ТЕОРИЯ КОНТЕНТ-АНАЛИЗА ДЛЯ ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЙ

Б.Ж. Ельмуратова, П.С. Ельников
ПГУ, г. Павлодар, Казахстан

В настоящее время историки приходят к осознанию того, что изучать отдельную эпоху или отдельный регион, страну более логично в длительной перспективе, опираясь на компаративные подходы. В этом случае становится понятным и очевидным, что возникновение источников имеет свою внутреннюю логику: они появляются и уходят с исторической арены закономерно. Поэтому возможно определение принципов и методов работы с источниками. И, как правило, эти принципы и методы разрабатываются на стыке смежных гуманитарных и естественных дисциплин [1].

Как писал А. Гершенкрон, «вклад историка заключается в том, чтобы указать на потенциально релевантные факторы и на потенциально значимые сочетания факторов, которые не столь легко определить изнутри сферы более ограниченного опыта. Но это – вопросы. Ответы – это уже другой предмет [2].

Самым распространенным видом информации является информация, представленная в виде текстов на языке данной страны, поэтому анализ текстов является одним из наиболее распространенных видов научного и научно-практического анализа. Более того, существуют науки, которые целиком или в основном опираются именно на анализ текстов. Традиционные формы сжатия информации – аннотирование, конспектирование, реферирование – уже давно не считаются какими-то специфическими видами работы с информацией и ими владеют любые специалисты.

На практике анализ большинства источников выражен лишь одними качественными (атрибутивными) признаками. Это и приводит к перенасыщению биографического раскрытия темы исследования. Хотя используется традиционный арсенал методов исследования, проверенный временем, выводы остаются противоречивые и порождают сомнения [3, с. 2].

Понятие анализ текстов иногда употребляется как синоним понятия контент-анализ, хотя последнее шире. Контент-анализ относится к более широкой области исследований, затрагивающих не только текст, но и информацию различного рода – изображения, аудио и видеоинформацию. Кроме того, контент-анализ, в отличие от других способов изучения документов, позволяет вписать содержание документа в социальный контекст, осмыслить его как проявление или как оценку социальной жизни. Понимание социального контекста документа предполагает выявление того, что именно получило в нем отражение, какой резонанс этот документ получил или может получить в общественной жизни и, наконец, степень оригинальности документа, отличие его от других документов такого рода.

Необходимым условием применения методики анализа содержания является наличие материального носителя информации. Во всех случаях, когда существует или может быть воссоздан такой носитель, допустимо использование методики контент-анализа [4].

Он подразумевает поиск, регистрацию и числовую обработку определенных смысловых единиц, зафиксированных в документе (или в совокупности документов), и дальнейшую интерпретацию полученных данных (описание обнаруженных закономерностей, установление взаимосвязи между выявленными тенденциями и т.д.) [5, с. 31].

Наивысшего пика в своем развитии контент-анализ достиг при Робер те Мертоне, который разработал уникальную классификацию типов метода. Первый тип основан на элементарном подсчете символов. Второй представляет собой дальнейшую разработку подсчета символов, при котором символы классифицируются в соответствии с использованием в благоприятном или неблагоприятном аспекте. Третий тип контент-анализа представляет собой анализ по элементам, который заключается в классификации отдельных частей материалов и подразумевает выбор главных и второстепенных частей материалов [6, с. 85].

Сначала исследователи пользовались двумя определениями – «традиционный» и «формализованный», затем – «формализованный» и «неформализованный». В данной работе остановимся на следующих определениях.

Неформализованный (традиционный или классический) анализ – это цепочка умственных, логических построений, направленных на выявление сути анализируемого материала. Он позволяет улавливать основные мысли и идеи, прослеживать их генезис, выяснять их обоснования, взвешивать следствия, охватывать глубинные, скрытые стороны содержания документа: этот анализ стремится как бы до конца проникнуть в глубь документа. При проведении неформализованного анализа исследователь должен ответить на ряд вопросов, среди которых: что представляет собой документ, каков его контекст, кто его автор, какова надежность самого документа и т.д.

Кроме того, в неформализованном анализе выделяют внешний и внутренний анализ. Внешний анализ – это анализ «исторического контекста» документа, т.е. анализ контекста документа в собственном смысле этого слова и всех тех обстоятельств, которые сопутствовали его появлению. Тогда как внутренний анализ – это исследование содержания документа. Недостатком метода классического анализа содержания текстов является излишняя субъективность исследователя.

Суть формализованного, или, как его называют, количественного анализа заключается в том, чтобы найти такие легко подсчитываемые признаки, черты, свойства документа, которые с необходимостью отражали бы определенные существенные стороны содержания.

К основным процедурам контент – анализа относятся:

1) формирование выборки (только в очень малом количестве случаев можно использовать весь материал по определенной проблеме, как правило, происходит многоэтапный процесс отбора). Под выборочным методом подразумевается такая система отбора единиц для наблюдения, при которой полученные результаты были бы присущи явлению в целом. Эти системы могут быть различны и зависят от поставленной исследовательской задачи. Как уже было сказано ранее, этап определения фильтров поиска полезных в исследовании источников информации и отсеивание ненужных, несвязанных с темой, либо не имеющих весомых связей с ней, является очень важным и не должен рассматриваться как формальный, механический и малозначимый этап работы. Ошибки при выборе типологических групп источников могут отрицательно сказаться на всех этапах последующего исследования;

2) выявление смысловых единиц контент-анализа (понятия, выраженные определенными терминами; темы; имена, фамилии людей; события, факты, смысл апелляций к потенциальному адресату);

3) интерпретация полученных данных (перевод количественной информации в качественное представление рассматриваемого результата работы);

4) дальнейшая перепроверка результатов исследования (полученные результаты не являются каноном, а есть более точное представление рассматриваемого исторического события).

Контент-анализ как экстенсивный, обобщающий анализ, успешнее всего работает с массовыми источниками, отличительными признаками которых являются постоянность, однородность, одномасштабность, повторяемость. Именно в массовой исторической документации частота повторения смысловой единицы, категорий анализа имеет тенденцию отражать ее значимость. В повествовательном источнике связь между частотой и значимостью нередко нарушается, а может и вообще не наблюдаться.

М.Й. Лауристин отмечает, что контент-анализ не может быть только качественным или количественным.

В дискуссии о качественном или количественном контент-анализе существует и объединяющая точка зрения, которая представляется наиболее продуктивной. Ее защитники утверждают, что должно использоваться некоторое сочетание количественного и качественного анализа.

Тематический или семантический подход. Основной отличительной особенностью тематического подхода является то, что выводы делаются на основании данных о встречаемости концептов, понятий в тексте. Основное допущение подхода состоит в том, что существует связь между наличием в тексте тех или иных тем и интересом к этим темам у автора текста. Сделав это допущение, исследователь может ставить вопрос об уровне интереса, проявляемого тем или иным коммуникатором к той или иной теме. Можно также сравнивать интерес к теме у разных коммуникаторов или в различные периоды времени.

Применимость тематического анализа определяется следующими основными критериями:

1) можно сформулировать четкие и однозначные правила кодирования;

2) единицей анализа выступает слово или устойчивое словосочетание, выводы будут делаться на основе частот встречаемости слов или словосочетаний в тексте;

3) в анализ включено большое количество концептов. Основное различие между тематической и семантической стратегиями – в ответе на вопрос «что считать?». В тематической стратегии кодируются концепты, семантическая предполагает подсчет и кодирование отношений между концептами. Преимуществом семантического анализа по сравнению с тематическим является то, что он позволяет сохранить

больше изначальной сложности, больше оригинальных смыслов исследуемого текста.

Семантика – смысловая сторона языка, отдельных слов и частей слова.

Семантический анализ в значительно большей мере, чем тематический, опирается на свойства естественного языка. Поэтому анализ (особенно компьютерный) в значительной степени специфичен для каждого языка. Каждый блок текста, попавший в выборку, суммируется при помощи нескольких кодов, взаимосвязанных в соответствии с общим пониманием предмета изучения.

Кодирование – это процесс систематической трансформации и агрегации исходных данных в категории, которые позволяют точно описать характеристики текста, релевантные для исследования. Основная проблема кодирования – неоднозначность текста. Слова могут иметь различные коннотации. Набор категорий создается на основе теории, которая есть у исследователя еще до начала исследования. В таком случае, целью исследования может быть подтверждение или опровержение теории.

Выделяется два базовых подхода к кодированию: кодирование открытого, эксплицитно выраженного содержания текста (манифестное или инструментальное) и кодирование скрытых, неявно выраженных смыслов, которые, по мнению исследователя, подразумевались создателем текста (латентное, репрезентационное).

Манифестное кодирование.

Кодирование содержания текста, лежащего на поверхности, называется манифестным (открытым, явным). Исследователь подсчитывает количество появлений фразы или слова.

Латентное кодирование.

Исследователь, который использует латентное (скрытое) кодирование, ищет скрытые, имплицитные значения содержания текста. Например, он прочитывает весь абзац целиком и решает, присутствует ли в его содержании та или иная направленность.

Следующим этапом проведения контент-анализа является составление словаря.

Условно можно разделить словари на два вида: частотные и семантические. Первые подразумевают выделение единиц контент-анализа на основе частоты их использования по отношению к суммарному количеству слов в массиве и к другим единицам, потенциально подходящим для включения в словарь.

Второй вид представляет собой включение в словарь категорий и единиц на основе заранее проработанных текстов, максимально точно

описывающих предмет исследования, а потому уже содержащих большинство единиц будущего словаря.

Квантификация и интерпретация результатов проведения контент-анализа. По завершению подготовительного периода следует этап работы с единицами подсчёта, выбранными в соответствии с установленной системой категорий, опирающейся в свою очередь на цель, задачи и гипотезы исследования. Здесь исследователь прибегает к помощи таких инструментов как регистрационная карточка или кодировальная матрица, бланк контент-анализа, также называемый протоколом итогов контент-анализа. На основе систематизированного и дифференцированного материала исследователь пишет работу – записку по результатам контент-анализа, опираясь, главным образом, на протокол итогов, полученный в ходе полной дистрибуции категорий текстового массива.

Контент-анализ – это наиболее надежный метод текстового анализа среди эмпирических методов исторического исследования. Однако в связи с немалым количеством литературы, описывающей этот метод, однородное понимание контент-анализа затруднено. Для историка важен не сам процесс разгадывания сложного механизма применения метода (с учетом всех его процедур), для него более интересна возможность подтверждения, уточнения имеющегося результата исследования, проведенного традиционными методами исторического исследования, такими как историко-генетическим, историко-сравнительным, историко-типологическим, историко-системным, диахронийческим, ретроспекцией [3, с.6].

Контент-анализ не должен заменить обычный анализ документов. Он его дополняет, а их сочетание углубляет понимание смысла текста. Контент-анализ позволяет обнаружить в документе то, что ускользает от поверхностного взгляда при его традиционном изучении, однако имеет важный социальный смысл.

Литература

1. Шотбакова Л.К. Место и роль источника в исторических исследованиях [Электронный ресурс] / Л.К. Шотбакова. – Электрон. текстовые дан. – Караганда: Вестник КарГУ, 2005. – Режим доступа: https://articlekz.com/article/4658, свободный, 2019 г.

2. Economicus.Ru. Экономическая отсталость в исторической перспективе [Электронный ресурс] / А. Гершенкрон. – Электрон. текстовые дан. – Москва: ИСТОКИ, 2004. – Режим доступа: http://gallery.economicus.ru/cgi-bin/frame_rightn_newlife.pl?type=in&links=./in/gershenkron/works/gershenkron_w1.txt&img=works.jpg&name=gershenkron, свободный, 2019 г. – Электрон. версия печ. публикации. Перевод Г.Д. Гловели

3. Приборович А.А. Контент-анализ – форма исторического исследования [Электронный ресурс] / А.А. Приборович. – Электрон. текстовые дан. – г. Минск: Белорусский государственный университет, 2011. – Режим доступа: http://elib.bsu.by/bitstream/123456789/3474/1/%d0%90.%d0%90.%20%d0%9f%d1%80%d0%b8%d0%b1%d0%be%d1%80%d0%be%d0%b2%d0%b8%d1%87%202.pdf, свободный. – Электрон. версия печ. публикации, 2019 г.

4. Контент-анализ [Электронный ресурс] / Коллектив авторов. – Электрон. текстовые дан. – США: Фонд Викимедиа, 2019. – Режим доступа: https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BD%D1%82%D0%B5%D0%BD%D1%82-%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7, свободный. – Интернет-энциклопедия, 2019 г.

5. Чурашева О.Л. Применение метода контент-анализа в исследовании «публичная библиотека в обеспечении информационно-психологической безопасности читателей» [Электронный ресурс] / О.Л. Чурашева. – Электрон. текстовые дан. – г. Москва: Электронная библиотека, 2019. – Режим доступа: https://cyberleninka.ru/article/n/primenenie-metoda-kontent-analiza-v-issledovanii-publichnaya-biblioteka-v-obespechenii-informatsionno-psihologicheskoy-bezopasnosti, свободный. – Кибер Ленинка-научная электронная библиотека, 2019 г.

6. Винокурова И.С. Применение компьютера для контент-анализа социологической информации [Электронный ресурс] / И.С. Винокурова. – Электрон. текстовые дан. – SocioLogos, 2019. – Режим доступа: http://www.sociologos.ru/metody_i_tehnologii/Razdel_Analiz_dannyh/Kontent-analiz_i_desk_research/Vinokurova_I_S_Primenenie_kompyutera_dlya_kontent-analiza_sociologicheskoj, свободный. – SocioLogos – Портал российской прикладной социологии, 2019 г.