

# Новые тенденции в компьютеризованном анализе текстов: концепции, методы, технологии

## New Tendencies in Computer-Assisted Text Analysis: Concepts, Methodology, Technology

Авторы: Ирина Гарскова

Ссылка для цитирования:

Гарскова И. М. Новые тенденции в компьютеризованном анализе текстов: концепции, методы, технологии // Электронный научно-образовательный журнал «История», 2015. Т.6. Выпуск 8 (41) [Электронный ресурс]. Доступ для зарегистрированных пользователей. URL: <http://history.jes.su/s207987840001255-9-1> (дата обращения: 04.11.2015). DOI: 10.18254/S0001255-9-1

Ключевые слова: компьютеризованный анализ текстов, нарративные источники, качественный и количественный анализ, контент-анализ, парадигма смешанных методов, программное обеспечение, информационные ресурсы

Ключевые слова (англ): computer-assisted text analysis, historical narrative sources, qualitative and quantitative analysis, content analysis, mixed methods paradigm, software, information resources

Аннотация (рус): В статье дается обзор современного состояния методологии, методики и технологии исследований в области компьютеризованного анализа текстов. Это междисциплинарное направление востребовано в самых разных областях гуманитарного знания, в том числе и в исторических исследованиях. В работе с текстами историки многое заимствовали у социологов и у представителей других социально-гуманитарных наук. Однако есть определенные проблемы, и прежде всего, это отсутствие базовых монографических изданий для гуманитариев по современному компьютеризованному анализу текстов. В любых гуманитарных приложениях учебники по методам исследования уделяют анализу текста в лучшем случае один из разделов. Об отсутствии базовой литературы по применению контент-анализа в исторических исследованиях свидетельствует то, что в своих историографических обзорах многие авторы механически пересказывают одни и те же фрагменты классических публикаций 1970—1980-х гг., зачастую без ссылок на источники либо с недостоверными ссылками, взятыми с интернет-сайтов. Поэтому необходимо как осмысление накопленного за последние два-три десятилетия опыта, так и оценка возможных перспектив с учетом таких важных факторов, как появление парадигмы интеграции качественных и количественных подходов в гуманитарных исследованиях и перспективы компьютеризованного анализа текстов, обеспеченного новым поколением программных продуктов и быстро растущим объемом электронных текстовых архивов.

Аннотация (англ): The article gives an overview of the current state of concepts, methodology and technology in the field of computerized text analysis. The interdisciplinary methodology and technology of text analysis is developing in various fields of humanities, including historical research. In this research area historians borrowed much from sociologists and representatives of other social sciences and humanities. However, there are certain problems and, above all, it is the lack of basic monographs for the humanities on computer-assisted text analysis. Any textbook on research methods for the humanities includes the best, only one small section devoted to the analysis of texts. In the absence of the basic literature on the use of content analysis in historical research, many authors in their historiographical review mechanically regurgitate the same fragments of classical publications of 1970—1980's. These historiographical reviews often contain no references or contain unreliable references copied from Internet sites. It is therefore necessary to rethink the experience accumulated over the past two or three decades, and to evaluate the possible prospects considering such important factors as the emergence of the paradigm of the integration of qualitative and quantitative approaches in humanities research and prospects for computerized text analysis provided by the new generation of software products and the rapidly growing volume of electronic text archives.

Дата публикации: 04.11.2015

№	Текст	Примечания
1	Анализ текстов в настоящее время является чрезвычайно востребованным и популярным направлением в самых разных областях социально-гуманитарного знания: социологии, социальной психологии, политологии, лингвистике, журналистике и многих других. Самым известным направлением в нетрадиционном, формализованном анализе текстов является, безусловно, контент-анализ, который имеет разные дефиниции, имеющие в качестве общей характеристики метода то, что контент-анализ изучает содержание текстовых документов.	
2	Контент-анализ с самого начала заявил себя как междисциплинарный метод (или даже — система методов), поэтому представители многих социально-гуманитарных наук склонны считать этот метод «своим». Поэтому любой обзор	1. Krippendorff K. Content Analysis. An Introduction to its Methodology. London: Sage Publications, 2004. 2nd edition. P. 4.

истории зарождения контент-анализа, к какой бы гуманитарной области он ни относился, оперирует одними и теми же именами «классиков» (Г. Лассуэлла, Б. Берелсона, П. Лазарсфельда, П. Стоуна, Л. Лоуэнталя и других), развивавших различные приложения концепции контент-анализа и привлекает известные примеры ранних контент-аналитических исследований, когда еще не утвердился сам термин «контент-анализ», прослеживая их с начала XX в. или даже с XVIII в.<sup>1</sup> Большинство этих примеров иллюстрируют возможности контент-анализа на материалах массовой коммуникации, прессы.

3 В отечественную историографию контент-анализ пришел на рубеже 1960—1970-х гг. через публикации социологического семинара, организованного Сибирским отделением советской социологической ассоциации и рабочего совещания социологов «Методологические и методические проблемы контент-анализа»<sup>2</sup>. Первые обстоятельные обзоры контент-анализа в социологии в начале 1970-х гг. принадлежат перу А. Н. Алексеева<sup>3</sup>. Интересно, что в работах А. Н. Алексеева рассматриваются разные определения контент-анализа, включая такие, где он называется техникой исследования<sup>4</sup>, и такие, где он относится к методологии. Эволюция отношения к контент-анализу, по мнению автора, это движение от его инструментальной трактовки как техники к его трактовке как методологии<sup>5</sup>. Важным акцентом в упомянутых публикациях является междисциплинарность метода контент-анализа и постановка вопросов связи контент-анализа с теорией информации, семиотикой, лингвистикой, математической статистикой — в методическом и технологическом плане, и с историей, документоведением, литературоведением, библиографией, архивным делом — в плане взаимодействия. Так, наибольший интерес для А. Н. Алексеева во взаимодействии социологии и истории представляли методы источниковедения<sup>6</sup>. Упоминания о важности для социологов методического опыта смежных гуманитарных наук встречаются и в других публикациях. Например, при ретроспективном анализе документальных источников социологам необходимы знания основ источниковедения, архивоведения и документоведения. Однако становление «источниковедения для социологов», адаптация методов источниковедения к специфике предмета социологии пока находится в начальной стадии<sup>7</sup>.

4 И все-таки не случайно, что в исторические исследования контент-анализ пришел из социологии. Немаловажную роль в развитии исторических контент-аналитических исследований сыграла разработка И. Д. Ковальченко концепции явной и скрытой информации исторического источника<sup>8</sup>. Применение метода формализованного анализа текста, как одного из способов повышения информационной отдачи исторического источника, привлекло внимание многих исследователей. В 1980-х гг. появились работы, дающие представление о методе контент-анализа и обобщающие опыт первых контент-аналитических исторических исследований<sup>9</sup>. Многие из них проводились на материалах исторической прессы.

5 Поскольку в работе с текстовыми источниками историки во многом следовали подходам, апробированным социологами, нельзя обойти те методологические и методические проблемы, которые сегодня обсуждаются в социологических исследованиях<sup>10</sup>. Прежде всего, это проблемы дихотомии т. н. качественного и количественного анализа предметной области (хотя однозначного понимания этих терминов в литературе нет, и четкая граница между ними отсутствует)<sup>11</sup>. Качественный, традиционный подход чаще ассоциируется с микроанализом и опирается на неформализованные, «мягкие» методы сбора и анализа информации, «понимающее» восприятие содержания документа. В социологии источниками такой информации являются наблюдение, интервью, включающие исследователя непосредственно в процесс сбора информации.

2. Проблемы контент-анализа в социологии: Материалы Сибирского семинара социологов. Новосибирск, 1970; Методологические и методические проблемы контент-анализа. М.—Л., 1973.

3. См., например, Алексеев А. Н. Контент-анализ: его задачи, объекты и средства // Социология культуры. Вып. 1. М.—Л., 1974. С. 131—162.

4. Там же. С. 131—132.

5. Алексеев А. Н. Контент-анализ: техника или методология? (К постановке проблемы) // Методологические и методические проблемы контент-анализа. Вып. 1—2. М.—Л., 1973. С. 19—24.

6. Алексеев А. Н. Контент-анализ в социологии и точки соприкосновения с другими отраслями знания // Проблемы контент-анализа в социологии... С. 11—17.

7. Маслова О. М., Пугачева М. Г. Виды документальных источников и методы их анализа // Методы сбора информации в социологических исследованиях / отв. ред. В. Г. Андреев и О. М. Маслова. М., 1998. Кн. 2. С. 41—42.

8. Ковальченко И. Д. Исторический источник в свете учения об информации // История СССР. 1982. № 3.

9. Бородкин Л. И. Контент-анализ и проблемы изучения исторических источников // Математика в изучении средневековых повествовательных источников. М., 1986. С. 8—30; Бородкин Л. И. Многомерный статистический анализ в исторических исследованиях. М., 1986. С. 138—160, 160—173; Ковальченко И. Д., Бородкин Л. И. Современные методы изучения исторических источников с использованием ЭВМ. М., 1987. С. 11—35; Миронов Б. Н. Историк и социология. Л., 1984.

10. Сходные методологические проблемы, связанные с противопоставлением качественного и количественного подходов, обсуждаются и в смежных социально-гуманитарных науках.

11. Толстова Ю. Н., Масленников Е. В. Качественная и количественная стратегии. Эмпирическое исследование как измерение в широком смысле // Социс. 2000. № 10. С. 101—102.

12. Семенова В. В. Качественные методы: введение в гуманистическую социологию: Учеб. пособие для студентов

- Количественный подход чаще имеет дело с формализованными, «жестко» структурированными данными, такими, как анкеты, личные листки и др.; на смену прямого контакта исследователя с объектом исследования приходит «статистическое видение» и статистическое обобщение результата анализа на генеральную совокупность<sup>12</sup>.
- 6 В основе «качественно-количественной» дихотомии исследовательских стратегий можно видеть противостояние социологического реализма (количественный подход, структурализм) и социологического номинализма (качественный подход, феноменология), хотя на практике ни один ученый не является ни реалистом, ни номиналистом в чистом виде, в каждом сосуществуют оба качества<sup>13</sup>.
- 7 В историческом исследовании, прежде всего, в силу отсутствия прямого контакта исследователя со своим объектом, качественный анализ связывается с определенным происхождением данных, то есть с анализом источников личного происхождения, таких, как письма, дневники, мемуары и тму подобное, с «естественными» выборками, историко-культурной или историко-антропологической проблематикой исследования, отсутствием строгой категоризации. Здесь можно провести параллели с известной дискуссией «Методологические проблемы исторической информатики и квантитативной истории», которая состоялась международной ассоциации History & Computing в 1996 г.<sup>14</sup> Дискуссия затронула такие важные методологические вопросы исторического познания, как соотношения номотетического и идеографического подходов, взаимоотношения между историей, социальными и естественными науками. Обсуждались также возможности применения количественных методов в исторических исследованиях (проблемно-ориентированный подход) и их ограничения, связанные с неколичественным, нечетким и неформализованным характером исторических данных (источнико-ориентированный подход). Как показала упомянутая дискуссия, для исторического исследования актуальны и востребованы оба подхода, равно как и их комбинация.
- 8 Рассматривая проблему «Quality and Quantity», можно утверждать, что первой появилась парадигма качественного подхода: на рубеже XIX—XX вв. было очень мало исследований с применением простейшего количественного анализа, социологи и психологи в основном использовали полевые методы и работали с большими объемами визуальных данных (фотографии). Более серьезные статистические методы (проверка статистических гипотез) стали широко использоваться после 1915 г.<sup>15</sup>
- 9 Хотя еще в 1952 г. впервые был предложен термин «качественный контент-анализ»<sup>16</sup>, довольно долго продолжалось изначальное доминирование контент-анализа как количественного метода работы с текстами, и лишь через несколько десятилетий этот термин был введен в научный оборот<sup>17</sup>. В 1990-х гг. отмечается заметный рост интереса к исследованиям, выдвигающим на первый план «качественный анализ»<sup>18</sup>, и противостояние качественного и количественного подходов. Оценивая тенденцию противостояния «качественников» и «количественников» среди социологов как «некую болезнь научного социологического сообщества», Ю. Н. Толстова и Е. В. Масленников<sup>19</sup> предлагают более рациональный подход, основанный на понимании того, что выбор метода в каждом конкретном исследовании определяется исследовательской задачей, объектом и предметом, источниковой базой. Ученый всегда использует те методы, которые являются наиболее эффективными, не задаваясь вопросом, какую стратегию, количественную или качественную, следует предпочесть, а органически сочетая качественные и количественные методы.
- 10 В последние годы на смену противопоставлению качественных и

13. Толстова Ю. Н., Масленников Е. В. Указ. соч. С. 107.

14. Круглый стол «Методологические проблемы исторической информатики и квантитативной истории» // Информационный бюллетень Ассоциации «История и компьютер». 1996. № 19. С. 62—195.

15. Kuckartz U. Realizing Mixed-Methods Approaches with MAXQDA [Электронный ресурс]. URL: <http://www.maxqda.com/download/MixMethMAXQDA-Nov01-2010.pdf>

16. Kracauer S. The Challenge of Qualitative Content Analysis. Public Opinion Quarterly, 1952, № 16. P. 631—642. Цит. по: Kuckartz U. Qualitative Text Analysis. A Guide to Methods, Practice and Using Software. Sage Publications, 2014. P. 31.

17. Mayring Ph. Qualitative Inhaltsanalyse. Grundlagen und Techniken (1st ed.). Weinheim, 1983; Mayring P. Qualitative Content Analysis // Forum Qualitative Sozialforschung / Forum: Qualitative Social Research — [Электронный журнал]. Volume 1, No. 2, Art. 20 — June 2000. Mode of Access: <http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385>

18. См. список учебной и научной литературы по качественным методам в социологии в статье: Толстова Ю. Н., Масленников Е. В. Указ соч. С. 108.

19. Толстова Ю. Н., Масленников Е. В. Указ соч.

20. Teddlie Ch., Tashakkori A. Major issues and controversies in

количественных подходов в 1970—1990-х гг. приходит идея интеграции этих подходов, хотя и сегодня многие исследователи признают по-настоящему научными лишь количественные методы, а многие сторонники качественного подхода отвергают основные принципы подхода количественного. Тем не менее, сформировался новый подход или «третья парадигма»<sup>20</sup>, которую можно обозначить как интеграцию или комбинацию качественного и количественного подходов в исследовании. Эта вполне естественная интеграция может быть явной, основанной на методологическом выборе, даже если это просто дань современному дизайну научного проекта, так и неявной, естественно следующей за решением конкретной исследовательской задачи, без особого акцента на каком-то особом методологическом подходе.

the use of mixed methods in the social and behavioral sciences. Sage, 2003.

11 Комбинация качественных и количественных методов обычно связывается с такими терминами, как «триангуляция», «смешанные методы» или «методологическая интеграция». «Триангуляция» и «смешанные методы» ориентированы на практику конкретных исследовательских проектов, а «методологическая интеграция» на концептуальные, методологические аспекты исследования<sup>21</sup>. «Методологическая интеграция» была предложена У. Келле<sup>22</sup>, который рассматривал комбинацию различных методов под углом зрения философии и методологии науки, а не исследовательской практики в духе дискуссий естественных и гуманитарных наук об объяснении и понимании, которая восходит к началу XX в.

21. См.: Kuckartz U. Realizing Mixed-Methods Approaches...

22. Kelle U. Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte, Wiesbaden, 2007.

12 Термин «триангуляция»<sup>23</sup> обозначает комбинацию различных методов в изучении одного явления (объекта). Метафора «триангуляция» заимствована из геометрии (топографии), где она обозначает процесс определения положения объекта путем соотнесения его с известными объектами с помощью построения треугольников, измерения их углов и расчета расстояний; возможно, метафора является не слишком удачной, поскольку исследователи часто используют в своих исследованиях не три, а большее число методов.

23. Впервые предложен в работе: Denzin N. The research act. New York: McGraw-Hill, 1978. Идея триангуляции развивается в некоторых работах отечественных исследователей. См., например: Олейник А. Н. Триангуляция в контент-анализе. вопросы методологии и эмпирической проверки // Социологические исследования. 2009. № 2. С. 65—79.

13 Парадигма «смешанных методов» (mixed methods) является наиболее популярной, особенно в американской социологии. Ее протагонисты подчеркивают, что она преодолевает ограниченность двух прежних парадигм (качественной и количественной). Можно рассматривать различные типы стратегии «смешанных методов», исходя из следующих критериев: обрабатывается ли набор качественных и количественных данных; с помощью одного из методов или обоих сразу; является ли приоритетным использование одного из двух методов; имеет ли место интеграция данных и интеграция методов; какова роль теории — неявная или четко сформулированная с самого начала исследования<sup>24</sup>.

24. Creswell J. Research Design. Qualitative, quantitative and mixed methods approaches. Sage Publications, 2003.

14 В отечественной социологии и других социально-гуманитарных науках интерес к этим вопросам проблемам пока невелик. Например, Е. А. Каневский (Экономико-математический институт РАН, СПб) и Г. И. Саганенко (Социологический институт РАН, СПб) в своей статье<sup>25</sup> ставят проблему отсутствия в отечественной социологии монографических научных и учебных публикаций по контент-анализу, а также и отсутствия у социологов интереса к освоению соответствующих компьютерных программ. Статья, о которой идет речь, была опубликована в 1997 г. Через 10 лет, в 2007 г. Г. В. Каныгин (Институт социологии РАН) снова констатирует ту же проблему<sup>26</sup>. Между тем, программное обеспечение компьютеризованного анализа текстов заслуживает серьезного внимания.

25. Каневский Е. А., Саганенко Г. И. Концептуальное обоснование компьютерного анализа массивов с текстами // Социология: 4М. 1997. № 9. С. 1—8.

26. Каныгин Г. В. инструментальные средства и методологические принципы анализа качественных данных // Социология: 4М. 2007. № 25. С. 70—97.

15 Уже с конца 1960-х гг. в анализе текстов стали использоваться компьютерные программы анализа текстов. Появление таких программ радикально ускорило самые трудоемкие операции работы с текстами: в первую очередь, это создание т. н. инвертированных списков (словарей текста, содержащих все грамматические формы всех слов с подсчетом их частот и позиций в исходном тексте), поиск по заданным критериям слов и словосочетаний без контекста

(KWOC — KeyWords Out of Context) и в контексте (KWIC — KeyWords in Context).

- 16 Примеры таких программ, разработанных в 1960 — начале 1980-х гг., рассмотрены в обзорах Л. И. Бородкина и И. Д. Ковальченко<sup>27</sup>. В этих публикациях рассматриваются два направления компьютеризации анализа текстов: разработка систем с фиксированным словарем и систем с открытым (свободным) словарем. Основываясь на заранее заданном списке категорий или же комбинируя в категории слова из словаря текста, эти системы позволяли затем обрабатывать полученные в результате частотные данные с помощью методов статистического анализа. Наиболее известной программой с фиксированным словарем, содержащим ограниченное число смысловых категорий таких предметных областей, как социология и психология, была система GENERAL INQUIRER<sup>28</sup>. Примером системы со свободным словарем может служить программа WORDS<sup>29</sup>. Оба направления имели свои ограничения. Системы с фиксированным словарем предлагали ограниченное число категорий и индикаторов и не работали с остальными словами текста, системы со свободным словарем не способны решать задачи интерпретации автоматически созданных категорий<sup>30</sup>.
- 17 Продолжение обзора программ компьютеризованного анализа текста, разработанных в следующем десятилетии, можно найти в статье А. Н. Петрова<sup>31</sup>. Среди программ «нового поколения» наибольшее внимание автор уделяет программам Word Cruncher и TACT, которые обладают большими возможностями, чем программы 1970—1980-х гг., например, позволяют строить конкордансы, получать не только дескриптивную статистику (частоты встречаемости) категорий, но и изучать их взаимосвязи (частоту совместной встречаемости) с подсчетом статистического критерия значимости этих взаимосвязей. Например, в программе TACT на основе сравнения частоты встречаемости слова в полном тексте документа с частотой его встречаемости в минитексте, то есть «в окрестности» другого слова (как правило, в контексте некоторой семантической категории) подсчитывается величина Z-score. Если эта величина превышает определенное пороговое значение, это свидетельствует о том, что вероятность случайного «попадания» слова в «окрестность» данной категории очень мала, т. е. найденная связь является статистически значимой. В тексте данной статьи упоминаются такие характерные черты развития компьютеризованных методов контент-анализа, как растущее число доступных в сети Интернет полнотекстовых ресурсов, в том числе коллекций оцифрованных исторических источников, что связано, в частности с совершенствованием аппаратного и программного обеспечения для сканирования и оптического распознавания текстов и перевода их в электронный формат.
- 18 В полной мере возможности компьютеризованного контент-анализа появляются на рубеже XX—XXI вв. и в начале XXI в. Так, среди нескольких десятков программ анализа текстов в этот период лидируют те, которые полностью поддерживают различные виды интеграции качественных и количественных подходов, например, MAXQDA, QDA Miner, NVivo<sup>32</sup>. Авторы контент-аналитических исследований выбирают программное обеспечение, руководствуясь критериями доступности, удобства и простоты использования, возможности работы с текстами на разных языках.
- 19 Отметим также наличие в сети ресурсов, содержащих информацию о текстовых архивах, программном обеспечении, методические материалы по работе с программами текстового анализа, историографию и библиографию традиционного и компьютеризованного анализа текстов и даже предлагающих обработку текстовых массивов пользователей в режиме удаленного доступа<sup>33</sup>.

27. Бородкин Л. И. Контент-анализ и проблемы изучения исторических источников; Бородкин Л. И. Многомерный статистический анализ; Ковальченко И. Д., Бородкин Л. И. Современные методы изучения исторических источников.

28. Stone P. J., Dunphy D. C., Dexter C., Smith M. S., Ogilvie D. M. The General Inquirer: A Computer Approach to Content Analysis. Cambridge, 1966.

29. Iker H. P., Harway N. J. Computer Systems Approach toward the Recognition and Analysis of Content // The Analysis of Communication Content. Ed. by G. Gerbner, O. R. Holsti, K. Krippendorf, W. J. Paisley, Ph. J. Stone. New York, 1969.

30. См., например, Бородкин Л. И. Многомерный статистический анализ... С. 152—153.

31. Петров А. Н. Компьютерный анализ текста: историография метода. // Круг идей: модели и технологии исторической информатики. М., 1996. С. 255—277.

32. Известны и некоторые отечественные разработки компьютерных систем анализа текстов. Например, система ДИСКАНТ (Диалоговая Итерационная Система Классификации и Анализа Текстов) для обработки анкетной социологической информации (см. Каневский Е. А., Саганенко Г. И. Концептуальное обоснование компьютерного анализа массивов с текстами // Социология: 4М. 1997. № 9) или ВЕГА (Боярский К. К., Каневский Е. А. ВЕГА — компьютерная система классификации и анализа текстов // Вестник СПбГУ ИТМО. СПб., 2009, № 5 (63). С. 98—105).

33. Например, портал TAPoR (Text Analysis Portal for Research), предлагающий большой набор исследовательских инструментов для анализа и визуализации текстов (в основном, англоязычных) онлайн (<http://www.tapor.ca/>).

Приведем здесь несколько примеров таких ресурсов.

<sup>20</sup> Сайт, содержащий информацию о методах и технологиях анализа коммуникаций и о программном обеспечении преимущественно для работы с текстами, а также с аудио, видео и другими мультимедийными документами. Создан и поддерживается с 1999 г. д-ром Х. Клейном. (<http://www.textanalysis.info/>).

<sup>21</sup> Сайт, содержащий аннотированные описания ресурсов, связанных с анализом текстов и, в частности, с контент-анализом: библиография, программное обеспечение, исследовательские проекты, текстовые архивы и др. ресурсы. Создан и поддерживается д-ром М. Ромпелем (<http://www.content-analysis.de/>).

<sup>22</sup> Сайт, представляющий собой онлайн-приложение к известной книге проф. К. Нойендорфа<sup>34</sup>. Поддерживается Кливлендским университетом и содержит много информации о контент-аналитических исследованиях, программном обеспечении, тематических веб-сайтах и других ресурсах, полезных для исследователей, работающих в этой области ([http://academic.csuohio.edu/neuendorf\\_ka/content/cata.html](http://academic.csuohio.edu/neuendorf_ka/content/cata.html)).

34. Neuendorf, K. *The Content Analysis Guidebook*. Sage Publications, 2002.

<sup>23</sup> Сеть «The CAQDAS<sup>35</sup> Networking Project», созданная в 1994 г. одним из исследовательских центров департамента социологии университета Суррея (Великобритания) при поддержке ESRC. Сеть предоставляет информацию и обеспечивает поддержку исследований в области качественного анализа данных, а также онлайн курсы по наиболее популярным программам. На сайте проекта дается аннотированный список таких программ; наибольшее внимание уделяется интеграции качественных и количественных данных, анализу визуальных данных и расширению возможностей ГИС для работы с методами качественного анализа

35. Аббревиатура CAQDAS (Computer Assisted Qualitative Analysis) была предложена Р. Ли и Н. Филдингом в 1989 г. на конференции, собравшей исследователей, начавших развивать методологию, технологию и программное обеспечение этой научной области.

(<http://www.surrey.ac.uk/sociology/research/researchcentres/caqdas/about/index.htm>).

<sup>24</sup> Сайт «Online QDA Project» на сайте университета Хаддерсфилд (Великобритания) предлагает пользователям обучающие материалы по качественному анализу данных (QDA) и, в частности, по программному обеспечению компьютеризованного качественного анализа (CAQDAS), поддерживая тесное сотрудничество с CAQDAS Networking project. Кроме того, на сайте много ссылок на онлайн-ресурсы: книги, журналы, в том числе электронные, статьи, видеоматериалы, сетевые форумы и т. п. На сайте могут представлять свои материалы и члены научного сообщества, занимающиеся качественным анализом данных (<http://onlineqda.hud.ac.uk/>).

<sup>25</sup> Раздел сайта ReStore, финансируемого Национальным Советом Великобритании по социально-экономическим исследованиям (ESRC), который связан с проектом развития инновационных исследовательских методов контент-анализа медиа, в основном, новостных. Содержит информацию о различных ресурсах в этой области: тематических сайтах, методологических и прикладных разработках, программном обеспечении и т. д. (<http://www.restore.ac.uk/lboro/>).

<sup>26</sup> Последний ресурс — это библиографическая публикация Alexa M., Zuell C. *Text Analysis Software: Commonalities, Differences and Limitations: The Results of a Review // Quality and Quantity*, August 2000. Vol. 34, No. 3. Pp. 299—321.

<sup>27</sup> На основе информации упомянутых сайтов нами в 2010—2014 гг. были протестированы 79 программ, включая и самые простые, и наиболее мощные, с целью возможности их использования для работы с русскоязычными текстами и бесплатного доступа. В результате была выбрана программа MAXQDA немецкой фирмы VERBI. Программа относится к классу качественного анализа, но в равной степени эффективно работает с количественным анализом текстов, т. е. предоставляет функции комбинированного анализа (mixed methods). Достоинством программы является возможность обмена файлами с

офисными программами, в первую очередь — с табличным процессором MS Excel, не только как с инструментом визуализации результатов, но и как с универсальным контейнером табличной информации, а через него — со статистическими пакетами (SPSS, Statistica и др.). Собственные аналитические и графические возможности программы позволяют выполнять все основные функции работы с текстами: от отбора документов по различным критериям, формирования системы категорий и индикаторов и разметки текстов в соответствии с этой системой.

<sup>28</sup> Программа MAXQDA дает возможность работать с данными разных типов, анализировать их вместе или отдельно, как разные группы документов; разные исследователи могут независимо друг от друга анализировать одни и те же данные, используя различные схемы кодирования и обмениваться информацией и результатами; различные схемы кодирования позволяют сравнивать интерпретации, основанные на разных теоретических подходах, лежащих в их основе. Наконец, что очень важно, первичные, качественные по своей природе данные могут быть в ходе анализа визуализированы или трансформированы во вторичные, количественные данные, которые легко экспортировать в статические программы для дальнейшего углубленного анализа.

<sup>29</sup> Констатируя большие возможности современных компьютерных программ, базирующихся на парадигме смешанных, качественно-количественных методов анализа текстов, обратимся к характеристике контент-аналитических исследований, выполненных историками, начиная со второй половины 1980-х гг. и не отраженных в уже упомянутых историографических обзорах середины 1980-х гг. В обзоре работ, помимо описания исследовательской проблемы и результатов, мы будем обращать внимание на характер источниковой база исследования, принципы формирования системы категорий анализа, отбор объектов исследования, программное обеспечение для анализа текстов и статистической обработки.

<sup>30</sup> Отметим, что в 1990-е гг. в работах членов Ассоциации «История и компьютер» с текстами стала активно использоваться одна из популярных тогда программ — TACT, разработанная в конце 1980-х гг. в университете Торонто и переданная нам коллегами из международной Ассоциации History and Computing. Эта программа использовалась в контент-анализе стенограмм I Государственной Думы 1906 г. и постсоветских Дум, программных документов политических партий, материалов прессы. Прежде всего, это коллективная монография «Становление российского парламентаризма», где глава 6 полностью посвящена изучению стенографических отчетов заседаний Думы с помощью метода контент-анализа с помощью программы TACT<sup>36</sup>. Для анализа стенограмм, сканированных, распознанных и переведенных в формат полнотекстовой базы данных, была создана система категорий, соответствующих основным положениям программ партий и обсуждавшимся в Думе законопроектам. Изложению результатов анализа предшествовало подробное изложение методики работы, частично вошедшее в главу учебника по исторической информатике (глава 4, параграф 5 — «Анализ текста с помощью системы TACT»)<sup>37</sup>. Данное исследование ввело в научный оборот новые методы анализа делопроизводственных материалов, отражающих деятельность первого российского парламента, а результаты контент-анализа думских дебатов позволили авторам получить интересные выводы о политическом поведении депутатов Думы, складывании процедур работы парламента, механизма принятия решений российской законодательной элитой.

<sup>31</sup> Методические принципы анализа стенограмм, апробированные в рассмотренной выше работе по российскому парламентаризму, были в дальнейшем использованы в ряде работ А. Л. Кобринского, изучавшего Государственные Думы постсоветской России. На материалах хранящихся в

36. Селунская Н. Б., Бородин Л. И., Григорьева Ю. Г., Петров А. Н. Становление российского парламентаризма. М., 1996. С. 163—213.

37. Историческая информатика / под ред. Л. И. Бородина и И. М. Гарсковой. М., 1996.

38. См., например: Кобринский А. Л. Проблемы государственного строительства в Российской Федерации (по материалам стенограмм пленарных заседаний Государственной Думы 1993—

электронном виде стенографических отчетов Государственной Думы РФ первого созыва (1993—1995 гг.) с использованием программы ТАСТ автор рассмотрел дебаты по вопросам государственного строительства РФ: какой путь развития Российской Федерации как государства виделся Государственной Думе наиболее оптимальным в новых политических условиях: унитарное государство, конфедерация или федерация (конституционная, договорная или конституционно-договорная). На той же источниковой базе — стенограммах пленарных заседаний Государственной Думы первого созыва в электронной версии А. Л. Кобринский занимался изучением работы фракций, депутатских групп, отдельных депутатов, выявляя их позиции и взгляды по актуальным проблемам внутри- и внешнеполитической жизни страны и уделяя первоочередное внимание выступлениям руководителей основных фракций. Результаты этой работы отражены в серии публикаций<sup>38</sup>.

1995 гг.). М., 2001; Кобринский А. Л. Лидеры ЛДПР, КПРФ и Яблока в Государственной Думе 1995—1999. М., 2003.

32 Методика анализа работы со стенографическими парламентскими отчетами, апробированная предшественниками, была использована в исследованиях, предпринятых в Пермском университете в рамках проекта «Информационная система "Стенографические отчеты Государственной Думы (1906—1917)"». С помощью программы ТАСТ на материалах стенографических отчетов Государственной Думы Российской империи первого — четвертого созывов изучались вопросы, связанные с взаимодействием Думы и правительства, фракционными установками и их отражением в таком виде парламентской деятельности, как запросы депутатов, ответы правительства на запросы, их обсуждения и т. д.<sup>39</sup>. Контент-анализ использовался для выяснения тематики, структуры и основных функций запросов, активности отдельных депутатов и думских фракций в запросной деятельности. Основные семантические категории формулировались автором для каждого из четырех созывов на основе наиболее часто встречающихся словоформ из словаря текста. Была выявлена основная функция депутатских запросов — контроль за деятельностью правительства и обнаружены категории, которые присутствовали в запросной деятельности всех созывов: это запросы, связанные с нарушением законодательства, и запросы, связанные с нарушениями при издании того или иного закона или постановления.

39. Рябухин И. В. Контент-анализ запросной деятельности в Государственной Думе Российской империи, 1906—1912 гг. // Информационный бюллетень Ассоциации «История и компьютер». 2010. № 36. С. 118—119.

33 На базе упомянутого проекта «Информационная система "Стенографические отчеты Государственной Думы (1906—1917)"» проводился контент-анализ выступлений депутатов<sup>40</sup>. Данное исследование имеет и содержательный, и методический компонент. Формирование списка категорий проводилось автором по двум сценариям: со свободным словарем и с фиксированным словарем. В обоих сценариях проводился частотный анализ отдельных категорий и анализ из взаимосвязей на основе частот совместной встречаемости индикаторов категорий (лексем теста) в пределах определенного контекста, размер которого может варьировать. Далее матрица совместной встречаемости обрабатывалась с помощью инструментов сетевого анализа в программе UCINET (разработка Analytic Technologies), с вычислением для каждой категории (узла сети) такого показателя как *betweenness*, (показатель промежуточности, или посредничества), который равен суммарному количеству кратчайших путей (связей) между всеми категориями, проходящими на графе через данную вершину.

40. Сметанин А. В. Семантический контент-анализ выступлений депутатов Государственной Думы Российской Империи: методологические аспекты // Вестник Пермского университета. Серия «История». 2014. № 3 (26). С. 57—66.

34 Первый вариант сценария использовался автором для анализа позиций конкретных депутатов с целью сопоставления явных (имеющих высокие частоты) и скрытых категорий, которые могут иметь невысокие частоты, но высокие значения показателя *betweenness*. Например, анализ показал, что депутат от Дагестанской области И. Гайдаров, который в III Думе входил во фракцию социал-демократов, на уровне явных категорий выражал левые взгляды, типичные для социал-демократов, тогда как на латентном уровне — идеи, характерные для членов Мусульманской группы, для которой наиболее важными были национальные вопросы<sup>41</sup>.

41. Там же. С. 60—61.

- 35 Второй сценарий (работа с фиксированным словарем категорий) автор использует для выявления различий между членами одной фракции. Например, левое крыло кадетской фракции в 1912—1914 гг. представлял депутат Н. В. Некрасов, а правое — В. А. Маклакова. Для сопоставления речей обоих депутатов была сформирована система категорий, отражавшая основные положения Программы кадетской партии. Оказалось, что относительные частоты встречаемости многих категорий у обоих депутатов существенно различаются, что позволило автору выявить наиболее важное противоречие во взглядах двух депутатов: противопоставление народа и государства, было связано с конфликтом двух идей программы партии: революционной идеи народного освобождения и реформистской идеи правового государства<sup>42</sup>.
- 36 В целом, предлагаемая методика представляется перспективной для анализа выступлений политиков не только в стенах парламента.
- 37 Упомянутая в обзорах 1980-х гг. работа О. Г. Буховца<sup>43</sup> была успешно продолжена в серии следующих статей и завершилась защитой докторской диссертации и публикацией монографии<sup>44</sup>. Автор применял контент-анализ для изучения крестьянского менталитета на основе текстов приговоров, принимавшихся на крестьянских сходах в Самарской и Воронежской губерниях в период 1905—1907 гг. и направлявшихся в правительственные инстанции, Государственную Думу, редакции газет и т. д. Анализ разработанной системы категорий (суждений, требований, лозунгов), содержащихся в текстах этих документов, с помощью подсчета коэффициентов ранговой корреляции позволил сравнивать структуру «приговорного» движения в разных губерниях, а построение таблиц сопряженности и подсчет коэффициентов взаимосвязи Чупрова дал возможность выделить несколько групп тесно взаимосвязанных требований, которые были проинтерпретированы как обобщенные факторы общественного сознания крестьянства.
- 38 Аналогичное исследование (как по источниковой базе, так и по методике работы) было проведено на материалах наказов крестьян Екатеринославской, Таврической и Херсонской губерний<sup>45</sup>. Результаты этого исследования были использованы для сравнения с выводами О. Г. Буховца, для чего автору пришлось адаптировать набор содержательных категорий, отказавшись от части использованных в работе О. Г. Буховца и включив в анализ ряд новых категорий.
- 39 Отметим, что ряд исследований, связанных с анализом текстов, был ориентирован не столько на получение конкретных результатов, сколько на обсуждение теоретико-методологических проблем такого анализа. В первую очередь, это проблемы формализации (категоризации, концептуализации, кодирования), то есть построения системы семантических категорий, описывающих содержание текста. Разнообразные публикации на эту тему в основном обсуждают два базовых подхода: категоризация «сверху вниз», на основе фиксированного словаря категорий, основанного на внеисточниковом знании<sup>46</sup> и категоризация «снизу вверх», на основе свободного словаря всех слов текста и обобщения слов в категории на основе частотных данных с помощью методов статистики. Решение именно этих проблем, во многом, различает количественный и качественный подходы к анализу неструктурированных текстовых данных.
- 40 Так, в работе В. И. Тихонова<sup>47</sup> был предложен морфологический метод построения системы категорий через определение групп связанных словоформ текста, представленного как совокупность отдельных фраз в формате источник-ориентированной базы данных, созданной в СУБД KLEIO по материалам личных дел лишенцев, ходатайствовавших на рубеже 1920—1930-х гг. о восстановлении в избирательных правах. Работа была выполнена в рамках источник-ориентированного подхода, который, как упоминалось выше,
42. Там же. С. 64.
43. Буховец О. Г. К методике изучения «приговорного» движения и его роли в борьбе крестьянства в 1905—1907 годах // История СССР. 1979. № 3.
44. Буховец О. Г. Массовые источники по общественному сознанию российского крестьянства (Опыт применения контент-анализа при изучении приговоров и наказов 1905—1907 гг.) // История СССР. 1986. № 4; Буховец О. Г. Математика в исследовании общественного сознания: крестьянские приговоры и указы 1905—1907 гг. // Число и мысль. Вып. 9. М., 1986. С. 32—51; Буховец О. Г. Социальные конфликты и крестьянская ментальность в Российской Империи начала XX века: новые материалы, методы, результаты. М., 1996.
45. Топка Р. В. Контент-анализ наказов крестьян Украины в I Государственную Думу России // Информационный бюллетень Ассоциации «История и компьютер». 1997. № 21. С. 122—125.
46. При всей ограниченности подходов с использованием фиксированного словаря, следует признать, что этот метод имеет свои достоинства, если в качестве словаря использовать тезаурус соответствующей предметной области, что позволяет снизить уровень субъективности исследования на этапе категоризации.
47. Тихонов В. И. Проблемы категоризации при контент-анализе // Круг идей: модели и технологии исторической информатики. М., 1996. С. 278—289.

наиболее близок концепции качественного анализа текстов.

- 41 Задание порога (30) частоты встречаемости и грамматическая «нормализация» позволили автору сформировать список из 215 словоформ, связи между которыми были подсчитаны с помощью таблиц сопряженности и направленного коэффициента неопределенности<sup>48</sup>, позволяющего вычислять связи между этими словоформами. Полученные группы словоформ интерпретировались как категории контент-анализа, которые включили примерно 2/3 всего объема текста. Экспертная проверка полноты построенной системы категорий показала, что метод не обнаружил около 18 % категорий, что связано с выбором порога частоты встречаемости словоформ. Разработанный подход к категоризации использовался автором в следующих работах с личными делами лишенцев и с применением различных статистических методов<sup>49</sup>.
- 42 Работа С. Ф. Гребениченко также посвящена методологическим и методическим вопросам категоризации текстовых документов<sup>50</sup>. Автор стремился разработать автоматизированную процедуру объединения первичных индикаторов текста в смысловые категории при работе с комплексом актов документов за 1920-е гг. Процедура, предложенная в данной работе, как и методика В. И. Тихонова, ориентирована на выявление устойчивых словосочетаний, но имеет иерархическую организацию: с использованием словаря текста и списка словосочетаний (объектов анализа) строится таблица сопряженности и подсчитываются коэффициенты корреляции между всеми парами слов<sup>51</sup>. На основе матрицы корреляции проводится факторный анализ, и полученные факторы интерпретируются как смысловые категории на уровне словосочетаний. На следующем этапе в качестве объектов выступают предложения, а в качестве признаков — выявленные на предыдущем этапе факторы (словосочетания); вновь рассчитывается матрица корреляции, проводится факторный анализ и проводится интерпретация факторов более высокого уровня, которыми теперь являются предложения. На последнем, третьем, этапе процедура повторяется на уровне абзацев текста как элементов высшего смыслового уровня анализируемого текста. Полученные факторы рассматриваются как финальные обобщенные категории текста документа. С точки зрения автора, предложенная методика воспроизводит логику гуманитарного восприятия текста. Применение предложенной методики в конкретно-историческом приложении изложено в другой статье С. Ф. Гребениченко<sup>52</sup>.
- 43 Еще одним направлением контент-аналитических исследований историков является работа с письмами. Здесь можно назвать работы Д. Х. Ибрагимовой<sup>53</sup>, А. М. Маркевича<sup>54</sup>, О. С. и С. В. Поршневых<sup>55</sup>.
48. Направленный или асимметричный коэффициент неопределенности (Asymmetric Uncertainty Coefficient), имеющий теоретико-информационную природу, позволяет оценить среднее уменьшение неопределенности распределения значений одной номинальной переменной, если известны значения другой.
49. Тихонов В. И. Кто достоин избирательных прав? Типология и контент-анализ заявлений лишенцев // Круг идей: макро- и микроподходы в исторической информатике. Труды V конференции Ассоциации «История и компьютер». М.—Минск, 1998. С. 140—177.
50. Гребениченко С. Ф. Как повысить объективность контент-анализа? // Информационный бюллетень Ассоциации «История и компьютер». 1995. № 14. С. 114—117.
51. Мы не рассматриваем в данном случае корректность применения корреляционного анализа к альтернативным признакам (наличие / отсутствие) слова в словосочетании.
52. Гребениченко С. Ф. Власть, циркуль, село (Концептуальное обращение к законам и нормативным актам 1920-х гг. на принципах базы данных) // Круг идей: развитие исторической информатики. М., 1995.
53. Ибрагимова Д. Х. Перестройка, рыночные ориентиры, сельский менталитет: многомерный анализ массовой корреспонденции в центральные органы печати // Круг идей: модели и технологии исторической информатики. Труды III конференции Ассоциации «История и компьютер». М., 1996. С. 230—254.
54. Маркевич А. М. Солдатские письма во ВЦИК и Петросовет в 1917 г. К методике анализа // Круг идей: историческая информатика на пороге XXI века / Отв. ред. Л. И. Бородин, Ю. П. Смирнов и И. Ф. Юшин. М.—Чебоксары, 1999. С. 151—173.
55. Поршнева О. С., Поршнев С. В. К характеристике менталитета народных масс России: революция 1917 г. в фокусе массового сознания (опыт статистического анализа писем рабочих, крестьян и солдат в центральные органы Советов рабочих и солдатских депутатов) // Круг идей: историческая информатика на пороге XXI века. Труды VI конференции Ассоциации «История и компьютер». М.—Чебоксары, 1999. С. 119—150.

- <sup>44</sup> Д. Х. Ибрагимова поставила проблему выявления «структурной организации» сознания сельского населения в условиях экономических преобразований на рубеже 1990-х гг., а также факторов, ее детерминировавших. Источниковой основой изучения проблемы послужил массив неопубликованных писем жителей села в редакцию газеты «Сельская жизнь» с декабря 1989 г. по август 1990 г. Это несколько тысяч документов в архиве редакции. Для анализа из нескольких тысяч писем, содержание которых касается социально-экономических проблем на селе в условиях перехода к рынку, была сделана выборка. Письма изучались как с точки зрения анкетных характеристик авторов, так и с точки зрения их содержания, для чего был использован контент-анализ. Неструктурированный текст источника был представлен набором первичных смысловых единиц (индикаторов), сформированных на основе словаря текста, затем, путем агрегирования индикаторов было получено 111 категорий. На основе корреляционного анализа взаимосвязей категорий в текстах писем было показано, что сознание сельского населения в изучаемый период не было монолитным — по своей внутренней структуре менталитет сельского общества представлял собой «слоеный пирог»; были выделены шесть основных «пластов», представляющих собой комплексы взаимосвязей смысловых категорий: от «антиэтатистского» до «индивидуалистского». Затем с помощью факторного анализа (использовался метод главных компонент) было показано, что «многослойность» сельского менталитета определялась совокупностью нескольких главных факторов, таких, как долговременные традиции, связанные с общиной; недавнее историческое прошлое; новые идеи, которые проводились через официальные заявления властей и средства массовой информации.
- <sup>45</sup> Работа А. М. Маркевича также посвящена анализу источников личного происхождения — это солдатские письма (как правило, коллективные) в Петросовет и ВЦИК в 1917 г. Автор предпринял попытку выработать комплексную методику анализа этого весьма сложного источника. Поскольку количество сохранившихся солдатских писем в Петросовет и ВЦИК исчисляется несколькими тысячами, для апробации методики была сделана пробная случайная выборка. Методика автора включала изучение структуры писем и их содержания. Для изучения структуры использовалось представление каждого письма как набора определенных высказываний, реализующих одну или несколько целевых функций, и последующая группировка с помощью кластерного анализа, в результате чего было получено пять основных типов солдатских писем.
- <sup>46</sup> Для изучения содержания был использован контент-анализ: было выделено 25 тем-категорий, фиксирующих основные вопросы, затронутые в высказываниях. Частотный анализ их взаимной встречаемости выявил основные группы проблем, которым были посвящены солдатские письма: вопросы войны и мира; представления солдат о новой ситуации в тылу и о революционной России; процессы, происходившие в армии после революции; повседневная служба и солдатский быт; просьбы солдат об отпуске. Попытка связать основные проблемы, поднимаемые в письмах, с выявленными типами писем, в данном исследовании оказалась неудачной, поэтому автор сделал вывод о том, в общем случае содержание письма не определяет его структуру и цели, и наоборот. В целом, удалось выявить как достоинства, так и ограничения предложенной методики, которая, возможно, получит развитие в других исследованиях.
- <sup>47</sup> К исследованию А. М. Маркевича достаточно близка работа О. С. и С. В. Поршневых, источниковой базой которой стали письма в газету «Известия». Авторы поставили задачу реконструкции ряда важных аспектов содержания и структуры массового сознания рабочих, крестьян и солдат в период революции 1917 г., их важнейших экономических, социальных, политических предпочтений, идеалов и ценностей в динамике за весну, лето и осень. Авторами была изучена совокупность из 330 писем, которая рассматривалась как естественная

выборка. Экспертным путем были выделены 238 смысловых единиц — показателей в форме высказываний авторов писем по социальным, политическим, экономическим, культурным проблемам. Агрегирование этих показателей составило 64 смысловые категории. С помощью корреляционного анализа были подсчитаны коэффициенты связи смысловых категорий и выделены четыре основных «пласта» в массовом сознании: традиционалистский, революционно-оборонческий и социалистический. Была выявлена эволюция: весной—летом 1917 г. массовое сознание «низов» радикализируется в направлении общинно-уравнительных установок социальной справедливости, равенства.

48 Как и другие исследователи массового сознания, авторы данной работы отмечают высокую долю «диффузного» сознания, поскольку писем с категориями, относящимися более чем к одному пласту сознания, оказалось около половины. При этом, наряду с общей характеристикой менталитета, авторам удалось выявить существенные качественные отличия у различных социальных групп. Например, в письмах рабочих доминировало выражение поддержки Советам, большевикам, протест против их преследований; в целом позиция рабочих была довольно радикальна. В крестьянском сознании преобладали главным образом традиционалистские ценности; в первых рядах стоят призывы к заключению скорейшего мира с Германией и проведению аграрных преобразований в стране; сильны антигородские и антирабочие настроения. Для солдат были характерны стремление к миру, ненависть к буржуазии, революционное оборончество, призывы к мобилизации тыла, выступления против преследования большевиков.

49 Данное исследование позволило выявить преобладание в массовом сознании данных социальных групп двух пластов: радикально-антивоенного и революционно-оборонческого. Была выявлена эволюция: весной—летом 1917 г. массовое сознание «низов» радикализируется в направлении общинно-уравнительных установок социальной справедливости, равенства. Вместе с тем массовое сознание народа было достаточно противоречивым, многослойным: более трети корреспондентов сохраняли приверженность традиционным ценностям, часть из них выражала симпатии к прежним монархическим порядкам, другие, наоборот, ждали установления новой, революционной власти. Однако эти ожидания сочетались с отрицанием буржуазно-парламентской формы организации государства в пользу народовластия общинного типа.

50 Еще одним выводом авторов стало утверждение, что обострение социальных противоречий, бедствий народа в период войны вызвали к жизни социально-психологический механизм эскалации ненависти, стимулировавшей в условиях ослабления репрессивного аппарата после февраля 1917 г. акты жестокости и насилия, выработку новых идейно-ценностных установок по отношению к «классовым врагам». Элементы новой ментальности, которая станет основой многих поведенческих реакций «советского человека», зарождались под влиянием первой мировой войны и революции 1917 г.

51 Упомянем также несколько исследований, связанных с изучением социальной психологии и ментальности, в которых контент-анализ привлекался как вспомогательный инструмент формализации содержания таких источников, как письма<sup>56</sup>.

56. Марасинова Е. Н. Психология элиты российского дворянства последней трети XVIII века (по материалам переписки). М., 1999; Рафалюк О. Е. Применение контент-анализа в исторических исследованиях к источникам личного происхождения (На материале переписки русских писателей-модернистов рубежа XIX—XX вв.) // Вестник Института социологии. 2012. № 5. С. 97—108.

52 Но наибольшее число контент-аналитических исследований историков посвящено анализу прессы. В отличие от предыдущего периода в этих исследованиях проводится не «ручная», а автоматизированная обработка и анализ текстов газетных статей. Не последнюю роль в этом процессе играет

возможность использования обширных текстовых архивов, хранящих статьи не только электронных, но и традиционных («бумажных») СМИ. В Интернете можно найти множество каталогов таких архивов.

53 Продолжая традиции предшествующего периода, в 1990-е гг. контент-анализ наиболее часто применяется к изучению газетных материалов. В новых работах сочетаются приемы традиционного и компьютеризованного анализа. В отсутствие значительных электронных коллекций газетных статей исследователи по-прежнему прибегают к выборочному методу, но, с другой стороны, расширение возможностей применения компьютеризованного анализа проявляется в том, что авторы новых работ не ограничиваются частотным анализом, но переходят от дескриптивной статистики к использованию более сложных инструментов анализа взаимосвязей семантических категорий.

54 В работе А. Г. Голикова<sup>57</sup> для анализа отражения процесса монополизации российской промышленности на страницах газет сначала в несколько этапов проводится источниковедческий анализ и формируется представительный информационный массив сведений, позволяющий разработать систему семантических категорий для совокупности издания периодической печати. Затем проводится анализ частот встречаемости выделенных категорий и выявляются их взаимосвязи. Для оценки степени их взаимосвязи используется коэффициент сопряженности с пороговым значением, позволяющим выявлять статистически значимые связи на уровне доверительной вероятности 95 %, что дает возможность найти и интерпретировать наиболее существенные тенденции, нашедшие отражение на страницах российской прессы.

57. Голиков А. Г. Тенденции монополизации российской промышленности 1914—1917 гг. (опыт контент-анализа публикаций отечественных газет) // Круг идей: развитие исторической информатики. М., 1995. С. 184—198.

55 Контент-анализ содержания статей газеты «Правда» был предпринят О. Н. Блиновой для изучения отражения в газете экономической концепции КПРФ в 1993—1999 гг.<sup>58</sup> Это исследование было построено на принципах компьютеризованного анализа источника, представленного в традиционном бумажном формате, поэтому перед автором стояли задачи создания выборки статей, их сканирования и распознавания для последующего анализа оцифрованных текстов. Экспертным путем были сформулированы базовые вопросы экономического характера и на их основе сформированы категории: собственность, приватизация, бюджет, кризис, и социалистическая экономика, капиталистическая экономика, китайская модель экономики, инвестиции, инфляция, внешняя торговля, внешний долг, банки, налоги. Частотный анализ, проведенный с помощью программы TACT выявил наиболее острые проблемы: приватизация, собственность, федеральный бюджет, развитие банковской системы и земельный вопрос.

58. Блинова О. Н. Экономическая концепция КПРФ в 1993—1999 гг. в отражении газеты «Правда»: опыт применения контент-анализа // Круг идей: электронные ресурсы исторической информатики. М., 2003. С. 371—408.

56 Изучение динамики частот встречаемости показало, что на протяжении изучаемого периода отношение «Правды» к большинству проблем становится более утилитарным и менее эмоциональным; например, в 1993 г. она резко критикует проводимую приватизацию, в 1994 г. пропагандирует идею ее немедленного прекращения, однако в 1995 и 1996 гг. на страницах газеты активно обсуждаются итоги приватизации для экономики России, а в 1998 г. газета уже сетует на малые объемы финансовых поступлений в федеральный бюджет от приватизируемых объектов.

57 В конце 2000-х гг. историки начинают использовать для работы с материалами СМИ современные программы качественного контент-анализа. В частности, идет освоение уже упомянутой программы MAXQDA, которая интегрирует качественные методы формализации текстов с количественными процедурами классического контент-анализа.

58 Впервые это программное обеспечение было апробировано И. М. Гарсковой при анализе историографии исторической информатики<sup>59</sup>. Характеристика возможностей программы на примере контент-анализа полнотекстовой базы материалов конференций АИК<sup>60</sup> показали, что программа является полезным

59. Гарскова И. М. Тематический анализ историографии АИК // Информационный бюллетень Ассоциации «История и компьютер». № 35, декабрь 2008 г. С. 8—10; Гарскова И. М. Анализ историографии

инструментом работы с большими массивами неструктурированных нарративных источников. С помощью контент-анализа историографии можно на коротких временных интервалах проследить смену приоритетов в проблематике, методах и технологиях исследований, представительства региональных групп и школ, выявить «точки роста», которые в результате формируют основные тенденции в рамках более продолжительных периодов.

исторической информатики как научного направления // Харківський Історіографічний збірник. Х., 2010. Вип. 10. С. 138—171; Гарскова И. М. Библиометрический и сетевой анализ историографии // Информационный бюллетень Ассоциации «История и компьютер». № 37. Труды международной конференции «Компьютерные технологии и математические методы в исторических исследованиях» (Петрозаводск, 11—16 июля 2011 г.). Петрозаводск, 2011. С. 39—48; Гарскова И. М. Историография научного направления: возможности количественного анализа // Вестник РГГУ. Серия «Документалистика. Документоведение. Архивоведение». 2011. № 18 (80). С. 88—100.

60. Гарскова И. М. Историческая информатика как профессиональное сообщество: контент-анализ полнотекстовой базы по материалам конференций 2000—2010 гг. // Историческая информатика. 2012. № 1. С. 44—53.

59 И. В. Митрофанов, изучая образ России в итальянской прессе конца 1990-х гг.,<sup>61</sup> создал в программе MAXQDA полнотекстовую базы данных по материалам более 6 тыс. статей, извлеченных их электронных архивов трех ведущих итальянских газет, и получил ряд интересных выводов, в частности, о влиянии политической ориентации выбранных газет на освещение вопросов, связанных с Россией. Гипотеза о широком применении в итальянской публицистике стереотипов времен «холодной войны», сформулированная на основе эмпирических данных, не получила подтверждения. Среди общих тенденции было отмечено позитивное восприятие России «правой» итальянской прессой (по сравнению с «левой» прессой), отождествление образа России с образом ее лидера.

61. Митрофанов И. В. Россия в отражении итальянской прессы в конце 1990-х гг.: контент-анализ газетных публикаций // Информационный бюллетень Ассоциации «История и компьютер». 2012. № 38. С. 95—96.

60 Исследование Н. В. Раздиной посвящено контент-анализу газеты «За индустриализацию» и выявлению изменений в редакционной политике издания за период в 1930-е гг. Начало этого исследования было связано с традиционным анализом статей, включающим трудоемкую работу по формированию представительной выборки, вводу текстов редакционных статей в память компьютера и созданию системы категорий, отражающих их содержание<sup>62</sup>. Продолжение и завершение исследования выполнено автором с помощью программы MAXQDA<sup>63</sup>.

62. Раздина Н. В. Газета «За индустриализацию» как источник для изучения промышленной политики государства в годы первых пятилеток (опыт контент-анализа редакционных статей) // Историческая информатика. 2013. № 1. С. 41—60; Раздина Н. В. Изменения в промышленной политике государства в годы первых двух пятилеток в отражении газеты «За индустриализацию» (опыт контент-анализа редакционных статей). // Исторический журнал: научные исследования. 2013. № 3. С. 259—273.

63. Раздина Н. В. Компьютеризованный анализ материалов газеты «За индустриализацию» (опыт работы с программой MAXQDA) // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42. С. 93—95.

61 В качестве выборки были взяты редакционные статьи за март и октябрь 1932 и 1937 гг. (завершающие годы первых двух пятилеток), затем были добавлены март, октябрь, ноябрь 1940 г. — последний мирный год до начала Великой Отечественной войны.

62 В работе выявлены изменения показателей важности, актуальности проблематики газетных материалов за выбранные годы. Количественный (частотный) анализ смысловых категорий статей показал изменения в содержании газеты от первой пятилетки ко второй: в годы первой пятилетки подавляющее число статей было посвящено вопросам экономики, в годы второй пятилетки — социально-политическим вопросам и вопросам идеологического воспитания, т. е. выявить скрытую информацию, не очевидную при

использовании традиционных подходов к изучению источника. Полученные результаты показали, что издание можно изучать с точки зрения двух аспектов: с одной стороны, в нем описан реальный ход индустриализации, с другой — в статьях нашли отражение основные направления идеологического воздействия на массовое сознание читателей.

63 В работе А. С. Федорук метод компьютеризованного контент-анализа используется для изучения структуры и динамики гендерной проблематики на страницах местной печати в конце XIX — начале XX вв. В качестве источника для исследования выбрана газета «Крымский Вестник»<sup>64</sup>. Эта газета, как и газета «За индустриализацию», существует только в бумажном формате, поэтому и здесь стояла задача построения выборки, в данном случае — сплошной: были полностью просмотрены статьи всех номеров газеты, вышедших за период каждой второй недели каждого четвертого месяца (апрель, август, декабрь) с 1888 по 1918 гг. и найдены те публикации, которые соответствовали тематике исследования. Оцифрованные номера газет в программе MAXQDA были преобразованы в полнотекстовую базу данных. Система категорий была сформирована на основе частотного словаря, отобранные ключевые слова (индикаторы) образовали шесть категорий («образование», «благотворительность», «общественная деятельность», «медицина», «правовая сфера», «профессиональная занятость»). Автоматизированный лексический поиск и разметка текста позволили перейти к контент-анализу который проводился отдельно для статей и объявлений, и статистической обработке его результатов. Исследование показало специфику объявлений как вида публикации: информационная насыщенность объявлений и темп роста их количества заметно выше, чем для статей. Кроме того, была выявлена не только тематическая структура публикаций в целом за весь изучаемый период, но и динамика изменений: снижение доли таких категорий, как «образование» и «благотворительность», увеличение в статьях доли категорий «медицина» и «общественная деятельность», а в объявлениях — доли категории «профессиональная занятость».

64 Следует заметить, что проблематика контент-аналитических исследований прессы постоянно расширяется. Так, в работах Е. А. Еремеевой предметной областью стал официальный юмористический дискурс на страницах журнала «Перец» в аспекте гендерного подхода<sup>65</sup> и в аспекте сравнения официального юмора на страницах советского журнала и неофициального юмора политических анекдотов<sup>66</sup>. Частотный анализ проводился с помощью обработки реляционной базы данных с описаниями статей журнала. На материалах этого же журнала и по аналогичной методике изучается официальная политика формирования в советском обществе «правильного» отношения к моде<sup>67</sup>. В работе Ж. Г. Поповой<sup>68</sup> предметом исследования также является советская официальная сатира на страницах журнала «Крокодил» и ее трансформация в условиях период усиления партийного контроля журнала в период 1920—1930-х гг. Особенностью контент-аналитического изучения иллюстрированного журнала является то, что этот метод не может быть применен к графическим изображениям. Поэтому была разработана методика анализа, базирующаяся на метаданных, описывающих содержание изображений «на языке» предварительно разработанной системы категорий, преобразованной, с одной стороны, в поля реляционной базы данных для частотного анализа, а с другой стороны — в синтезированный текст, состоящий из ключевых слов, для анализа взаимосвязей (в данном случае с помощью программы TACT).

65 Наконец, упомянем, что объектом компьютеризованного анализа текстов в работах историков становились судебные материалы<sup>69</sup>, документы политических партий<sup>70</sup>, международные отношения<sup>71</sup> и даже сеть Твиттер<sup>72</sup>. Иногда в этих работах апробируются различные программы компьютеризованного контент-анализа, хотя не всегда объясняются принципы

64. Федорук А. С. Опыт компьютеризованного контент-анализа газеты «Крымский вестник» // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42. С. 100—101.

65. Еремеева Е. А. Гендерные представления в советском обществе сквозь призму официального юмористического дискурса (опыт контент-анализа текстов журнала «Перец») // Информационный бюллетень Ассоциации «История и компьютер». 2013. № 40. С. 63—70.

66. Еремеева Е. А. Советский юмористический дискурс: контент-анализ советских политических анекдотов и журнала «Перец» // Информационный бюллетень Ассоциации «История и компьютер». 2012. № 38. С. 94—95.

67. Корниенко О. В. Сатира как инструмент формирования отношения к моде в советском обществе: количественный анализ // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42. С. 61—62.

68. Попова Ж. Г. Журнал «Крокодил» 1920—1930-х гг.: информационный потенциал и методы количественной обработки контента // Информационный бюллетень Ассоциации «История и компьютер». 2010. № 36. С. 115—116.

69. Боленко К. Г., Долгих Е. В., Самовер Н. В. Новый подход к исследованию уголовности Верховного уголовного суда над декабристами // Информационный бюллетень Ассоциации «История и компьютер». 1997. № 21. С. 101—102; Зеленина Г. Стрелецкий бунт 1698 г.: опыт контент-

выбора этих программ.

анализа следственных материалов // Информационный бюллетень Ассоциации «История и компьютер». 1998. № 23. С. 58—60. В этих работах практически еще нет контент-анализа в строгом смысле слова, однако рассматриваются проблемы обоснования формализации текстовых источников и возможной статистической обработки системы категорий анализа.

70. Блинова О. Н. Опыт изучения экономических программ политических партий России 1990-х гг.: методика применения контент-анализа на примере программ КПРФ (1993, 1995) и АПР (1994) // Круг идей: историческая информатика в информационном обществе. М., 2001. С. 212—235.

71. Горобий А. В. Количественный контент-анализ периодической печати как источника по истории отношений СССР и ФРГ 1985—1991 // Историческая информатика. 2012. № 1. С. 54—71.

72. Баранова Е. В., Юрлов А. Р. Контент-анализ твиттера современных политиков. особенности интернет-источника // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42. С. 86—87.

#### Библиография:

1. Алексеев А. Н. Контент-анализ в социологии и точки соприкосновения с другими отраслями знания // Проблемы контент-анализа в социологии: Материалы Сибирского семинара социологов. Новосибирск, 1970.
2. Алексеев А. Н. Контент-анализ: его задачи, объекты и средства // Социология культуры. Вып.1. М.—Л., 1974.
3. Алексеев А. Н. Контент-анализ: техника или методология? (К постановке проблемы) // Методологические и методические проблемы контент-анализа. Вып. 1—2. М.—Л., 1973.
4. Баранова Е. В., Юрлов А. Р. Контент-анализ твиттера современных политиков. особенности интернет-источника // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42.
5. Блинова О. Н. Опыт изучения экономических программ политических партий России 1990-х гг.: методика применения контент-анализа на примере программ КПРФ (1993, 1995) и АПР (1994) // Круг идей: историческая информатика в информационном обществе. М., 2001.
6. Блинова О. Н. Экономическая концепция КПРФ в 1993—1999 гг. в отражении газеты «Правда»: опыт применения контент-анализа // Круг идей: электронные ресурсы исторической информатики. М., 2003.
7. Боленко К. Г., Долгих Е. В., Самовер Н. В. Новый подход к исследованию деятельности Верховного уголовного суда над декабристами // Информационный бюллетень Ассоциации «История и компьютер». 1997. № 21.
8. Бородкин Л. И. Контент-анализ и проблемы изучения исторических источников // Математика в изучении средневековых повествовательных источников. М., 1986.
9. Бородкин Л. И. Многомерный статистический анализ в исторических исследованиях. М., 1986.
10. Боярский К. К., Каневский Е. А. ВЕГА — компьютерная система классификации и анализа текстов // Вестник СПбГУ ИТМО. СПб. 2009. № 5 (63).
11. Буховец О. Г. К методике изучения «приговорного» движения и его роли в борьбе крестьянства в 1905—1907 гг. // История СССР. 1979. № 3.
12. Буховец О. Г. Массовые источники по общественному сознанию российского крестьянства (Опыт применения контент-анализа при изучении приговоров и наказов 1905—1907 гг.) // История СССР. 1986. № 4.
13. Буховец О. Г. Математика в исследовании общественного сознания: крестьянские приговоры и наказания 1905—1907 гг. // Число и мысль. Вып. 9. М., 1986.
14. Буховец О. Г. Социальные конфликты и крестьянская ментальность в Российской Империи начала XX века: новые материалы, методы, результаты. М., 1996.
15. Гарскова И. М. Анализ историографии исторической информатики как научного направления // Харківський Історіографічний збірник. Х., 2010. Вип. 10.
16. Гарскова И. М. Библиометрический и сетевой анализ историографии // Информационный бюллетень Ассоциации «История и компьютер». № 37. Труды международной конференции «Компьютерные технологии и математические методы в исторических исследованиях» (Петрозаводск, 11—16 июля 2011 г.). Петрозаводск, 2011.
17. Гарскова И. М. Историография научного направления: возможности количественного анализа // Вестник РГГУ. Серия «Документалистика. Документоведение. Архивоведение». 2011. № 18 (80).
18. Гарскова И. М. Историческая информатика как профессиональное сообщество: контент-анализ

- полнотекстовой базы по материалам конференций 2000—2010 гг. // Историческая информатика. 2012. № 1.
19. Гарскова И. М. Тематический анализ историографии АИК // Информационный бюллетень Ассоциации «История и компьютер». № 35, декабрь 2008 г.
  20. Голиков А. Г. Тенденции монополизации российской промышленности 1914—1917 гг. (опыт контент-анализа публикаций отечественных газет) // Круг идей: развитие исторической информатики. М., 1995.
  21. Горобий А. В. Количественный контент-анализ периодической печати как источника по истории отношений СССР и ФРГ 1985—1991 // Историческая информатика. 2012. № 1.
  22. Гребениченко С. Ф. Власть, циркуляр, село (Концептуальное обращение к законам и нормативным актам 1920-х гг. на принципах базы данных) // Круг идей: развитие исторической информатики. М., 1995.
  23. Гребениченко С. Ф. Как повысить объективность контент-анализа? // Информационный бюллетень Ассоциации «История и компьютер». 1995. № 14.
  24. Еремеева Е. А. Гендерные представления в советском обществе сквозь призму официального юмористического дискурса (опыт контент-анализа текстов журнала «Перец») // Информационный бюллетень Ассоциации «История и компьютер». 2013. № 40.
  25. Еремеева Е. А. Советский юмористический дискурс: контент-анализ советских политических анекдотов и журнала «Перец» // Информационный бюллетень Ассоциации «История и компьютер». 2012. № 38.
  26. Зеленина Г. Стрелецкий бунт 1698 г.: опыт контент-анализа следственных материалов // Информационный бюллетень Ассоциации «История и компьютер». 1998. № 23.
  27. Ибрагимов Д. Х. Перестройка, рыночные ориентиры, сельский менталитет: многомерный анализ массовой корреспонденции в центральные органы печати // Круг идей: модели и технологии исторической информатики. Труды III конференции Ассоциации «История и компьютер». М., 1996.
  28. Историческая информатика / под ред. Л. И. Бородкина и И. М. Гарсковой. М., 1996.
  29. Каневский Е. А., Саганенко Г. И. Концептуальное обоснование компьютерного анализа массивов с текстами // Социология: 4М. 1997. № 9.
  30. Каныгин Г. В. инструментальные средства и методологические принципы анализа качественных данных // Социология: 4М. 2007. № 25.
  31. Кобринский А. Л. Лидеры ЛДПР, КПРФ и Яблока в Государственной Думе 1995—1999. М., 2003.
  32. Кобринский А. Л. Проблемы государственного строительства в Российской Федерации (по материалам стенограмм пленарных заседаний Государственной Думы 1993—1995 гг.). М., 2001.
  33. Ковальченко И. Д. Исторический источник в свете учения об информации // История СССР. 1982. № 3.
  34. Ковальченко И. Д., Бородкин Л. И. Современные методы изучения исторических источников с использованием ЭВМ. М., 1987.
  35. Корниенко О. В. Сатира как инструмент формирования отношения к моде в советском обществе: количественный анализ // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42.
  36. Круглый стол «Методологические проблемы исторической информатики и количественной истории» // Информационный бюллетень Ассоциации «История и компьютер». 1996. № 19.
  37. Марасинова Е. Н. Психология элиты российского дворянства последней трети XVIII века (по материалам переписки). М., 1999.
  38. Маркевич А. М. Солдатские письма во ВЦИК и Петросовет в 1917 г. К методике анализа // Круг идей: историческая информатика на пороге XXI века / отв. ред. Л. И. Бородкин, Ю. П. Смирнов и И. Ф. Юшин. М.—Чебоксары, 1999.
  39. Маслова О. М., Пугачева М. Г. Виды документальных источников и методы их анализа // Методы сбора информации в социологических исследованиях / отв. ред. В. Г. Андреенков и О. М. Маслова. М., 1998. Кн. 2.
  40. Методологические и методические проблемы контент-анализа. М.—Л., 1973.
  41. Миронов Б. Н. Историк и социология. Л., 1984.
  42. Митрофанов И. В. Россия в отражении итальянской прессы в конце 1990-х гг.: контент-анализ газетных публикаций // Информационный бюллетень Ассоциации «История и компьютер». 2012. № 38.
  43. Олейник А. Н. Триангуляция в контент-анализе. вопросы методологии и эмпирическая проверка // Социологические исследования. 2009. № 2.
  44. Петров А. Н. Компьютерный анализ текста: историография метода. // Круг идей: модели и технологии исторической информатики. М., 1996.
  45. Попова Ж. Г. Журнал «Крокодил» 1920—1930-х гг.: информационный потенциал и методы количественной обработки контента // Информационный бюллетень Ассоциации «История и компьютер». 2010. № 36.
  46. Поршнева О. С., Поршнев С. В. К характеристике менталитета народных масс России: революция 1917 г. в фокусе массового сознания (опыт статистического анализа писем рабочих, крестьян и солдат в центральные органы Советов рабочих и солдатских депутатов) // Круг идей: историческая информатика на пороге XXI века. Труды VI конференции Ассоциации «История и компьютер». М.—Чебоксары, 1999.
  47. Проблемы контент-анализа в социологии: Материалы Сибирского семинара социологов. Новосибирск, 1970.
  48. Раздина Н. В. Газета «За индустриализацию» как источник для изучения промышленной политики государства в годы первых пятилеток (опыт контент-анализа редакционных статей) // Историческая информатика. 2013. № 1.
  49. Раздина Н. В. Изменения в промышленной политике государства в годы первых двух пятилеток в отражении газеты «За индустриализацию» (опыт контент-анализа редакционных статей) // Исторический журнал:

- научные исследования. 2013. № 3.
50. Раздина Н. В. Компьютеризованный анализ материалов газеты «За индустриализацию» (опыт работы с программой MAXQDA) // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42.
  51. Рафалюк О. Е. Применение контент-анализа в исторических исследованиях к источникам личного происхождения (На материале переписки русских писателей-модернистов рубежа XIX—XX вв.) // Вестник Института социологии. 2012. № 5.
  52. Рябухин И. В. Контент-анализ запросной деятельности в Государственной Думе Российской империи, 1906—1912 гг. // Информационный бюллетень Ассоциации «История и компьютер». 2010. № 36.
  53. Селунская Н. Б., Бородкин Л. И., Григорьева Ю. Г., Петров А. Н. Становление российского парламентаризма. М., 1996.
  54. Семенова В. В. Качественные методы: введение в гуманистическую социологию: Учеб. пособие для студентов вузов / Ин-т социологии РАН. М., 1998.
  55. Сметанин А. В. Семантический контент-анализ выступлений депутатов Государственной Думы Российской Империи: методологические аспекты // Вестник Пермского университета. Серия «История». 2014. № 3 (26).
  56. Тихонов В. И. Кто достоин избирательных прав? Типология и контент-анализ заявлений лишенцев // Круг идей: макро- и микроподходы в исторической информатике. Труды V конференции Ассоциации «История и компьютер». М.—Минск, 1998.
  57. Тихонов В. И. Проблемы категоризации при контент-анализе // Круг идей: модели и технологии исторической информатики. М., 1996.
  58. Толстова Ю. Н., Масленников Е. В. Качественная и количественная стратегии. Эмпирическое исследование как измерение в широком смысле // Социс. 2000. № 10.
  59. Топка Р. В. Контент-анализ наказов крестьян Украины в Государственную Думу России // Информационный бюллетень Ассоциации «История и компьютер». 1997. № 21.
  60. Федорук А. С. Опыт компьютеризованного контент-анализа газеты «Крымский вестник» // Информационный бюллетень Ассоциации «История и компьютер». 2014. № 42.
  61. Creswell J. Research Design. Qualitative, quantitative and mixed methods approaches. Sage Publications, 2003.
  62. Denzin N. The research act. New York, 1978.
  63. Iker H. P., Harway N. J. Computer Systems Approach toward the Recognition and Analysis of Content // The Analysis of Communication Content / ed. by G. Gerbner, O. R. Holsti, K. Krippendorf, W. J. Paisley, Ph. J. Stone. New York, 1969.
  64. Kelle U. Die Integration qualitativer und quantitativer Methoden in der empirischen Sozialforschung. Theoretische Grundlagen und methodologische Konzepte. Wiesbaden, 2007.
  65. Kracauer S. The Challenge of Qualitative Content Analysis. Public Opinion Quarterly, 1952, № 16.
  66. Krippendorf K. Content Analysis. An Introduction to its Methodology. London: Sage Publications, 2004. 2nd edition.
  67. Kuckartz U. Qualitative Text Analysis. A Guide to Methods, Practice and Using Software. Sage Publications, 2014.
  68. Kuckartz U. Realizing Mixed-Methods Approaches with MAXQDA [Electronic resource]. Mode of Access: <http://www.maxqda.com/download/MixMethMAXQDA-Nov01-2010.pdf>
  69. Mayring P. Qualitative Content Analysis // Forum: Qualitative Social Research [Electronic resource]. Volume 1. No. 2, Art. 20. June 2000. Mode of Access: <http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385>
  70. Mayring Ph. Qualitative Inhaltsanalyse. Grundlagen und Techniken (1st ed.). Weinheim: Beltz, 1983.
  71. Neuendorf, K. The Content Analysis Guidebook. Sage Publications, 2002.
  72. Stone P. J., Dunphy D. C., Dexter C., Smith M. S., Ogilvie D. M. The General Inquirer: A Computer Approach to Content Analysis. Cambridge, 1966.
  73. Teddlie Ch., Tashakkori A. Major issues and controversies in the use of mixed methods in the social and behavioral sciences. Sage Publications, 2003.

Ссылка для цитирования:

Гарскова И. М. Новые тенденции в компьютеризованном анализе текстов: концепции, методы, технологии // Электронный научно-образовательный журнал «История», 2015. Т.6. Выпуск 8 (41) [Электронный ресурс]. Доступ для зарегистрированных пользователей. URL: <http://history.jes.su/s207987840001255-9-1> (дата обращения: 04.11.2015). DOI: 10.18254/S0001255-9-1