

Олешкова Анна Михайловна

кандидат исторических наук,  
доцент кафедры гуманитарных  
и социально-экономических наук  
Нижегородского государственного  
социально-педагогического института  
филиала Российского государственного  
профессионально-педагогического университета

## ПЕРСПЕКТИВЫ ИЗУЧЕНИЯ СОЦИАЛЬНЫХ МЕДИА ПОСРЕДСТВОМ КОМПЬЮТЕРНОГО АНАЛИЗА

### Аннотация:

*В работе рассматривается феномен больших данных в контексте анализа социальных медиа как актуального источника, позволяющего изучать разные аспекты современного общества. Актуализируется роль автоматизации процедуры контент-анализа в двух аспектах. С одной стороны, применение данного метода является необходимым для сбора и интерпретации большого объема качественных данных, что является трендом междисциплинарных исследований. С другой стороны, значимость использования метода очевидна при обращении к такому источнику, как социальные медиа. Именно последние позволяют решать широкий спектр исследовательских задач – от развития маркетингового таргетирования до изучения сознания посредством лингвистики, философии и социологии языка. Предлагаются способы и принципы исследования социальных медиа с позиции социологического знания и междисциплинарных тенденций; раскрывается метод контент-анализа с точки зрения перспектив его использования в социально-гуманитарном знании с применением ресурсов компьютерной обработки данных; показаны ограничения метода и предложены способы организации качественных исследований на материалах социальных медиа с учетом разного программного обеспечения.*

### Ключевые слова:

*большие данные, контент-анализ, качественные исследования данных, интеллектуальный анализ текста, социальные медиа, социальные сети, программное обеспечение для анализа качественных данных*

Oleshkova Anna Mikhailovna

PhD in History, Associate Professor,  
Humanities and Social  
and Economic Sciences Department,  
Nizhny Tagil State Social-Pedagogical Institute  
branch of the Russian State  
Vocational Pedagogical University

## PROSPECTS FOR APPLICATION OF COMPUTER ANALYSIS TO SOCIAL MEDIA STUDIES

### Summary:

*The present research examines the phenomenon of big data in the context of social media analysis, as it can serve a major instrument in studying various aspects of modern society. The role of automated content analysis is presented in two aspects. On the one hand, the application of this method is necessary for collection and interpretation of large amounts of qualitative data, which is a general trend in interdisciplinary research. On the other hand, the importance of using the method is even more apparent when it is applied to social media studies. It is the latter that enables a researcher to investigate a wide range of scholastic subjects starting from target marketing to the study of consciousness through linguistics, philosophy and sociology of language. The author proposes methods and principles of social media research from the standpoint of sociological knowledge and interdisciplinary trends. There is revealed the content analysis with regard to the prospects for their application in computer-aided social and humanitarian studies. The author outlines limitations of the proposed method and suggests other ways of organizing qualitative research of social media data using various software.*

### Keywords:

*big data, content analysis, qualitative data research, text mining, social media, social networks, qualitative data analysis software*

Современное информационное пространство постоянно расширяется, что определяет важность способности субъекта обрабатывать большой массив данных на уровне как повседневной, так и профессиональной культуры. Если в повседневной жизни одним из основных требований к пониманию сложной социокультурной реальности является критическое мышление, то профессиональный анализ требует учета развития компьютерных технологий и появления большого массива неструктурированных данных, что одновременно расширяет тематическое поле исследования и обуславливает использование как новых источников, так и методов их обработки. Относительно новой объектно-предметной областью социальных исследований становится интернет-пространство, которое не только является важным контекстом социального бытия современного человека, но и в тех условиях, с которыми столкнулось человечество в период пандемии 2019–2020 гг., может считаться основным социокультурным пространством, аккумулирующим экономические, культурные, политические и иные процессы. Кроме того, даже без учета новых

экстремальных факторов развития общества в условиях его информационной стадии медиапространство играет важную роль в репрезентации и конструировании картины мира субъекта, а также в презентации его самого другим субъектам.

Социальные медиа, в отличие от традиционных, обладают рядом особенностей, которые обуславливают новый статус субъекта. Уже десять лет назад исследователи отмечали специфику сетевого взаимодействия индивидов [1, р. 64], которое не только определяет иную циркуляцию информации, но и фактически конструирует новую информационную реальность. В определенной степени «человек из телевизора» и «человек из социальной сети» не только живут в разной повестке дня, но и формируют разные информационные миры. Как следствие, они могут обладать разными ценностными установками и ориентирами, используют разную логику и аргументы в обосновании своей позиции, имеют противоположные оценки происходящих событий.

Если в традиционных медиа информация течет однонаправленно, то в социальных медиа статусы публикатора и потребителя информации сходятся на одном субъекте. Социальные медиа строятся на социальном взаимодействии, подразумевающим, что каждый может общаться с каждым. Следует выделить некоторые особенности социальных медиа, к которым можно отнести следующие типичные платформы, акцентирующие разные аспекты мультимедийности: Facebook, Twitter, Instagram, YouTube и др. В контексте развития российского информационного пространства с учетом двух факторов (русскоязычный контент и использование молодым поколением новинок, которые затем распространяются в разновозрастных средах), на наш взгляд, на вышеобозначенном фоне следует выделить социальную сеть «ВКонтакте», мессенджер Telegram и сервис TikTok, который можно также квалифицировать как социальную сеть. Особенности разных социальных медиа являются возможность всегда оставаться на связи; быстрая коммуникация с другими субъектами; оперативный поиск информации и наличие обратной реакции на событие, в том числе оценочного характера; формирование контента, а также его трансляция и ретрансляция.

«Большие данные» (big data) – термин, который активно используется учеными с конца 2000-х гг. В социально-гуманитарном аспекте возможности использования больших данных уже обозначены и реализуются в экономике, истории, антропологии, урбанистике, а также в маркетинге, менеджменте и бизнесе.

Применительно к социологии стоит говорить о перспективах их использования в настоящем времени и в будущем, поскольку современная социология переживает смену акцентов в развитии своей теоретической и эмпирической составляющих. Если в XX в. социологи были ориентированы на разработку опросных технологий, совершенствование метода наблюдения, анкетирования, интервью, то представляется, что XXI в. с учетом общенаучного междисциплинарного тренда переориентирует социальные исследования на использование новых способов производства и обработки данных, позволяющих понять поведение современного человека. Большие данные дают возможность обратиться к тем сторонам человеческого поведения, которые недоступны классическим социологическим методам. Среди прочих свойств, характерных для больших данных, можно выделить так называемые три «V»: volume, velocity, variety [2, р. 1].

Неструктурированные данные представляют особый интерес для исследователя. Информация, посредством которой субъект репрезентирует себя и транслирует модели поведения в незафиксированном и достаточно подвижном пространстве (в сравнении с вопросами анкеты, которая так или иначе представляет собой искусственную среду, достаточно четко очерченную, воспринимаемую субъектом как нечто внешнее и постороннее), позволяет выявить и проанализировать факторы социального влияния в динамике их развития. Онлайн-данные (комментарии в социальных сетях, паблики, микроблоги, чаты) дают возможность не только выявить паттерны поведения, но и определить доминирующие эмоции, оценки и даже болезненные состояния, свойственные субъекту или социальной группе. В этой линии исследований важной оказывается научная кооперация социологов, психологов и лингвистов. В другом аспекте развития данной темы следует выделить компьютерные технологии, которые не просто помогут оптимизировать социальное исследование, а сделают его возможным. В связи с этим стоит отметить системообразующую роль метода контент-анализа, который можно реализовывать при помощи математических алгоритмов и на основе разнообразного программного обеспечения. Контент-анализ и подходящее программное обеспечение позволяют исследователю отбирать, систематизировать и интерпретировать полученные данные. Под качественными исследованиями данных следует понимать разработку концепций, которые будут способствовать пониманию социальной реальности в максимально естественных условиях, акцентировать внимание на смыслах и взглядах конкретных субъектов. Важным положением таких исследований является учет роли реально-

сти, находящейся вне поля наблюдения. Таким образом, нужно говорить о видимом поверхностном уровне как той сфере, которая отображает только часть социальной реальности, а ее глубокие структуры необходимо выявлять и анализировать [3].

Качественный анализ данных (QDA) подразумевает совокупность процессов и процедур, с помощью которых исследователь объясняет и интерпретирует изучаемые ситуации. Языковой материал, представленный в социальных медиа, может стать основой для самых разнообразных тем, в которых исследуется фактическое и символическое содержание качественных данных, например анализ реакции пользователей Twitter на резонансные примеры суицидов [4]. В контексте изучения социальных медиа важно понять, какую именно информацию исследователю будет важно извлекать из данного комплексного источника и как именно это сделать. При работе с программным обеспечением, позволяющим провести качественный анализ данных, следует обратить внимание на основные этапы исследования, которые одновременно являются основными задачами исследователя и непосредственно проектом анализа текста.

При отсутствии программного обеспечения эти этапы можно проводить вручную, собранный материал в этом случае будет являться более локальным и менее масштабным.

1. Сбор данных – подразумевает выделение текстовых данных, необходимых для анализа, например комментарии пользователей определенного паблика. Фильтрация контента обусловит эффективность процедуры предобработки.

2. Синтаксический анализ, преобразование собранного текстового материала. Формирование словаря.

3. Уточнение переменных, связанных с текстом: автор, пол, возраст, дата.

4. Категоризация и кодировка. Процедуры предполагают формулировку критериев и соотнесение корпусов текста с несколькими группами. Критериями могут быть, например, тема или стиль.

Контент-анализ дает возможность преобразования качественных данных в формат, удобный для последующей обработки. Метод контент-анализа можно реализовать на основе Microsoft Excel и Microsoft Word. Обработка естественного языка при помощи автоматизированных процессов позволяет использовать термин «интеллектуальный анализ текста», или text mining. Использование автоматизированной обработки не исключает просмотра информации вручную. Если кратко представить схему реализации качественного анализа данных (QDA), то можно выделить стадии сбора информации, кодирования, поиска и представления результатов. Фактически автоматизированным является процесс поиска данных, все остальные стадии реализуются при участии самого исследователя.

Следует отметить большое разнообразие типов программ, которые позволяют провести качественные исследования: например, Atlas Ti, Nvivo, Taguette, Dedoose, QDA Miner, MaxQDA, QCoder. Важным является наличие бесплатного периода для тестирования возможностей программы. Выбор предпочтительного варианта связан с теми целями, которые ставит перед собой исследователь, и со степенью междисциплинарности его темы. Так, философские и культурологические работы обычно не предполагают количественных подсчетов, и наличие эмпирического материала, скорее, укажет на важные для ученого тренды.

Обозначенные выше ресурсы располагают разными возможностями относительно типов текста (табл. 1), которые можно импортировать (предпочтительнее оказывается MaxQDA, в котором предполагается возможность обработки веб-сайтов, YouTube, Twitter) и экспортировать (например, Atlas Ti позволяет включать кодовые книги, отчеты и визуализации с использованием параметров CSV или XML; информацию можно напрямую экспортировать в SPSS, Endnote или Zotero).

**Таблица 1 – Сильные стороны программного обеспечения для проведения качественных исследований**

Программа	Преимущество
Atlas Ti	Предоставляет возможность автоматического кодирования документов
Dedoose	Включает разнообразие вариантов визуализации информации
QDA Miner	Имеет онлайн-автоматизированные учебные пособия
Atlas Ti, Nvivo, Dedoose	Позволяют осуществлять научную работу коллективу авторов
Nvivo	Поддерживает импорт опросов
MaxQDA, QCoder	Дают возможность использовать смешанные методы анализа данных

В отношении ограничений использования больших данных применительно к анализу социальных медиа мы обозначим не технические стороны программного обеспечения, а факторы, определяющие работу исследователя, и выделим три аспекта. К первому можно отнести субъективную составляющую, в том смысле, что она корректируема: речь идет о повышении научной культуры исследователя и увеличении возможностей использования компьютерных технологий,

а также ресурсов не только смежных областей знания, но и точных и естественных дисциплин. Вторые два аспекта связаны с традиционными, во многом объективными проблемами научного знания, затрагивающими вопросы позитивистских и этических сторон исследования.

Использование больших данных может обусловить активизацию позитивистской методологии, в которой формализм будет важнее интерпретации. Следует учитывать, что объяснительные модели социологии должны сопровождать развитие технологий больших данных и помещать полученные результаты в социальный контекст.

Этические проблемы, характерные для использования больших данных, можно представить в виде следующего перечня, учитывая их применение не столько в исследовательских целях, сколько в коммерческих, экономических и политико-правовых: границы применения полученных сведений и конфиденциальность; наблюдение и контроль, ощущение отсутствия разделения на приватное и публичное; стигматизация и опасность воспроизводства стереотипов в культуре.

В данном списке непосредственно к анализу социальных медиа относится последний пример, однако именно исследование языкового материала социальных сетей, созданного в максимально естественной среде, и его обработка исследователями позволят понять эволюцию стереотипов, циркулирующих в обществе, и, возможно, будут способствовать их развенчанию посредством публикации полученных данных и составлений социологических рекомендаций.

Таким образом, следует говорить о вызове, который испытывает традиционная эпистемология, и необходимости научной рефлексии сложившейся ситуации как в социологическом, так и в научном знании в целом. Анализ больших данных, организация качественных исследований помогают сформировать крупномасштабные модели социального поведения, выйти на анализ причин поступков субъекта, интерпретацию субъективности его восприятия действительности и разнообразие форм выражения чувств и эмоций. При организации таких исследований следует учесть трудоемкость процедуры и сложность ее реализации, однако актуальность и перспективность данного направления особенно очевидны в контексте развития междисциплинарных исследований.

#### **Ссылки:**

1. Kaplan A.M., Haenlein M. Users of the World, Unite! The Challenges and Opportunities of Social Media // *Business Horizons*. 2010. Vol. 53, iss. 1. P. 59–68. <https://doi.org/10.1016/j.bushor.2009.09.003>.
2. Kitchin R. Big Data, New Epistemologies and Paradigm Shift // *Big Data & Society*. 2014. Vol. 1, iss. 1. P. 1–12. <https://doi.org/10.1177/2053951714528481>.
3. Sunday Ch.E. Division for Postgraduate Studies (DPGS) Post-graduate Enrolment and Throughput Program (PET) // SCRIBD [Электронный ресурс]. URL: <https://ru.scribd.com/document/420789502/The-role-of-theory-in-research-pdf> (дата обращения: 19.11.2020).
4. Bridging Big Data and Qualitative Methods in the Social Sciences: a Case Study of Twitter Responses to High Profile Deaths by Suicide / D. Karamshuk, F. Shaw, J. Brownlie et al. // *ScienceDirect*. 2017. Vol. 1. P. 33–43. <https://doi.org/10.1016/j.osnem.2017.01.002>.

Редактор: Фетисова Ирина Викторовна  
Переводчик: Герасимова Валентина Евгеньевна