

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра социальной информатики

КОМПЬЮТЕРНЫЕ МЕТОДЫ КОНТЕНТ-АНАЛИЗА

(автореферат бакалаврской работы)

студента 4 курса 451 группы
направления 09.03.03 - Прикладная информатика
профиль Прикладная информатика в социологии
Социологического факультета
Толстова Даниила Николаевича

Научный руководитель
Кандидат физико-математических наук

_____ М.Г.Плешаков
подпись, дата

Зав. кафедрой
кандидат социологических наук, доцент

_____ И.Г. Малинский
подпись, дата

Саратов 2020

ВВЕДЕНИЕ

Актуальность проблемы. Информации в современном обществе отводится значимое место, которое она не занимала ни в традиционном, ни в модернизирующемся мире. Динамику роста объема информации тяжело представить. Благодаря возрастающим возможностям средств массовой информации (СМИ) информация с высокой скоростью доходит до потребителя, многократно тиражируется и активно внедряется в сознание масс. Это породило новую форму общения – массовую коммуникацию, подразумевающую процесс производства массовой информации и ее дальнейшее распространение при непосредственном общении или с помощью СМИ на численно большие и децентрализованные аудитории. Трансляция одинаковой информации и возможность ее неоднократного воспроизведения дает возможность регулировать воздействие СМИ на массу – субъекта-носителя массового сознания.¹

В информационную эпоху роль СМИ в формировании общественных ценностей, ориентаций и воззрений значительно возрастает. Они овладели способностью целенаправленно воздействовать на массовое сознание, что приводит индивидов, составляющих массу, к жизни в мире «информационных фантомов», транслируемых СМИ. На современном этапе СМИ превратились из средств поиска, переработки и передачи информации в средства, способные контролировать и трансформировать внутренний, духовный мир человека. Поэтому очень важно анализировать получаемую информацию, но из-за ее огромного количества актуальным остается вопрос, какой метод анализа использовать.

Контент-анализ — это метод количественного систематического подхода к изучению текста. Важным является то, что он формализован. Формализованность, систематичность и строгость контент-анализа проявляется в том, что исследование проводится на основании методологически

¹ Ольшанский Д. В. Психология масс. — СПб.: Питер, 2002. — 368 с.

обоснованной программы, по определенным процедурам и служит для получения информации, отвечающей некоторым критериям качества.

С помощью контент-анализа изучались религиозная символика и популярные песни, устанавливались отличия эротических кинокартин от порнографических, устанавливалась мера эффективности политических слоганов, реклам и вражеской пропаганды, определялись особенности суицидального поведения, проявившиеся в предсмертных записках самоубийц, стереотипы сознания различных социальных групп, выявлялась направленность демонстрации людей определенной национальности на телеэкранах, логическая подоплека передовиц газет, отличия в трактовках одного и того же события в разных СМИ, исследовались многие другие темы.

Степень научной разработанности данной проблемы.

В последнее время в отечественной научной среде наблюдается заметное повышение интереса к методу контент-анализа. И, несмотря на то, что подчас бывает непросто найти книгу по данной тематике, интерес к данному способу интерпретации информации возрастает, проводятся дискуссии по теме, предлагаются решения, проводятся исследования. На сегодня помимо упоминания метода в работах монографического плана, посвященных общесоциологическим вопросам, а также в публикациях по теме коммуникации проблематика контент-анализа присутствует в профессиональной периодике: в журналах «Вестник МГУ. Сер.10, Журналистика», «Социологический журнал», «Социология – 4М», «Социологические исследования», и на ежегодных конференциях на факультете журналистики МГУ по проблемам массовых коммуникаций. Важный вклад в развитие процедур контент-анализа внесли российские и эстонские социологи, особенно Алексеев А.Н., Вооглайд Ю., Вихалемм П., Грушин Б.А., Дридзе Т.М., Лауристинь М., Баранов А.В., Шляпентох В., Федотова Л.Н. и многие другие. Небольшие работы по контент-анализу можно найти сейчас и у студентов различных ВУЗов. Так, например, достаточно интересными, на мой взгляд, являются работы Элеазера А.В. «Исследование избирательных стратегий – метод контент-анализа», Рытова А.,

«Опыт применения контент-анализа в биографическом исследовании цесаревича Александра Николаевича Романова» и Четверикова С. «Контент-анализ экономической информации».

Кроме того институт социологии РАН создает в настоящее время программные системы для обработки информации - экспертно-диагностическая система «МАКС», программа «АЛЕКС», система контент-анализа, программа «КСЕНИЯ» и др.

Цель исследования заключается в том, чтобы с помощью применения компьютерных методов контент-анализа показать каким образом СМИ освещают поведение болельщиков во время проведения Чемпионата Европы по футболу 2016.

Объект данной работы - метод контент-анализа как метод количественных исследований.

Предмет данной работы – исследовательские возможности, условия, границы и процедуры применения метода контент-анализа в оценке роли, функций и механизмов воздействия СМИ на поведение болельщиков во время проведения соревнований по футболу под руководством УЕФА.

Теоретическая значимость исследования заключается в возможности использования основных положений и выводов данной работы для дальнейшего изучения выбранной проблематики.

Структура выпускной квалификационной работы представлена введением, тремя главами, заключением, списком использованных источников и приложением.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе «Контент-анализ: история, применение, классические методы» рассматривается история возникновения метода, назначение, область применения и особенности контент-анализа.

Методика контент-анализа нашла широкое применение в информационную эпоху, однако история метода не ограничивается эрой автоматической обработки текста. Так первые примеры использования контент-

анализа датированы XVIII веком, когда в Швеции частота появления в тексте книги определенных тем служила критерием её еретичности. Однако всерьёз говорить о применении контент-анализа как полноценной методики можно лишь начиная с 30-х годов XX века в США. Термин контент-анализ впервые начали применять в конце XIX – нач. XX вв. американские журналисты Б. Мэттью, А. Тенни, Д. Спид, Д. Уипкинс.

Использовался контент-анализ преимущественно в социологических исследованиях, в том числе при изучении рекламных и пропагандистских материалов.

В сфере политических исследований начало использованию методики контент-анализа положил Г. Лассуэл, который занялся анализом пропагандистских материалов периода Второй мировой войны. В 1960-е годы, во время так называемого «методологического взрыва» исследования с применением методики контент-анализа особенно активизировались. Это способствовало развитию методики, разнообразило её варианты. Именно в этот период начинается активное использование компьютерной техники в исследованиях.

В настоящее время к базовым процедурам контент-анализа относятся выявление смысловых единиц и выделение единиц счета.

1) Выявление смысловых единиц контент-анализа. Смысловыми единицами могут быть:

- понятия, выраженные в отдельных терминах;
- темы, выраженные в целых смысловых абзацах, частях текстов, статьях, радиопередачах и т. п.;
- имена, фамилии людей;
- события, факты и т. п.;
- смысл апелляций к потенциальному адресату.

Единицы контент-анализа выделяются в зависимости от содержания, целей, задач и гипотез конкретного исследования.

2) Выделение единиц счета

Единицы счета могут совпадать либо не совпадать с единицами анализа. В 1-м случае процедура сводится к подсчету частоты упоминания выделенной смысловой единицы, во 2-м — исследователь на основе анализируемого материала и здравого смысла сам выдвигает единицы счета, которыми могут быть:

- физическая протяженность текстов;
- площадь текста, заполненная смысловыми единицами;
- число строк (абзацев, знаков, колонок текста);
- длительность трансляции по радио или ТВ;
- метраж пленки при аудио- и видеозаписях,
- количество рисунков с определенным содержанием, сюжетом и пр.

3) Процедура подсчета

Подсчет результатов в общем виде схож со стандартными приемами классификации по выделенным группировкам. Применяется составление специальных таблиц, применение компьютерных программ, специальных формул (напр., «формула оценки удельного веса смысловых категорий в общем объеме текста»), статистические расчеты понятности и аттрактивности текста.

Метод контент-анализа широко применяется как вспомогательное средство в психологии, социологии и др. науках при анализе ответов на открытые вопросы анкет, материалов наблюдений, результатов психологического тестирования (в частности, в проективных методиках), для анализа результатов в методе фокус-групп.

Подобные методы используются также в исследованиях массовых коммуникаций, в маркетинговых и многих др. исследованиях.

Контент-анализ может использоваться для исследования большей части документальных источников, однако лучше всего он работает при относительно большом количестве однопорядковых данных.

Во второй главе «Современные компьютерные технологии контент-анализа» представлены современные компьютерные технологии контент-анализа, обзор доступных продуктов и конкретно выбранной для исследования программы.

Вероятно, самой первой попыткой проведения контент-анализа при помощи компьютера следует считать исследования Томаса Себеока и его коллег, изучивших в 1958 г. около 4 тыс. сказок черемисов. Одной из первых компьютерных программ, созданных специально для проведения контент-анализа, следует считать General Inquirer, изобретенный в 1962 г. Ф. Стоуном. Эта программа использует в настоящее время 182 семантических категории и пользуется словарем в сотни тысяч слов, которые поставлены в соответствие этим категориям. При этом исследуются все смысловые оттенки каждого термина.

Большими возможностями в области анализа документов обладают современные технологии. Так, компьютерный контент-анализ официальных документов компаний позволяет обнаружить скрытые факторы, никак не отражаемые в их основном содержании. В случае кризисных ситуаций они могут оказаться гораздо более информативными, чем обычные показатели. В конце 2001 г. компания ENRON объявила о своем банкротстве. Оно тут же было названо крупнейшим за всю историю США.

Таким образом, весьма трудоемкие процедуры обработки данных контент-анализа существенно облегчаются при использовании компьютерных технологий. Благодаря расширяющемуся применению персональных компьютеров и введению в их память текстовой информации у социологов появилась возможность автоматизировать традиционно ручной фронт работ — контент-анализ ответов респондентов на открытые вопросы в анкетах.

Таким образом, весьма трудоемкие процедуры обработки данных контент-анализа существенно облегчаются при использовании компьютерных технологий. Благодаря расширяющемуся применению персональных компьютеров и введению в их память текстовой информации у социологов

появилась возможность автоматизировать традиционно ручной фронт работ — контент-анализ ответов респондентов на открытые вопросы в анкетах.

В своем исследовании я использовал программное обеспечение от ATLAS.ti. Это мощный инструмент для качественного анализа больших массивов текстовых, графических, аудио и видео данных. Сложные инструменты помогут вам систематизировать, собирать и управлять материалом творчески, но и систематически. ATLAS.ti позволяет сосредоточиться на самом материале. Независимо от того, занимаетесь ли вы антропологией, экономикой, криминологией или медициной: ATLAS.ti удовлетворит ваши потребности в качественном анализе.

Работа возможна с текстовыми документами во всех основных форматах, включая txt, .doc, .docx, .odt и, конечно же, .pdf. Можно выполнять полностью автоматический поиск в одном или нескольких документах, автокодирование и другие мощные семантические операции для извлечения смысла. Достаточно просто работать с десятками графических и аудиоформатов (.wav, mp3, .wma и т. Д.), А также с наиболее распространенными типами видео (.avi, .mp4, .wmv и т. Д.).

Центральное рабочее пространство в ATLAS.ti служит контейнером для данных проекта. Доступ ко всем основным компонентам проекта, таким как документы, выделенные/закодированные сегменты данных, коды, заметки, гиперссылки, группы или сети, являются быстрым и удобным. Кодирование может быть выполнено простым перетаскиванием кода на выбранный фрагмент данных. Диспетчеры объектов, обозреватель проектов и обозреватель совместных экземпляров позволяют просматривать данные проекта и перемещаться по ним.

В третьей главе «Контент-анализ освещенности поведения болельщиков на Чемпионате Европы по футболу 2016 в СМИ» представлен контент-анализ освещенности поведения болельщиков на Чемпионате Европы по футболу 2016 в СМИ, выводы по результатам исследования.

Для анализа были отобраны 5 интернет-порталов новостей: «Ведомости», «Лента.ру», «РБК», «РТ» и «Спорт-Экспресс».

Для анализа разрабатывалась система категорий, которые можно разделить на две структурные группы. Первая группа включает формальные характеристики публикации, вторая – тематические характеристики публикации.

После разработки категорий был произведен сбор данных. Согласно описанию объекта были изучены материалы СМИ, которые публиковались в период проведения Чемпионата Европы по футболу 2016, то есть с 10 июня 2016 года по 10 июля 2016 года. Поиск необходимых публикаций в указанных СМИ осуществлялся через поисковой тэг «беспорядки болельщиков»

По итогам исследования 5 интернет-порталов тема поведения болельщиков, как основная тема публикации, затрагивалась в 304 статьях. В «Ведомостях» - 20 статей, в «Лента.ру» - 62 статьи, в «РБК» - 62 статьи, в «РТ» - 46 статей, а самое большое количество статей, что обусловлено тематикой интернет-портала у «Спорт-Экспресс» - в нем тема поведения болельщиков затрагивалась в 114 статьях.

Проведя исследование жанровой направленности публикаций за весь изучаемый период можно сказать, что чаще всего использовались «статьи» (33,6%), «информация» (28,6%) и «репортаж» (26,6%). Наименее используемый жанр «интервью» - всего 11,2%

При освещении проблемы доминирует нейтральный тон, он зафиксирован в 184 публикациях (60,5%). Следом за ним идет позитивный тон, который отмечен в 75 публикациях (24,6%). И негативный в 45 публикациях (14,8%).

В ходе исследования выяснилось, что чаще всего в новостной ленте упоминался Александр Шпрыгин – 275 упоминаний (57,3%). Это не удивительно, ведь он является создателем общероссийского общественного движения «Всероссийское объединение болельщиков». Второй, по количеству упоминаний, на втором месте оказался Владимир Путин с 65 упоминаниями в

анализируемых публикациях (13,5%) – президент Российской Федерации. Это также объективно понятно, так как он неоднократно давал оценку и комментировал действия футбольных болельщиков в ходе Чемпионата Европы по футболу 2016 года.

Чаще всего в СМИ упоминались российские (61,8%) и английские болельщики (27,6%). российские (61,8%) и английские болельщики (27,6%). Также упоминались украинские (7%), венгерские (2,6%) и польские (0,9%) болельщики.

Самая частая тема, освещаемая в средствах массовой информации, стала «дисциплинарные наказания». 236 раз (38,3%) СМИ касались этого вопроса в своих публикациях. Такое частое упоминание обусловлено тем, что в ситуациях, когда происходят подобные беспорядки необходимо давать знать, что ответственность за это понесут не только болельщики, но и сборные их стран. Что именно их действия влекут за собой неблагоприятные последствия для тех, за кого они болеют. Следующее по частоте упоминаний – жесткие меры со стороны полиции, а именно применение силы, слезоточивого газа и водометов против агрессивно настроенных болельщиков. Далее идут – действия со стороны полиции Франции. Под этим подразумевается любые упоминания работы французских полицейских такие, как задержания, подготовка к предшествующим футбольным матчам и возможным беспорядкам. Как мне кажется, такое частое касание темы полиции в СМИ в период расстройств и разногласий среди болельщиков является важной задачей работы прессы. Весь агрессивно настроенный контингент обязан понимать, что им может быть оказано сопротивление, зачастую очень жесткое. Далее 45 (7,3%) упоминаний о заключениях под стражу – реального лишения свободы. Это тоже является важной составляющей в ходе подобных событий. Все болельщики должны видеть настоящие кейсы, к чему приведет их подобное поведение. Также часто поднималась тема получения ранений в ходе беспорядков – 38 упоминаний (6,2%) и тема госпитализации болельщиков – 29 упоминаний (4,7%).

Подводя итоги проведенного исследования, можно сделать вывод, что проблема поведения болельщиков во время Чемпионата Европы по футболу 2016 года и уровень ее освещенности был достаточно высок. На примере 5 интернет-порталов можно увидеть, насколько часто издавались публикации, посвященные этой теме. А также можно обратить внимания, как ярко освещались конкретные темы в каждой публикации. Безусловно, подобный резонанс в СМИ благотворно влияет на предшествующие крупные футбольные турниры. Стоит отметить, что опыт 2016 года положительно сказался на Чемпионате Мира по футболу 2018 года, проводившегося в России. Тогда не случилось ни одной стычки футбольных болельщиков. А все благодаря нововведениям после ЕВРО-2016. Все это не было бы возможно, если не четкая и своевременная работа СМИ.

ЗАКЛЮЧЕНИЕ

На проведение исследования проблемы освещенности поведения болельщиков в СМИ могло бы уйти намного больше затрат: как временных, так и материальных. Но с использованием компьютерных методов контент-анализа, а именно программного обеспечения ATLAS.ti 9 работ существенно сократилась. В настоящее время невозможно представить проведение контент-анализа без использования программного обеспечения. Особенно это касается анализа СМИ, где объем новостей, статей, репортажей по совершенно разным темам с каждым днем только растет.

Как мне кажется, в постоянном росте информационного потока очень тяжело найти репрезентативную выборку. Компьютерные методы контент-анализа становятся неотъемлемой частью современного инструментария любого деятеля социогуманитарных наук. Ведь с помощью этого мощного инструмента можно анализировать огромные массивы информации такие, как различные типы текстов, сообщения в СМИ, заявления политических деятелей, программы партий, правовые акты, рекламные и пропагандистские материалы, исторические источники и даже литературные произведения.