

## Контент-анализ — методические основания исследования в электронных СМИ\*

В. Г. Погорецкий

Контент-аналитическое исследование, избирающее в качестве документа-объекта материалы прессы, измеряет и оценивает представленное в публикациях социальные явления-объекты, отражение (характер отражения) их в исследуемых текстах и методы, используемые авторами текстов для анализа явлений и объектов. Относительно исследовательских подходов следует сказать, что в настоящее время типовые методики социологического контент-анализа отсутствуют. Общие принципы количественного оценивания содержания текстов реализуется различными способами. Ввиду многочисленных проблем формализации методические задачи контент-анализа решаются применительно к конкретным эмпирическим исследованиям или в теоретическом плане сосредотачиваются на сформулированных проблемах, последовательно развивая отдельные аспекты формализации знаковых структур, именуемых текстами.

В прикладном плане реализация общих теоретических принципов контент-анализа даже для одних и тех же задач может быть осуществлена различными способами — эквивалентными или дающими отдельные преимущества в эффективном решении прикладных задач. В настоящее время идет накопление форм обработки текстов, и чтобы их суммировать, осознать их общность, объективно взвесить достоинства и недостатки разновидностей инструментария во всей совокупности, необходимо сочетать конкретный контент-анализ с обсуждением используемых методических подходов и техники измерения, рассматривая в таком случае объект и сам процесс прикладного исследования как **тест** для инструментария и методический пилотажный проект для собственно техники и технологии измерения.

Поскольку пластичность текстов сообщений безгранична, что, в частности, дает повод для объяснения их природы с позиций взаимоисключающих теорий дескриптивного и порождающего типа, а тип методик контент-анализа может варьировать от простых, очевидных способов непосредственного измерения до сложного семиотического, интенционального анализа, обсуждение методических аспектов контент-анализа в процессе представления прикладного эмпирического результата вполне оправдано.

---

\* Работа выполнена при финансовой поддержке РГНФ (проект № 02-03-18007а).

Обсуждение метода в процессе получения содержательного результата имеет для контент-анализа принципиальное значение в силу его особого положения в совокупности ресурсов социологического изучения реальности. Рассматривая его как одну из форм анализа документов, мы поддерживаем позицию, согласно которой контент-анализ в качестве метода имеет универсальное значение как средство анализа дискурса. Поскольку все мышление есть мышление в понятиях, то есть дискурсивно, значит, соответственно, исследовательский аппарат дискурса имеет общее значение и может применяться за пределами той сферы, для нужд которой он был впервые разработан.

Под общим понятием «дискурс» подразумевается собственно дискурс как обсуждение, выдвижение альтернатив, сопоставление альтернатив, взвешивание, предпочтение, выбор и идентификация образа объекта в соответствии с дискурсивно определенным понятием, что есть логика понятийной квалификации объекта. В то же время общее понятие «дискурса», наряду с понятием собственно «дискурса» включает и понятие «нарратива» [1] — повествования, рассказа, изложения последовательности, смены событий, логики развития ситуации от истока к цели, к результату, логики истории реализации события. Эти две логики «дискурса» и «нарратива» могут быть разделены только в абстракции, но в реальном мышлении существуют нераздельно. И как в профессиональном социологическом анализе, так и в общей интеллектуальной практике вне зависимости от характера предметности интеллектуальной деятельности присутствуют одни и те же формы дискурсивного (в общем смысле понятия «дискурс»), то есть понятийного сознания, социологическая версия контент-анализа приобретает общеинтеллектуальное значение. Приоритет же социологического контент-анализа в качестве метода с компонентами общеинтеллектуальной значимости состоит в том, что, будучи ориентированным на эмпирические прикладные задачи, он, преимущественно, в первую очередь, опирается на формы дискурса (в узком смысле слова), то есть понятийной идентификации, так как они непосредственно содержатся в текстах сообщений, служащих для него первичными документами, исходным материалом анализа, и в ходе исследования и измерения этих первичных понятийных форм сам метод наиболее последовательно воспроизводится и, как это бывает при многократном интеллектуальном воспроизводстве, устойчиво фиксируется, внутренне дифференцируется, обретает стройную последовательную непротиворечивую структуру, в то же время создавая оппозиции сложившимся формам идентификации и квалификации понятий. В этом опосредована его общеинтеллектуальная, а не только собственно социологическая значимость.

Но такая общедискурсивная значимость метода приобретает смысл только тогда, когда он может быть усвоен в процессе восприятия социологического предметного результата за пределом собственно професси-

ональной социологической сферы. В нашем случае, случае конкретного социологического анализа феномена коррупции в обществе, внимание со стороны общественности к предмету исследования носит самый заинтересованный характер, поскольку явление коррупции затрагивает интересы каждого. Вместе с результатами социологического анализа усваивается и его дискурсивная логика, в которой, в отличие от хроники текущих событий, представленной непосредственно в материалах публикаций, присутствуют характеристики логических форм обобщения, дистинктивности и вывода, для которых череда конкретных фактов, повествующих о сути социального явления (у нас коррупции) только череда иллюстраций общих логических форм дискурсивного (понятийного) сознания. Так, усваивая предметность социальных событий, субъект-реципиент сообщения усваивает и логическую ментальную фактуру анализа, если фокус его внимания попеременно акцентируется то на предметности, то на понятийном строе анализируемых событий. Но здесь мы еще раз должны подчеркнуть, что такой контакт восприятия общественностью формально-логических структур дискурсивного вывода конкретного социального явления может состояться, если метод опирается на непосредственность дискурсивной логики понятийной идентификации реальности, что, конечно, не должно ослаблять внимания специалистов к фундаментальным проблемам, формулируемых в языкознании, семантике, социолингвистике, модальных и интенциональных логиках. То, что определяется как акт дискурсивной идентификации реальности есть логический акт суждения о тождестве образа реальности и понятия. Если образ реальности выступает в суждении в качестве субъекта, то факт принадлежности ему свойства, представляемого предикатом, констатирует наличие тождества в случае общеутвердительного или частноутвердительного суждения, или, напротив, констатирует отсутствие тождества в случае общеотрицательного или частноотрицательного суждения. Тот же принцип тождества распространяется на конъюнктивные и дизъюнктивные суждения.

Относительно методик контент-анализа мы знаем, что значимость, вес определенной тематики в общем объеме текста может определяться просто по относительному объему площади, посвященной данной проблематике в общем объеме текста, или по частоте употребления слов, характеризующих данную тематику в общем объеме текста. Но эта, так сказать, приблизительные, брутальные оценки. Оценки же более точные, непосредственно соотносимые с содержанием текста, становятся возможными тогда, когда объект исследования, соотносимый с понятием субъекта суждения, оценивается по совокупности тех определений (свойств), которые приписываются ему в последовательном ряду суждений в качестве предикатов этих суждений. Очевидно, что в дискурсивной структуре сообщения тематический объект в качестве субъекта последовательного ряда суждений оказывается устойчивым фрагментом, тогда как приписываемые

объекту характеристики, фиксируемые предикатом суждений, представляют фрагменты изменчивые. Но каковы бы они ни были, они фиксируют факт тождества образа объекта подлинному реальному объекту на основании свойств, приписываемых субъекту, повторяющемуся в последовательном ряду суждений, совокупности предикатов, содержащихся в этих суждениях. И если в конкретном контент-аналитическом исследовании мы имеем ряд текстов, написанных по одному и тому же поводу, посвященных одной тематике, то объект у них один и тот же, и он фиксируется в качестве понятия-субъекта суждения некоторого содержательного дискурса, но тождество образа объекта (субъекта суждения) объекту реальности, который он представляет, в суждении констатируется на основании совокупности предикатов всей совокупности суждений отдельного тезиса. Самостоятельные тексты, независимо представляющие одну и ту же тематику, описывающие один и тот же объект или одну и ту же проблему объекта, констатируют факт тождества на основании возможных и достаточных характеристик объекта (субъекта суждения), характеристик-предикатов. Вот эти-то совокупности характеристик, фиксирующие тождество объекта и его понятийного образа, и являются главным предметом регистрации и измерения контент-анализа в конкретном эмпирическом исследовании. У разных авторов сообщений (отдельных текстов) эти характеристики (предикаты) будут различны. Они-то при регистрации и выявят различные содержательные и оценочные позиции авторов текстов, что, собственно, и являются одной из целей контент-анализа.

При проведении эмпирических контент-аналитических измерений, также как при разработке программ для автоматизированной обработки текстов исследователи сталкиваются с целым рядом технических проблем. Так, например, при первичной регистрации категорий (понятий-объектов исследования) мы обнаруживаем, что они многозначны, могут быть выражены посредством синонимов, словосочетаний, эквивалентных исходному понятию-категории, использовать вместо общего исходного понятия его еще более общие определения или, наоборот, частные предметные определения общего понятия, после первоначально полного обозначения понятия, которое может состоять из нескольких слов, может употребляться его сокращенное, контекстуально идентифицируемое обозначение или аббревиатура. Стилистика текста часто требует использования местоимений, вместо фиксируемых существительных, субъект суждения не обязательно может оказываться подлежащим предложения, предложение может состоять из нескольких суждений. Все это создает технические трудности реализации формальных контент-аналитических принципов, и на преодоление их необходимы значительные усилия. Наличие в текстах аннотаций, пояснительных подписей под фотографиями и тому подобное искажают статистическую картину регистрируемых единиц, также как избыточное содержание текстов, исследуемых в рамках конкретной исследовательской

задачи. Все это делает методики контент-анализа трудоемкими, недостаточно формализуемыми, находящимися в значительной зависимости от человеческого фактора и, следовательно, незащищенными от ошибок. Тем не менее, методики контент-анализа, базирующиеся на трактовке дискурса сообщений как совокупности суждений о тождестве образа объекта (субъект суждения) на основании совокупности предикатов ряда суждений, относимых к одному и тому же объекту реальности представляют собой один из эффективных инструментов метода. Именно он реализует в себе логику непосредственной информационной коммуникации субъектов социального действия. Ни сложный дедуктивный вывод, приводящий к результату через серию сложных абстрагирующих акций, ни последовательная и исчерпывающая индукция, позволяющая придти к неопровержимому заключению, ни анализ противоречий в ходе рассуждений оппонента, но непосредственное утверждение принадлежности предиката субъекту суждения, выдвижение ряда суждений, в которых одному и тому же понятию приписывается последовательный ряд характеристик, и оценка этой совокупности характеристик, служат основанием для решения о признании соответствия или несоответствия исходного утверждения фактической, реальной действительности.

Воспроизводя в методической процедуре контент-анализа элементарный акт естественной обыденной логики информационной коммуникации, мы получаем возможность и количественно оценить и сопоставить суждения, высказанные одним субъектом сообщения по одному поводу, по разным поводам, связанным с одним объектом или проблемой, или осуществить ту же последовательность действий в отношении нескольких субъектов сообщений без ограничения их числа. При этом могут быть зафиксированы интенсивность утверждений (степень категоричности придаваемой суждению значительности); экстенсивность (количество доводов в пользу исходного тезиса); наличие альтернативных квалификаций исходного тезиса в виде прямого отрицания или выдвижения несовместимых с ранее выдвинутыми содержательными доводами, что и служит материалом для последующей аналитической деятельности исследователя.

Реализация данного подхода в методике контент-анализа имеет еще один аспект. Поскольку регистрируемые единицы непосредственно совпадают с элементарными актами массового рационального сознания, то техника анализа текста социологической методики может быть перенесена на процедуры анализа сообщений массовой практики интеллектуальной деятельности в системах массовых коммуникаций. Внедрение технологий анализа текстов, основанных на данных принципах для воспроизведения и усвоения содержания сообщений, формирует еще один вектор социальной информатики, который наряду с известными обращается к обработке качественной информации в целях ее эффективного усвоения, ее лексической удобочитаемости с точки зрения рефлексии в рамках формальной

логической модели, это во-первых, и, во вторых, ее лексической удобочитаемости посредством формального аппарата контент-аналитической обработки текста, возможны даже две параллельных версии текста — для субъективного человеческого прочтения и машинного прочтения для определенных целей, в таком же ключе может быть составлен, возможно, опять-таки посредством контент-аналитических программ реферат текста.

Используемая нами программа обработки текстов автоматически составляет словарь всех слов, использованных в исследуемом тексте, причем, если слово относится к изменяемым частям речи, фиксируются все словоформы данного слова. Падеж, к которому относится данная словоформа, можно определить по окончанию, что представляет дополнительную информацию для анализа, если, конечно, словоформы в тексте, употребленные в значении различных падежей, не совпадают по написанию. Далее будем использовать примеры из текстов проведенного нами исследования о коррупции. Так в родительном и винительном падежах слово «чиновник» записывается одинаково «чиновника». Слова, встречающиеся в тексте в кавычках, как, например, характерные для коррупционной проблематики «оборотни», «крыша», фиксируются отдельно, если то же слово записано без кавычек, оно фиксируется независимо от того, употреблялось ли это слово до того в кавычках или нет. Так первоначально «оборотни» пишется в кавычках, но затем, прижившись в тематическом лексиконе, войдя в обиход словоупотребления, воспроизводится без кавычек, слова, написанные через дефис, трактуются как единое слово. Так «вымогатель-чиновник» будет записано как одно слово. Числа, не разделенные пробелом, трактуются как одно слово. Это важно, поскольку введенные в отдельные места текста числа могут нести некоторую извне заданную информацию об этом тексте. Например, при обработке в единый текст объединяются статьи из разных газет, с целью фиксирования источника отдельной публикации, перед ее началом мы вводим пятизначное число соответствующее индексу по каталогу Роспечати, что будет информативным свидетельством источника и вне рамок отдельного, локально проводимого нами исследования публикаций о коррупции, а, скажем, при сохранении материалов публикаций в архиве и использовании их в других контент-аналитических исследованиях. Те же цифровые индексы, проставляемые перед началом каждой из публикаций, вошедших в объединенный массив, выполняют и еще одну функцию — фиксируют начало и конец отдельного текста.

Дата публикации является еще одним необходимым идентификатором текста. В публикуемых текстах статей дата иногда отсутствует, иногда присутствует и может быть записана множеством различных способов, в начала или в конце текста, поэтому используется стандартный формат записи даты год-месяц-число (гггг-мм-чч). Программой такая запись интерпретируется как одно слово. Фиксирование даты в формате

используемой операционной системы предоставляет дополнительные возможности поиска, сортировки, формирования выборок из общего массива текстов. Исчерпывающей информацией о происхождении конкретного отдельного текста могло бы быть еще и указание веб-адреса текста.

В сформированном словаре слова размещаются в алфавитном порядке. При этом использованные в тексте словоформы одного слова оказываются в непосредственной близости, в относительной близости оказываются и те же слова, но с различающейся основой (при факторе беглых гласных, или словах с разной основой, но относящихся к одному корневому гнезду). Это значительное обстоятельство. Например, «коррупция» и «коррупцированность», поскольку и «коррупция» и «коррупцированность» обозначают одну и ту же категорию в контент-аналитическом исследовании, несмотря на то, что непосредственно воспринимаемые как носители оттенков смысла одного понятия, в некотором контексте могут обозначать достаточно различные понятия, или, наоборот, употребляться как параномические эквиваленты. Однако, в любом случае оба слова относятся к одной категории, и неучет одного из них становится источником искажения измерения параметров текста. Обычно круг исследуемых категорий определяется на предварительной стадии контент-аналитического исследования, в процессе непосредственного знакомства с первичным материалом. Но поскольку слово «коррупцированность» заведомо реже употребляется, чем «коррупция» оно может ускользнуть от внимания исследователя на этой предварительной стадии. В этом случае контроль слов, фиксирующих исследуемые категории по словарю, восполняет этот пробел, поскольку в словаре оба слова размещаются в относительной близости одно от другого. Таким образом, словарь выполняет определенную сервисную функцию контроля и пополнения терминов, обозначающих исследуемые категории или их отдельные значения. Для осуществления подобных функций в полном объеме необходимо просмотреть весь словарь, что, конечно, является рутинным и достаточно трудоемким делом.

В то же время просмотр словаря является исходной позицией для дальнейшей обработки текста. Выше мы обращали внимание на то, что перечень терминов, характеризующих категории, формируется на предварительной стадии в ходе осознания проблемы, формулирования целей и концепции исследования, выдвижения гипотез, и, как мы говорим, предварительного ознакомления с текстом. Но современный контент-анализ, опирающийся на машинную обработку больших массивов текстов, не предполагает полного, а тем более многократного их прочтения. В этом случае всегда остается вероятность, что часть значимых для категориального анализа терминов останется неопознанной на предварительной стадии, не будет включена в анализ, что, естественно, скажется на качестве результата. Чтобы избежать этого, можно выделить активно используемые в тексте термины по их частотным характеристикам. Этот путь, однако,

не свободен от проблем. Наибольшую частоту употребления будут иметь служебные слова, слова функциональные в чисто грамматическом плане. В практике контент-анализа используются различные способы для преодоления подобных трудностей.

Вторая фаза обработки текста состоит в вводе тех терминов, которые определены в качестве представляющих исследуемые категории и выделения тех предложений, в состав которых эти термины включены. Как мы уже говорили, категория может быть выражена одним словом, или несколькими словами в тексте, или одной категории могут соответствовать несколько слов с разным набором значений. Так в достаточно простом варианте в качестве субъекта коррупционного действия может рассматриваться понятие «чиновник». В этом случае слово-термин совпадает с категорией, категория непосредственно представлена в тексте, ее присутствие в тексте непосредственно индицируется данным словом, и поиск категории в тексте происходит непосредственно. Правда, у термина, эквивалентного категории, может быть синоним, например, «чинуша». В зависимости от задачи мы его можем рассматривать как тождественный категории или размещаем оба слова на шкале оценочной характеристики, и тогда «чинуша» будет зафиксирован как вариант с большей интенсивностью отрицательной эмоциональной выразительности. В общем плане, категория в тексте может быть представлена термином или выражением как общеутвердительным, частноутвердительным, общеотрицательным или частноотрицательным понятием.

Если категория «субъект коррупционного действия» предполагает ряд конкретных типов субъектов, то она рассматривается как класс понятий, где «глава администрации города» — один тип; «заместитель главы администрации города» — другой тип; и у них могут быть эквиваленты: «глава администрации города» — «мэр»; «заместитель главы администрации» — «замглавы администрации». Поиск категорий должен осуществляться по всем возможным вариантам выражения, в которых категория может присутствовать в тексте. Очевидно, что здесь появляются технические проблемы автоматической регистрации категорий в тексте. Просто, когда категория выражена одним словом, для регистрации же категории в форме совокупности слов программа должна быть существенно усложнена. Но каждый уровень решения технических проблем открывает новый уровень нерешенных, проблемы количественного анализа текста бесконечны. Поэтому в прагматическом плане реализации эмпирического исследования целесообразно следовать правилу: то, что на данном уровне реализации технической программы автоматизированного анализа текста не может быть решено, решается на уровне неформализованного редактирования и неформализованного анализа технического результата.

Так, если категория квалификации «инкриминируемого деяния» может фиксироваться в тексте в форме «превышение служебных полномочий»



или «злоупотребление служебными полномочиями» — очевидно, что это два разных типа одного класса или две конкретные версии одной общей категории анализа. Используемая же программа может осуществлять поиск только по одному слову, но не по сочетанию слов. Тогда мы должны выбрать, по какому из трех слов осуществлять поиск категории, какое бы мы не выбрали — результаты будут разными. Это слово мы обозначили как «индикатор категории». Получив в результате обработки список предложений, содержащих, например, слова «служебных», мы можем сначала удалить все, не относящиеся к делу, те, в которых слово «служебных» сочетается со словом «расследование», а затем разделить оба типа одного класса путем непосредственного редактирования. Но есть и другой путь. Найденные в результате фразы подвергнуть вторичному поиску, который будет искать слова «превышение» и «злоупотребление».

Поскольку в словаре фиксируются все словоформы, в которых слово присутствует в тексте, поиск может осуществляться как по списку словоформ, так и по основе слова. В первом случае в полученном в результате, списке предложений порядок их следования будет соответствовать порядку словоформ в списке поиска, во втором порядке следования предложений в исходном тексте. Если процедура поиска запускается для нескольких слов-индикаторов категорий одновременно, то на выходе будет аналогичная картина как в случае поиска по основе слова.

На третьей стадии окончательно отселектированные предложения, содержащие исследуемые категории, анализируются и редактируются, приближаются к форме логического суждения, удаляются вторичные и избыточные характеристики, если необходимо, изменяется порядок слов, естественно, при сохранении смысла редактируемой фразы. Наличие внутреннего редактора позволяет это сделать. Оценка категорий в контексте приписываемых им значений в отдельных фразах на основании различения субъекта и предиката фразы является заключительным этапом конкретного контент-аналитического исследования.

Наличие необходимого интерфейса позволяет осуществлять нужные операции неподготовленному пользователю, но поскольку обработка текстов происходит с использованием базы данных, то ряд процедур может осуществляться при использовании ресурсов СУДБ, что расширяет возможности опытного пользователя.

Дублирование способов подготовки результатов имеет и то преимущество, что, прибегая к использованию ресурсов СУДБ в процессе проведения контент-анализа, пользователь по мере накопления опыта отчетливее сознает потребность в тех или иных процедурах и операциях, необходимых для развития и совершенствования программ. Когда такая потребность в модернизации инструментария вызревает, исследователь и разработчик метода чувствует себя психологически более готовым к контактам с профессиональными программистами, которыми и будет такая модернизация

проведена. В результате приобретенных навыков социолог способен более отчетливо, предметно операционально формулировать задачи развития техники контент-аналитической обработки текста, в чем едва ли не главная составляющая успеха предстоящего предприятия.

Объект контент-аналитического исследования в электронных СМИ как бы непосредственно доступен: есть интернет-адреса газет; есть регулярный поток публикаций; есть архивы отдельных изданий; есть поисковые системы, непосредственно встроенные в информационные сети издателей. Все это облегчает задачу доступа к непосредственному предмету социологического исследования и анализа. В то же время задача выделения генеральной совокупности объектов исследования затруднена, поскольку лавина информации по первоначально обозначенной проблеме, ориентированной на отчетливо обозначаемый эмпирический объект, оказывается связанной множеством ассоциаций, которые оказываются существенными и воспринимаются как принципиально важные для объективного и исчерпывающего результата, с более широкой информационной реальностью, чем это представляется при первоначальном формулировании проблемы и вытекающей отсюда формулировкой предполагаемой локализации объекта в информационном пространстве. Помимо этого информационная реальность, это вторичная реальность, реальность отражения, констатации, описания и рефлексии, и как таковая она предполагает множество типов и жанров отображения, фактически представляющих разные жанры информационных сообщений, требующих особенных подходов и способов регистрации.

Поскольку сплошное обследование невозможно, уже потому что мы имеем дело с некоторым подобием «галактики Гуттенберга», границы которой как минимум неотчетливы, а также в силу обычных для эмпирического социологического объекта причин, совершенно необходимо сформировать обеспечивающую репрезентативность исследования выборку. Возникают и проблемы временного ограничения периода наблюдения. Очевидно, что они диктуются как содержательными аспектами проблемы, так и прагматичными целями прикладного эмпирического исследования. В качестве примера приведем решения данных вопросов в нашем исследовании материалов электронной прессы о коррупции.

Для контент-анализа текстов о коррупции нами были выбраны публикации трех центральных газет, находящихся, как нам представлялось, в разных секторах спектра выражения общественного мнения. Одной из них являлась «Российская газета» — официальный орган правительства России, самое массовое независимое корпоративное издание «Комсомольская правда», ориентированная на массового читателя, которое, в частности, служило трибуной или кафедрой для обращения к населению главы государства, и «Независимая газета» — издание, на период обследования главный собственник которого, а, следовательно, и лицо,

имеющее возможность влиять на редакционную политику издания, публично заявлял, что находится в оппозиции к проводимой в настоящее время в стране государственной политике. Таким образом, если в освещении данной проблематики в этих изданиях имеются различия, обусловленные политической позицией и политическими мотивами, они должны некоторым образом проявиться и в отношении освещения в публикациях проблемы коррупции.

Период, к которому относятся рассматриваемые публикации о коррупции, ограничивается временем с 1 января 2004 г. по 30 сентября 2004 г. Интервал, за который взяты публикации, определяется не произвольно выбранными сроками, а связан с временными рамками, в некотором роде, однородного развития общественно-политического процесса в стране, ограниченного, с одной стороны, выборами в Государственную Думу, с другой, масштабными террористическими актами, показавшими, что эти террористические акты носят характер системной стратегии и требуют мобилизации всех ресурсов страны для борьбы с ними.

Противостояние продажности в структурах государственного и общественного управления помимо прямой необходимости сделать социально-политическое управление эффективным и действительно базирующимся на гуманитарных ценностях гражданского общества многократно являлось средством заявить о своей позиции различных политических сил, борющихся за представительство в органах государственной власти. Поэтому особое значение накануне выборов в Государственную Думу имело заявление о раскрытии преступлений «оборотней в погонах» — широко известного дела о взятках и вымогательствах группы офицеров МВД. Скептический взгляд мог усмотреть в сопровождающих уголовное расследование заявлениях политических лидеров пиаровскую акцию, протекающую по канонам выборных технологий, которая окончится сразу после подсчета голосов. Но известные политические заявления оказались действительно началом планомерной политики борьбы с коррумпированностью во всех эшелонах общественного управления, в том числе и в высших эшелонах государственной власти. Метафорическое выражение «оборотни» получило распространение в качестве клише, когда автор хотел привлечь особое внимание к факту коррупции в рядах милиции. Поэтому последовательное освещение коррупционных явлений в прессе начинается после окончания праздников Нового Года. События 1 сентября 2004 г. в Беслане явились потрясением для всего общества и временно переместили фокус общественного внимания с коррупции на вопросы безопасности. Очевидно, что тема коррупции поднималась и до, и после указанного нами периода, и в этом смысле определенные нами границы в известной степени условны, хотя, по нашему мнению, фиксируют некоторый однородный период в состоянии общественного сознания. Публикации о коррупции отбирались нами в Интернет-изданиях указанных

газет, отбирались по названиям, сопровождающим названия аннотаций и, если это было необходимо, просматривался текст публикации. Первоначально было отобрано 208 публикаций. Тексты, с точки зрения сопоставления их в рамках единой процедуры, как и следовало ожидать, оказались неоднородны. Часть из них представляла собой выраженный тип «Анализа случая». Мы берем в кавычки выражение «анализ случая», поскольку такой жанр публикаций ассоциируется с определенным типом объекта социологического исследования и, фактически, подпадает под сложившийся подход к анализу в других сферах социологии. Его разновидностью является описание одного коррупционного события в двух и более публикациях. Такой тип был определен как «Журналистское сопровождение события». Выделение типа статей по формальному признаку двух и более публикаций, посвященных одному событию, объединяло два разных повода обращения к одному и тому же объекту. Во-первых, это могут быть публикации, следующие одна за другой в непосредственной временной близости, развивая, уточняя, детализируя первоначально высказанную позицию, во вторых, информируя читателей о расследуемом уголовном деле по мере развития, связанных с ним событий. Часть публикаций описывают явления коррупции за рубежом и, таким образом, прямого отношения к нашему исследованию не имеют. В отдельных случаях публикации посвящены уголовным преступлениям, одним из условий совершения которых и явились имевшие место контакты преступников с коррумпированными элементами в органах власти. В таких публикациях собственно анализ фактов коррупции отсутствует, имеются только отдельные упоминания о ней, но которые тем не менее дополняют картину коррумпированности в обществе.

Отдельный разряд публикаций представляют собой статьи аналитического типа, содержащие обзорные материалы, обращающиеся к статистическим данным, с преобладающей оценочной составляющей. Сюда же относятся сообщения об установках на противодействие коррупции официальных лиц, выступающих в ходе публичных политических и деловых встреч.

Часть публикаций первоначально собранного массива подверглась выбраковке, поскольку, обладая компонентами, характерными для коррупции (корыстный мотив преступления должностного лица, использование служебных полномочий), квалифицировались как более тяжкие преступления, например, государственная измена. В ряде случаев типологизация сообщения оказывалась проблематичной. К какой группе следовало отнести получение должностным лицом незаконных доходов при продаже самолетов Миг в Колумбию, если уголовное дело против него возбуждено в Италии?

Мы придерживаемся в ходе изложения результатов избегать персонально указывать имена и фамилии участников, поскольку это уже

сделано в текстах публикаций, а задача социологического исследования состоит в исследовании характера и природы социальных явлений, а не рассмотрение конкретных деяний. Одним из общих принципов социологических обследований является принцип анонимности, хотя в данном случае, в силу происхождения первичного материала исследования, вопрос об анонимности не стоит. При необходимости делаются ссылки на источник информации о конкретных фактах, а имя и фамилия конкретных участников событий могут указываться в исключительных случаях, когда это обусловлено необходимостью представления сути дела узнаваемого уникального события.

Отказ от персональной идентификации событий обусловлен также и тем, что авторы публикаций не застрахованы от тенденциозности, предвзятости, подчас дают повод для обвинений в ангажированности. В ряде случаев первоначальная трактовка событий, послуживших предметом обвинения и судебного разбирательства, в дальнейшем не подтверждается.

Первоначально контент-аналитическое обобщение и оценка совокупности социальных фактов может быть произведена путем соотнесения отдельных элементов текстов публикаций с некоторой моделью, обобщенным представлением о событии, соответственно которой фиксируется ряд параметров конкретных событий, по оценке элементов текстов согласно позициям классификатора, который разработан в соответствии с принятой моделью и является фактической структурой формального измерения и регистрации отдельного события.

В случае исследования текстов о коррупции наш классификатор фиксирует характеристики субъекта коррупционного действия, характеристики деятельности, квалифицируемой как коррупционной, характеристики субъекта противодействия коррупции, характеристики стадии пресечения коррупционной деятельности.

В соответствии с этим общим контуром классификатор включает 11 позиций, имеющих от трех до девяти значений регистрируемых характеристик. Значения измеряемых характеристик вносятся в таблицу базы данных, в которой поля создаются соответственно числу позиций классификатора, каждая запись соответствует одному анализируемому случаю, в каждой ячейке соответствующего поля фиксируется цифровой код значения регистрируемой характеристики по определению кодировщика. Вводятся технические поля номера записи, кода источника текста, регистрируемого каждой записью и поле даты публикации, которых практически достаточно для возврата к исходному тексту в случае возникновения сомнений в точности классифицирования характеристик текста.

В целях информационной поддержки нами были введены поля, дублирующие кодируемые характеристики. В эти поля переносились фрагменты текстов публикаций, послужившие основанием для определения классификационной индексации конкретной характеристики события.

Такой способ двойной регистрации должен был сократить количество ошибок при кодировании, допускать возможность контроля при возникающих сомнениях по поводу точности проведенной идентификации, представлять материал для последующего создания словарей слов, несущих основную смысловую нагрузку при избранной модели контент-анализа. Однако, это все-таки дублирующий прием, хотя в методиках контент-анализа параллельное кодирование, дублирование измерение одних и тех же объектов разными методами достаточно обычная мера.

От качества классификации зависит качество результата. Реализация на практике этого очевидного постулата вызывает серьезные трудности. Только абстрактные аналитические понятия, заведомо фиксируемые в системе формальных отношений, естественно вписываются в структуру однозначной типологизации. Естественный же язык оперирует понятиями многозначными даже в пределах единого контекста, не говоря уже о тех случаях, когда контексты по одному и тому же поводу генерируются авторами с различным индивидуальным интеллектуальным опытом, ценностными позициями, ролевыми статусами в описываемых социальных ситуациях. Даже фундаментальные понятия культуры и наук о культуре не имеют строго разграничения. Для примера сошлемся на Й. Хейзингу. По поводу разграничения понятий Средневековья и Возрождения он пишет: «Далеко в Средневековье обнаруживали формы и явления, которые, казалось, уже несли на себе печать Ренессанса, и само понятие „Возрождение“ постоянно растягивали, заставляя его вбирать все эти признаки, пока оно в конце концов не утратило свою упругую силу. Но и обратно: тот, кто воспринимает дух Ренессанса без всякой предвзятой схемы, находит в нем гораздо больше „средневекового“, чем вроде бы допустимо с теоретической точки зрения. — и далее — И все же мы не можем отказаться от противопоставления: Средневековье и Возрождение стали для нас понятиями, в которых мы ощущаем на вкус различие в сути той или иной эпохи настолько же ясно, насколько мы отличаем вкус яблока от вкуса земляники, хотя вряд ли мы смогли бы описать это различие» [2, с. 306–307]. Действительно, ясное в целом становится проблематичным при накладывании схем на конкретное событие.

Констатирование некоторой характеристики события в классификаторе может быть зафиксировано на основании непосредственного свидетельства автора текста, либо как вывод кодировщика, возникающего как следствие некоторых наблюдаемых признаков данной характеристики. Но такой вывод уже косвенное заключение, помещаемое в классификатор, и уже сам кодировщик принимает решение о достаточности или недостаточности признаков для фиксирования кодировочной позиции. То есть субъект исследования привносит смысл в первичный текст, регистрируя некоторое элементарное, с точки зрения кода, значение. Если

подобная практика достаточно распространена, то объективность эмпирического измерения оказывается под вопросом.

Вообще, при создании программы контент-аналитического исследования социолог оказывается перед дилеммой: недостатка информации для разветвленной системы кодирования и избытка при ее ограничении. Очевидно, что публикации не создаются по одному шаблону, модель описания явления существует только в сознании исследователя, поэтому при тотальной регистрации текстов по единому стандартному классификационному перечню существенные характеристики в одних случаях отсутствуют, поскольку их нет в первичных источниках, в других, выпадают, поскольку нет соответствующей позиции классификатора.

Эта проблема может быть решена путем придания контент-анализу статуса качественного социологического исследования. Для статистически обоснованного вывода требуются достаточно большие выборки. Объем информации, привлеченной для контент-аналитического исследования публикаций о коррупции недостаточен для подобного результата. Вместе с тем охватываемый нашим исследованием период характеризуется определенной целостностью, все публикации о коррупции за этот период по трем избранным источникам вовлечены в круг первичных документов анализа. Таким образом, мы можем рассматривать их как полный комплект данных об объекте, ограниченном этими тремя источниками информации. Подобного рода проблемы может испытывать любое исследование, претендующее на статистическую достоверность. Например, формирование трендов экономического развития на основании показателей ВВП (валового внутреннего продукта) хотелось бы иметь за сорок, пятьдесят и более лет, отрезки времени сопоставимые с длинными волнами Кондратьева, но исходные данные ВВП на душу населения не будут объективно отражать тенденций экономического роста, поскольку уже за двадцатилетний период экономического развития условия производства претерпевают существенные изменения. С многофакторными моделями экономического роста дела обстоят еще сложнее, специалисты отмечают, что отрезок времени правдоподобного прогнозирования тенденций экономического роста весьма ограничен.

Контент-анализ локальной общности социальных фактов характеризует эту общность как таковую, но также эта общность обладает чертами типа явления в целом. Ряд характеристик имеет значение общезначимых, но в то же время их число может быть расширено за счет черт существенно значимых, но варьирующих в пределах данной локальной общности или даже проявляющихся в форме знаменательных, но уникальных символов. Регистрируемые коллизии могут расцениваться не только как феномены качественного порядка, но и требовать феноменологической интерпретации.

И если частотный анализ терминов может рассматриваться как дублирующая процедура получения результатов, полученных путем регистрации элементов текста кодировщиками, то фиксация текстовых фрагментов, выявляющих интенсивности идентификации категорий, альтернативных квалификаций актов, образующих регистрируемое событие, выделение участниками событий или авторами публикаций разных признаков и разных этапов совершающихся событий, явно относятся к сфере качественной и, в ряде случаев, к феноменологической социологии.

Первоначальная квалификация элементов текста в соответствии с классификатором позволяет идентифицировать единичные события актов коррупции по территориальному признаку по типу поселения, типу организации, иерархическому статусу субъектов коррупции в организации, сфере и типу коррупционных деяний и других признаков, а также ряд вторичных признаков, относящихся к сфере редакционной политики отдельных изданий. Здесь проблема исследования и цель исследования могут быть фиксированы достаточно четко и определенно. Но, как мы уже обращали внимание выше, классификатор всегда оказывается то слишком широк, то слишком узок для той наличной информации, которую представляют в распоряжение исследователя первичные документы. Легенда о Прокрустовом ложе фактически притча о возможностях классификации. Сама же задача содержательной постановки проблемы, первоначально сформулированной для целей объективной регистрации и измерения некоторого социального феномена шире ее формальной постановки, адаптированной к методу регистрации. В содержательном плане решение социологической проблемы подобно проблеме литературной и шире художественной критики, тем более, когда, как в случае контент-анализа, речь идет об исследовании текстов. Здесь подобие двух видов объектов становится больше чем просто аналогия. Задачи же художественной критики таковы. Что они не могут быть сведены к ответам на ряд заранее сформулированных вопросов, как это имеет место при формулировании гипотез социологического исследования. Художественная критика должна реагировать возможно более разносторонним образом на идею художественного произведения, его образы, на реальности, которую символизируют и представляют эти образы. Ее задачи — задачи рефлексивной интерпретации, рефлексивной в смысле реакций на многообразие образов и многообразие значений образов. Ее характерная черта первоначальная незаданность к выбору объектов внимания и критической трактовки. Она задана только соотношением с ценностями, вкусовыми и стилистическими комплексами эстетического кредо критика. Объективно же критика должна реагировать **на все**, то есть на любые факты художественного выражения, сколь прихотливым бы ни было воображения автора художественного произведения.



Первоначальная незаданность критики в отношении реакций на заранее сформулированный вопрос — не только привилегия эстетической критики. В предисловии к «Критике чистого разума» Кант, в отношении задачи исследования, констатирует незаданность конечного результата, метафорически определяя процесс метафизического рассуждения как «суд» [3, с. 9]. То же самое применимо и к дискурсивной методике социологического анализа документов. Рассматривая термины текста, соотносимые с исследуемыми категориями в контексте их непосредственных определений, то есть с категориями как субъектами суждений их предикативное обрамление, исследователь реагирует на ту ситуативную реальность, которая может быть соотносима с исходными гипотезами, а может и выходить за их пределы, или даже никоим образом с ними не соотноситься, что не обязательно снижает их содержательную значимость для целей исследования, но, безусловно, его качественно обогащает при последовательной интерпретации. Мы с уверенностью можем утверждать, что соотношение, совмещение и сочетание в методической практике компьютерного контент-анализа альтернативных подходов — традиционного, основанного на идентификации элементов текста вместе с частотным анализом элементов, с одной стороны, и, с другой стороны, дискурсивного анализа то есть качественного анализа категорий в контексте их ситуационных определений, или непосредственно связываемых с ними предикатов, является перспективным направлением развития метода в целом.

## Литература

1. Бронкмейер Й., Харе Р. Нарратив: проблемы и обещания одной альтернативной парадигмы // Вопросы философии. 2000. № 3.
2. Хейзинга Й. Осень Средневековья. М.: Наука, 1988.
3. Кант И. Критика чистого разума. М.: Мысль, 1994.