРАЗОВАНИЕ»
(Сравнение структуры ответов «о других» и «о себе»)

	Почему люди получают высшее образование (Л)			Почему респондент получает высшее образование (Р)			Разница Л – Р	ВСЕГО	
	ответов	% ответов	% людей	ответов	% ответов	% людей		% людей	ответов
ДИПЛОМ	120	17,0	82	78	13,4	53	29	198	68
РАБОТА	91	12,9	62	74	12,8	51	11	165	57
САМОРАЗВИТИЕ	77	10,9	53	75	12,9	51	2	152	52
ЗНАНИЯ	69	9,8	47	81	14,0	55	-8	150	51
СТАТУС, УВАЖЕНИЕ	78	11,1	53	40	6,9	27	26	118	40
ДЕНЬГИ	52	7,4	36	40	6,9	27	9	92	32
РОДИТЕЛИ	39	5,5	26	28	4,8	18	8	67	23
ПРОФЕССИЯ	25	3,5	17	38	6,6	26	-9	63	22
АРМИЯ	43	6,1	29	9	1,6	6	23	52	18
ЗНАКОМСТВА	18	2,6	12	30	5,2	21	-9	48	16
КАРЬЕРА	20	2,8	14	15	2,6	10	4	35	12
СТУДЕНЧЕСКАЯ ЖИЗНЬ	23	3,3	16	12	2,1	8	8	35	12
ВОЗМОЖНОСТИ	12	1,7	8	16	2,8	11	-3	28	10
УЧИТЬСЯ В ГОРОДЕ	5	0,7	3	13	2,2	9	-6	18	6
УХОД ОТ НЕГАТИВА	4	0,6	3	9	1,6	6	-3	13	4
ЛИЧНЫЙ ВЫБОР	3	0,4	2	8	1,4	5	-3	11	4
ПСИХ. УСТОЙЧИВОСТЬ	6	0,9	4	2	0,3	1	3	8	3
РАЗНОЕ	20	2,8	14	12	2,1	8	6	32	11
Итого	705	100		580	100			1285	

Пояснение: Ранжирование по наибольшей частоте ответов — предпоследний столбец.

Мы имели возможность сравнения с результатами 2003 года. Обнаружилось, что произошли следующие существенные изменения: на первое место вышел «диплом» как таковой и прописался фактически у каждого респондента. Существенно стало больше идей, связанных с работой. Значительные изменения произошли по фактору «деньги, материальная обеспеченность, зарплата» — раньше это тема составляла только около 3% упоминаний, сейчас, как мы видим — 7,6 %.

Покажем средства анализа и сравнения структур нескольких текстовых массивов (таблица 2).

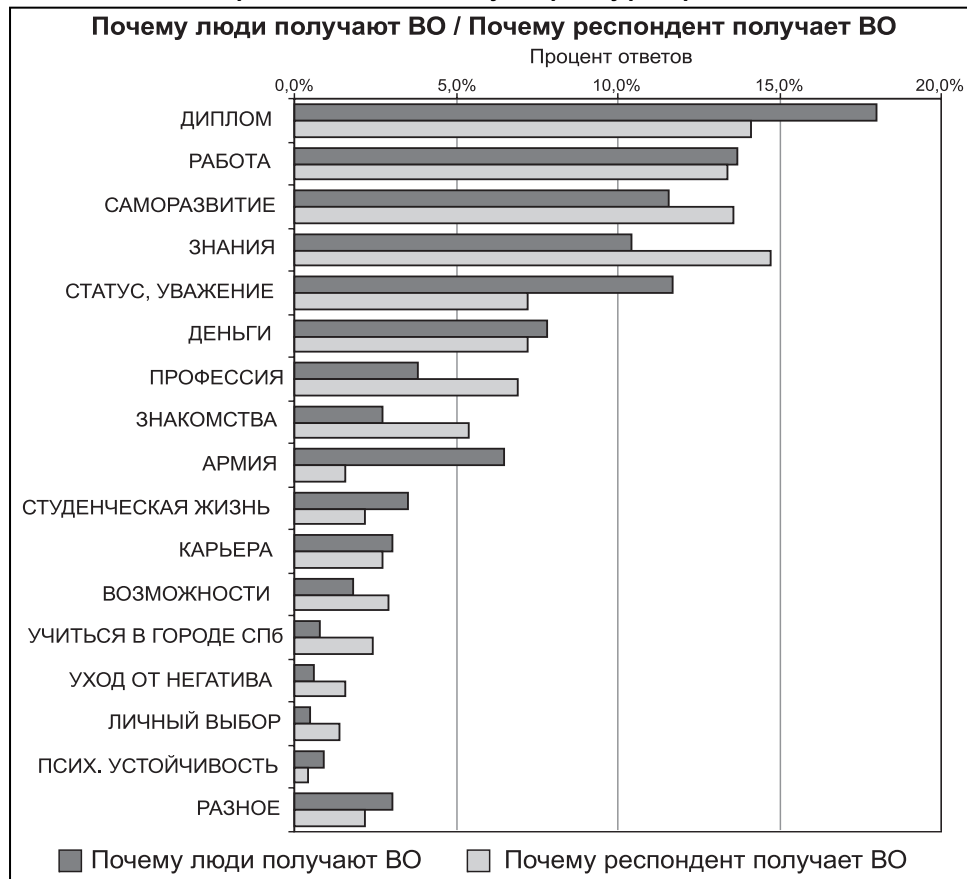
Если посмотреть на столбец «РАЗНИЦА» — разницы со знаком плюс означают, что респонденты чаще в ответах называли эти вещи применительно к «другим», чем к «себе». И наоборот — разница с минусом означает, что респонденты чаще записывали такие ценности «себе» и реже «другим». Так учатся ради ДИПЛОМА (1-я строка) 82% — другие и только 53% — сами респонденты, так что 82% — 53% дает 29%. И так резоны получать высшее образование у других, значительно чаще, чем у респондентов такие: другие чаще учатся ради дипломов, ради приобретения статуса, чтобы избежать службы в армии. А также относительно чаще им приписывают получение высшего ради получения работы, денег, по настоянию родителей. Респондент же чаще, чем другие люди, учится ради получения знаний, ради получения профессии.

Гистограмма (рис. 1) дает возможность наглядно увидеть значительные отличия в структурах текстовых полей и уточнить, каковы они. Оказалось, что «другие» в представлениях респондентов чаще отправляются в вуз под давлением родителей, чтобы избежать армии, ради получения диплома, ради статуса. Сам же респондент о себе чаще говорит — что он шел в вуз ради знаний, саморазвития, новых знакомств.

Система ВЕГА имеет эффективные возможности анализа больших концептуальных текстов. Уже упоминались биографические тексты, Антинаркотическая концепция для страны.

Система ВЕГА имеет средства статистической обработки цифровых признаков, среди таковых: возможности многократного шкалирования признаков, получения одномерных, двумерных распределений, многомерных сочетаний для многоальтернативных признаков.

Рис. 1. Сравнение мотивации на получение высшего образования «у других» и «у респондента». Ранжирование мотивов по суммарному распределению.



Заключение

Рассмотрена система макро-составляющих, которые сопровождают реализацию любого значимого социологического исследования. В нашей концепции они считаются определенными «субъектами», поскольку влияют на замысел, масштабы, процесс и результаты исследования. Отмечен ряд новых методологических коллизий, которые не рассматриваются и даже не называются социологами — многосоставной контекст исследования, «корреспондирующие» объекты и исследования, гуманистические интенции эмпирического исследования, коммуникационная и познавательная специфика «стационарных» аудиторий и др., специфика исследований на базе открытых вопросов.

Показана часть основных приемов, которые разработаны в системе ВЕГА, чтобы добиваться продуктивного и выразительного анализа многосоставных исследований, построенных на открытых вопросах и тем самым на множестве текстовых массивов. В рамках данной публикации была возможность только предъявить некоторые модули анализа текстовых массивов, не останавливаясь на процессе поиска решений, не имея возможности комментировать и интерпретировать полученные результаты.

Литература

1. Алексеев А.Н. Драматическая социология и социологическая ауторефлексия. Т. 1,2,3,4. СПб.: Норма, 2003, 2005
2. Батыгин Г.С. Обоснование научного вывода в прикладной социологии. М.: Наука, 1986
3. Готлиб А. С. Качественное социологическое исследование: познавательные и экзистенциальные горизонты. Самара. Изд-во Универс-групп. 2004.
4. Ильин В.И. Драматургия качественного полевого исследования. СПб: Социополис, 2006. 256 с.
5. Каневский Е.А., Саганенко Г.И., Гайдукова Л.М., Клименко Е.Н. Система анализа текстов // Социология: 4М, 1997. №9. С. 198-216.
6. Каневский Е.А., Саганенко Г.И. Концептуальное обоснование компьютерного анализа массивов с текстами // Социология: 4М, 1997. №9. С. 65–81.
7. Саганенко Г.И. Лицом к лицу с наркоманией: перемены идут, перемены возможны. (Социология для всех). СПб: ГНУ «ИОВ РАО». 2005. 326 с.
8. Саганенко Г.И. Многообразие возможностей социологического изучения сферы образования. Книга 1. Прикладные технологии в социологическом изучении сферы образования: Обоснование и технологии социологических исследований, базированных на текстах.
9. СПб: ГНУ «ИОВ РАО». 2004. 116 с.
10. Саганенко Г.И. (науч. ред.). Многообразие возможностей социологического изучения сферы образования. Книга 2. Прикладные технологии в социологическом изучении сферы образования / Под ред. Г.И. Саганенко. СПб: ГНУ «ИОВ РАО». 2005. 163 с.
11. Саганенко Г.И. Надежность результатов социологического исследования. Л.: Наука, 1983.
12. Саганенко Г.И. Социальное время как фактор эмпирической методологии в социологии // Журнал социологии и социальной антропологии. 2002. № 1. С. 129-136
13. Саганенко Г.И., Воронцова О.А. Представления молодежи о стране эпохи СССР // Социология — 4М, 2008, № 27. С. 158-188.
14. Ядов В.А. Социологическое исследование: методология, программа, методы. Самара, Изд-во Самарского ун-та, 1995. 331 с.