

DOI: <https://doi.org/10.23670/IRJ.2023.135.19>

МЕТОД BIG DATA ДЛЯ КОНТЕНТ-АНАЛИЗА (НА ПРИМЕРЕ РЕГИОНАЛЬНОГО СМИ ОБЩЕСТВЕННО-ПОЛИТИЧЕСКОЙ ГАЗЕТЫ «ВЕЧЕРНЯЯ УФА»)

Научная статья

Полякова Е.В.^{1,*}, Демичева А.С.²

¹ ORCID : 0000-0001-6606-6342;

² ORCID : 0009-0002-1437-7581;

^{1,2} Уфимский университет науки и технологий, Уфа, Российская Федерация

* Корреспондирующий автор (fotodarr[at]mail.ru)

Аннотация

Целью данной статьи является сравнение традиционного контент-анализа и его аналога на основе анализа больших данных (big data) для выявления особенностей последнего, его возможностей и ограничений. Для наглядности в работе приводится результат машинного анализа новостной колонки региональной газеты «Вечерняя Уфа» на предмет влияния Специальной военной операции на Украине на новостную повестку издания. В основной части статьи проведенное исследование анализируется с точки зрения его методологических особенностей, раскрываются его преимущества (скорость, работа с генеральной совокупностью) и недостатки (грубая работа с языком, невозможность учесть контекст). Делается вывод об эвристической ценности анализа больших данных, но также о его недостаточности для полноценного исследования.

Ключевые слова: анализ больших данных, контент-анализ, научная методология, газета «Вечерняя Уфа», Специальная военная операция.

BIG DATA METHOD FOR CONTENT ANALYSIS (ON THE EXAMPLE OF THE REGIONAL MASS MEDIA OF THE SOCIO-POLITICAL NEWSPAPER "VECHERNAYA UFA")

Research article

Poliakova E.V.^{1,*}, Demicheva A.S.²

¹ ORCID : 0000-0001-6606-6342;

² ORCID : 0009-0002-1437-7581;

^{1,2} Ufa University of Science and Technology, Ufa, Russian Federation

* Corresponding author (fotodarr[at]mail.ru)

Abstract

The aim of this article is to compare traditional content analysis and its analogue based on big data analysis in order to identify the specifics of the latter, its capabilities and limitations. For illustrative purposes, the work presents the result of machine analysis of the news column of the regional newspaper "Vechernaya Ufa" on the impact of the Special Military Operation in Ukraine on the news agenda of the publication. In the main part of the article the conducted research is analysed from the point of view of its methodological characteristics, its advantages (speed, work with the general population) and disadvantages (rough work with language, inability to take into account the context) are disclosed. The conclusion is made about the heuristic value of big data analysis, but also about its insufficiency for a full-fledged research.

Keywords: big data analysis, content analysis, scientific methodology, "Vechernyaya Ufa" newspaper, Special Military Operation.

Введение

Современные научные исследования редко обходятся без применения новых цифровых инструментов и методов – развитие технологий и техники диктует ученым новые подходы к проведению исследований. Одним из таких инструментов являются анализ больших данных и основывающиеся на нем конкретные научные методы. Например, метод контент-анализа известен уже давно, но появление специальных программ и средств расширило его возможности и сделало более доступным. Так, например, значительную популярность набирает анализ социальных сетей ([7], [8], [9] и др.) и электронных СМИ и прочих источников информации ([3], [5], [14] и др.).

Контент-анализ в первую очередь является количественным методом – его характеризуют как формализованный, систематизированный и строгий [11, С. 15], что облегчает его трансфер в машинную среду, также оперирующую формализованными единицами, однако для применения этого метода важно выявить, как именно перевод от исследователя к машине влияет на него, какие дает возможности и какие появляются ограничения.

В данной статье для наглядности проводится контент-анализ новостной колонки регионального СМИ машинным способом с целью выявить, насколько в нем отразилась тема Специальной военной операции на Украине, и последующий разбор особенностей и отличий по сравнению с традиционным методом проведения подобного исследования.

Методы и процедура исследования

Используя возможности программы RStudio, с помощью пакета rvest с сайта газеты «Вечерняя Уфа» [10], были собраны сообщения из раздела «Новости» за десять месяцев (октябрь 2021 – июль 2022) – пять месяцев до СВО и пять месяцев после, чтобы можно было сравнить, как изменилась лексика публикаций.

```
> paths_allowed(path=c("http://vechufa.ru/"))
vechufa.ru

[1] TRUE
warning message:
In get(name, envir = env, inherits = FALSE) :
  strings not representable in native encoding will be translated to UTF-8
> |
```

Рисунок 1 - Проверка политики сайта подтвердила, что данные с него собирать разрешено
DOI: <https://doi.org/10.23670/IRJ.2023.135.19.1>

Отдельно стоит отметить, что в сфере работы с цифровой информацией уже сложились (и продолжают складываться) собственные этические нормы. Так, владельцы открытых источников (в данном случае редакция газеты) по своему усмотрению и на основе правовых норм устанавливают, можно ли и какую именно информацию допустимо собирать с их ресурса машинным способом. Например, сайт «Вечерней Уфы» (см. рис. 1) запрещает к сбору только разделы, связанные с пользователями сайта (их регистрационные данные, записи/комментарии и т.п.), что является достаточно частой политикой открытых источников. В то же время большинству таких источников затруднительно контролировать исполнение подобных запретов, и оно остается на совести исследователя. Таким образом, в этике ученого появляются новые пункты, связанные с добросовестным использованием открытой информации, о чем не следует забывать при проведении исследований цифровыми методами на основе такой информации.

Возвращаясь к теме статьи: всего было спарсено (собрано) 210 новостей. Для последующего анализа была проведена предобработка возможностями пакета tm. Предобработка включала в себя стемминг (обрезку окончаний), замену заглавных букв на прописные, удаление чисел, знаков препинания, лишних пробелов и стоп-слов. К стандартному списку стоп-слов для русского языка (союзы, предлоги и т.п.) были добавлены имена авторов статей. Кроме того, из перечня слов, входивших в заголовки, были удалены слова повторяющегося объявления о поиске редакцией корреспондента. Все эти операции являются рутинными при работе с текстовыми массивами и не представляют собой чего-то исключительного. По сути, они позволяют отсеять языковой «шум», унифицировать и взять в работу только значимые лексемы, на основании которых можно говорить о содержании текстов и объеме содержащихся в них тем.

По итогам предобработки фрейм заголовков насчитывал 509 слов, фрейм непосредственно текстов – 4156 слов. На рисунке 2 представлены первые 15 строк фреймов заголовков (рис. 2а) и текстов статей (рис. 2б). Разумеется, для больших данных это очень незначительный объем – в этой области оперируют на порядки превосходящими числами – сотнями тысяч и миллионами. Чем больший объем данных обрабатывается, тем более детальный и точный результат можно получить. Однако для целей этой работы нет необходимости расширять выборку, поскольку нас интересует методологическая часть, а не непосредственно результаты анализа.

	freq	word		freq	word
	17	уф		148	уф
корреспондент	14	корреспондент	-	118	-
вечерн	13	вечерн	республик	83	республик
больш	4	больш	город	81	город
нов	4	нов	улиц	80	улиц
лет	3	лет	год	71	год
чест	3	чест	башкортоста	64	башкортоста
муз	3	муз	работ	59	работ
лидер	3	лидер	подготов	52	подготов
дет	3	дет	росс	47	росс
подорожа	3	подорожа	нов	47	нов
башкир	3	башкир	центр	42	центр
смен	3	смен	тысяч	40	тысяч
пенсионер	3	пенсионер	администрац	39	администрац
нужн	3	нужн	государствен	39	государствен

а)

б)

Рисунок 2 - Первые 15 слов каждого фрейма – заголовки (а) и тексты статей (б)
DOI: <https://doi.org/10.23670/IRJ.2023.135.19.2>

Далее с помощью пакета wordcloud были построены облака слов к каждому из фреймов. Облако слов представляет собой визуализацию частотности употребления слов в тексте: чем чаще встречается какой-либо элемент в исследуемом объеме, тем более крупным шрифтом он написан, и, соответственно, наоборот. Поэтому облака слов

достаточно наглядны и интуитивно понятны – даже короткий взгляд на этот вид графика позволяет легко определить, какие темы преобладают, какие имеют меньшее значение, какие примерно равны.

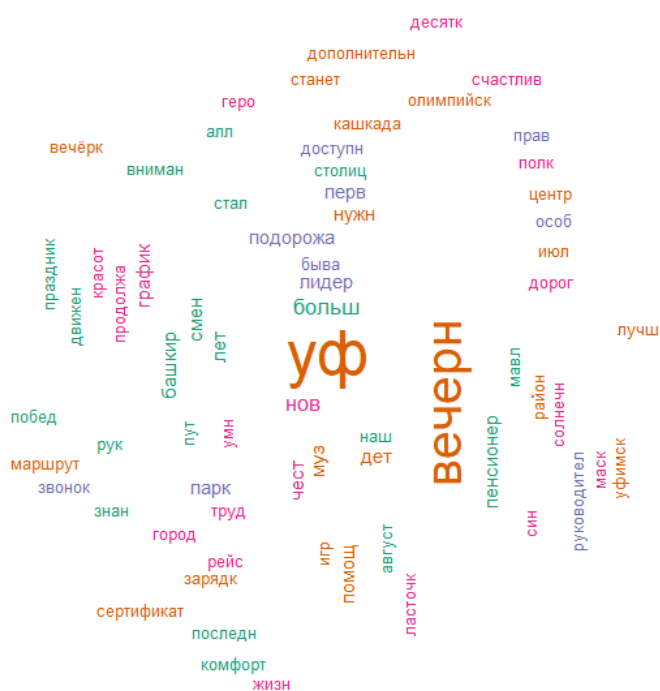


Рисунок 3 - Облако слов, упоминаемых в заголовках за весь выбранный период
DOI: <https://doi.org/10.23670/IRJ.2023.135.19.3>

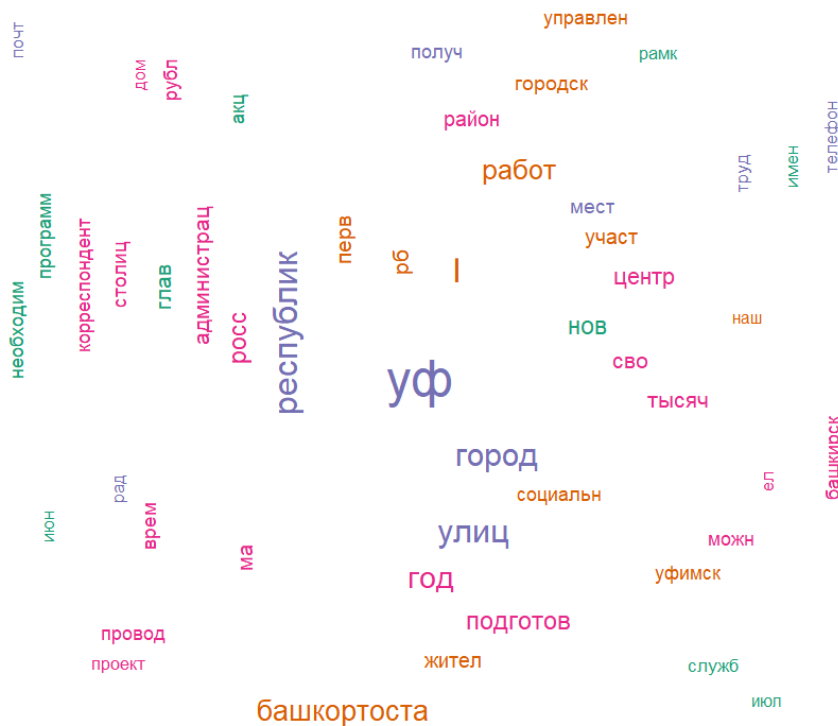


Рисунок 4 - Облако слов, используемых в текстах статей
DOI: <https://doi.org/10.23670/IRJ.2023.135.19.4>

Как видно из рисунков 2а и 3, частота упоминания слов в заголовках не особо различается. Выделяются «вечерняя» и «Уфа», поскольку они входили в повторяющееся объявление о поиске корреспондента, остальные слова на инфографике примерно одного размера, их частота употребления – 2-3 раза.

Облако слов из текстов статей (рис. 4) дает больше информации. Например, очевидно, что стемминг проводится очень грубо: такие слова, как «республика» и «республиканский», «Уфа» и «уфимский» восприняты системой как разные, хотя для контент-анализа разница несущественна. Кроме того, в списки попали неинформативные обрезки слов – «ел», «можн», «наш».

Любопытно, что имя Главы Республики Радий вошло в облако в форме «рад», в то время как фамилия Хабиров не попала. Если посмотреть в самом фрейме, то «рад» встречается 19 раз, а «хабир» – только 9. Полагаем, это может быть связано с тем, что обработанная форма образована не только от имени Главы, но и от краткого прилагательного «рад», предлога «ради» или даже от упоминания украинского парламента.

Также в облаке вызывает интерес форма «сво». Не вызывает особого сомнения утверждение, что это стемминг различных форм местоимения «свой», однако вполне допустимо предположить, что какая-то часть могла быть образована аббревиатурой СВО – специальная военная операция, проводимая в настоящее время Российской Федерацией на территориях ДНР, ЛНР и Украины. О том, что эта часть, если она действительно есть, не особо значительна, свидетельствует контекст – остальные слова в облаке не отсылают однозначно к военной тематике. Отчасти к таковым могут быть отнесены «защита» (но – опять же «защита детей», «защита прав потребителей» и т.п.), «участник» (устойчивое в настоящее время выражение «участник СВО» не перекрывает «участников» любых других мероприятий, тем более что в облаке также фигурирует слово «конкурс») и «служба» (но и здесь нет однозначной отсылки к военной службе – это может быть государственная или муниципальная служба, служба спасения, служба занятости и т.п.).

Если посмотреть только новости за период после начала СВО, то картина остается практически идентичной (см. рис. 5).

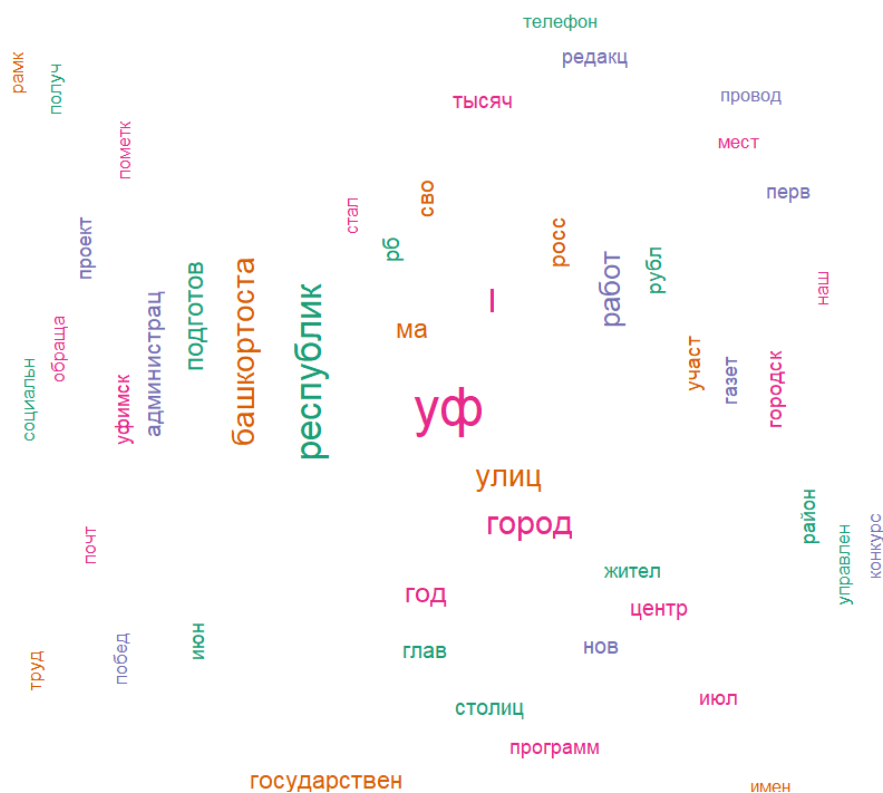


Рисунок 5 - Облако слов, используемых в текстах статей после начала СВО
DOI: <https://doi.org/10.23670/IRJ.2023.135.19.5>

Из новых слов появились «пометк», «обраща», «побед», «государствен», «стал», «редакц», «газет», «конкурс». Скорее всего, в тренд вышли темы, связанные с 9 мая (День Победы) и что-то, связанное с редакционными новостями. Можно сделать вывод о том, что СВО отразилась на региональной новостной повестке очень слабо (если вообще отразилась).

Основные результаты

Представляется очевидным, что машинный анализ обладает гораздо более высокой степенью детализации исследования, однако следует отметить, что она достигается в ущерб целостному видению ситуации. Частотный подсчет отдельных единиц не позволяет восстановить контекст, который воспринимает исследователь при обращении к традиционному контент-анализу. То есть результатом машинного анализа можно считать мозаику тем, для дальнейшей разработки которых нужно более углубленное исследование. Для некоторых исследований этого может быть достаточно – например, если нужно проанализировать, какой объем занимает тема в совокупности текстов, ее соотношение с другими темами и т.п. Но если целью исследования является что-то более сложное, например, оценка

отношения автора к теме, языковой портрет, определение влияющих факторов и т.д., без дополнительных процедур, причем в меньшей степени цифровизированных, не обойтись. Есть и промежуточные задачи, решаемые машинным способом, но углубляющиеся в анализ – сравнение текстов по интересующим темам (их близость или отдаленность друг от друга), поиск связей между темами или текстами и т.д. Однако для их решения, опять же, необходимо более серьезное погружение в исследование, не ограничивающееся написанием кода для машины, а требующее понимания проблемной области и постановки задачи.

Необходимо отметить также, что машинный анализ базируется на принципиально ином подходе к анализу текста, чем традиционный контент-анализ. Оставляя в стороне формирование выборки – при использовании последнего сначала определяется тема, формируются ключевые теоретические понятия, выявляются субкатегории и единицы счета и только затем происходит подсчет единиц, который позволяет анализировать тексты по заданной теме. В случае машинного анализа алгоритм практически противоположен: сначала происходит подсчет, а после выделяются темы и категории. С учетом этого можно ли утверждать, что машинный анализ все еще остается контент-анализом, или это какой-то новый метод, требующий дифференциации от традиционного контент-анализа?

Применение метода big data для контент-анализа новостных публикаций регионального издания выявило как плюсы, так и минусы такого подхода. К его достоинствам, безусловно, можно отнести скорость сбора и обработки материала: написанный код проводит указанные операции за несколько минут, легко адаптируется под уточненные задачи (общий фрейм и фрейм за определенный период), визуализирует результат.

К выявленным недостаткам можно отнести грубую работу с языком, высокую погрешность образования форм и иные сложности, вызванные в том числе особенностями русского языка. Стоит отметить, что существуют более совершенные инструменты, например, вместо стемминга может быть использована лемматизация – т.е. слова не обрезаются, а приводятся в первоначальную форму (именительный падеж единственного числа для существительных, инфинитив для глаголов и т.д.). Да и проведенные операции могут быть осуществлены с помощью других пакетов и других команд, что может дать отличающийся результат. Если говорить образно, то это как использовать для измерения два теоретически одинаковых прибора, но из-за погрешности настройки они могут дать разный результат. Кроме того, существуют специальные программы, предназначенные непосредственно для контент-анализа (например, MaxQDA, [13]), чей функционал шире, а настройки гибче и в большей степени способствуют решению поставленных задач.

Обсуждение

Если говорить о сравнении традиционного контент-анализа и исследования методом анализа больших данных, то можно выделить следующие отличия:

1. Нет необходимости формировать выборочную совокупность – методика позволяет провести контент-анализ всей генеральной совокупности. Во-первых, это позволяет сократить время на подготовку к исследованию. Во-вторых, помогает избежать ошибок, которые могут быть совершены при формировании выборочной совокупности. В-третьих, дает более точный результат, поскольку анализ на основе выборочной совокупности по умолчанию является неполной индукцией, в то время как генеральная совокупность (по умолчанию) исчерпывает множество исследуемых объектов, а значит, ее анализ представляется полной индукцией;

2. Нет необходимости выделять единицы анализа. При проведении традиционного контент-анализа особо кропотливая и трудоемкая часть работы связана с подбором смысловых единиц, наличие которых и анализируется в исследуемом объекте. Анализ больших данных, по сути, превращает каждое слово, встречающееся в каждом тексте совокупности, в такую единицу и считает частоту его употребления. Схожие алгоритмы есть и для более крупных единиц анализа (например, устойчивые словосочетания, темы и т.п.);

3. При машинном анализе затруднительно выявить и учесть синонимичные слова, метафорические обороты и использование местоимений при подсчете единиц анализа. Также представляет определенную сложность оценка контекста для них. В этих ситуациях необходимо оценивать слова не как отдельные единицы, а как часть целого (словосочетания, предложения, абзаца), а машина именно что разрывает связи между единицами и ведет учет каждой отдельно. Некоторые программные решения [13], специализирующиеся именно на контент-анализе, позволяют обратиться к контексту в каждом конкретном случае – выдают подборку фрагментов текстов, где встречается выбранная единица анализа, на установленную исследователем «глубину» – количество слов вокруг, предложение, абзац, в котором оно употребляется, что частично снимает указанную трудность, однако не избавляет от нее полностью, поскольку требует уже «немашинного» подхода к анализу и, в каком-то смысле, возвращает к традиционной форме исследования.

Отдельно следует отметить тот факт, что традиционный контент-анализ требует от исследователя четкой постановки цели и задач исследования – и, соответственно, длительную процедуру его подготовки. Он требует постановки проблемы. Машинный же анализ в этом отношении более лоялен к исследователю. Он позволяет взять любой массив информации и посмотреть, что в нем есть интересного и проблемного. Разумеется, исследователь должен понимать, куда смотреть и что искать, и, тем не менее, машинный анализ дает куда больший простор при постановке целей и задач [4, С. 15].

Заключение

Резюмируя сказанное, машинный анализ представляется удобным инструментом на начальном этапе для первичной оценки текстов. Он позволяет анализировать всю совокупность текстов, а не выборку, выделить основные темы (единицы анализа), оценить общий «объем» интересующих тем в генеральной совокупности. По сути, строит наглядную макрокарту исследуемых текстов, отталкиваясь от которой можно переходить к более глубокому и конкретному исследованию интересующей темы, в том числе к качественному контент-анализу.

Если говорить об анализе больших данных, как о научном методе, то его научным результатом является эмпирическое обобщение (на основе полной или неполной индукции – в зависимости от характеристики генеральной совокупности), что с уверенностью позволяет отнести данный метод к количественным методам эмпирического уровня научного познания [6].

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Akhmetyanova N. State and Legal Regulation of the Flow of MediaInformation about Coronavirus Infection COVID-19 / N. Akhmetyanova, I. Saitbattalova, E. Poliakova // Proceedings of the IX International Scientific and Practical Conference "Current Problems of Social and Labour Relations (ISPC-CPSLR 2021)". — Atlantis Press, 2022. — P. 10-13. — DOI: 10.2991/assehr.k.220208.002.
2. Ахметьянова Н.А. Современные печатные издания Республики Башкортостан: учебное пособие / Н.А. Ахметьянова. — Уфа: Издательский дом «Республика Башкортостан», 2018. — 113 с.
3. Березовская А.А. «Вот-вот войдет в Scopus»: контент-анализ научных электронных журналов / А.А. Березовская // Актуальные проблемы лингвистики и литературоведения: Сборник материалов VI (XX) Международной конференции молодых ученых, Томск, 18-19 апреля 2019 года / Под ред. Е.О. Третьякова. — Томск: СГТ, 2020. — С. 334-335.
4. Бодров А.А. Философия больших данных в парадигме междисциплинарных исследований / А.А. Бодров // Вестник Самарского государственного технического университета. Серия: Философия. — 2020. — № 1(3). — С. 14-21.
5. Борисова У.С. Имидж мэра Якутска: контент-анализ интернет СМИ / У.С. Борисова, Н.Р. Попов // Современные научные исследования: актуальные вопросы, достижения и инновации: сборник статей XIV Международной научно-практической конференции, Пенза, 5 октября 2020 года. — Пенза: Наука и Просвещение, 2020. — С. 211-213.
6. Демичева А.С. Анализ больших данных как эмпирический метод научного познания / А.С. Демичева // Пятая зимняя школа по гуманитарной информатике: сборник докладов, Калининград, 16-18 декабря 2021 года. — Калининград: Балтийский федеральный университет имени Иммануила Канта, 2021. — С. 3-7.
7. Маранчак А.Г. Социальные сети как новый политический инструмент: проект контент-анализа сети Instagram / А.Г. Маранчак // Молодежь XXI века: образ будущего: Материалы Всероссийской научной конференции с международным участием, Санкт-Петербург, 14-16 ноября 2019 года / Отв. ред. Н.Г. Скворцов, Ю.В. Асочаков. — Санкт-Петербург: Скифия-принт, 2019. — С. 203-204.
8. Некрылова В.П. Контент-анализ профилей Instagram с учебными материалами по РКИ и немецкому языку / В.П. Некрылова // Лингвистика и вызовы современной парадигмы общественных отношений: междисциплинарное, межкультурное, межъязыковое взаимодействие: материалы международного лингвистического форума, Воронеж, 24-28 ноября 2021 года. Часть 2. — Воронеж: Воронежский государственный университет, 2022. — С. 96-101.
9. Никольская А.В. Применение метода контент-анализа в изучении отношения пользователей социальных сетей к современной российской медицине / А.В. Никольская, А.А. Костригин // Южно-российский журнал социальных наук. — 2019. — Т. 20. — № 1. — С. 72-90. — DOI: 10.31429/26190567-20-1-72-90.
10. Общественно-политическая газета «Вечерняя Уфа». Новости. — URL: <http://vechufa.ru/news/> (дата обращения: 06.08.2022).
11. Пашинян И.А. Контент-анализ как метод исследования: достоинства и ограничения / И.А. Пашинян // Научная периодика: проблемы и решения. — 2012. — № 3. — С. 13-18. — URL: <https://cyberleninka.ru/article/n/kontent-analiz-kak-metod-issledovaniya-dostoinstva-i-ogranicheniya> (дата обращения: 08.05.2023).
12. Полякова Е.В. Массмедийный дискурс: специфика и особенность / Е.В. Полякова, Н.А. Садыкова // Филологические науки: состояние, перспективы, новые парадигмы исследований: Материалы IV Международной научной конференции / отв. ред. А.Р. Мухтаруллина. — Уфа: РИЦ БашГУ, 2021. — С. 118-121. — DOI: 10.33184/fnspnpi-2021-11-18.17.
13. MaxQDA 2022 User Manual // MAXQDA. — URL: <https://www.maxqda.com/help-mx22/welcome> (accessed: 06.05.2023). [in Russian]
14. Студенцов О.Р. Контент-анализ онлайн-версий информационных программ на чувашском языке государственной телерадиокомпании «Чувашия» (обзор) / О.В. Студенцов, Г.В. Хораскина, Л.А. Васильева // Медиасреда. — 2018. — № 13. — С. 212-216.

Список литературы на английском языке / References in English

1. Akhmetyanova N. State and Legal Regulation of the Flow of MediaInformation about Coronavirus Infection COVID-19 / N. Akhmetyanova, I. Saitbattalova, E. Poliakova // Proceedings of the IX International Scientific and Practical Conference "Current Problems of Social and Labour Relations (ISPC-CPSLR 2021)". — Atlantis Press, 2022. — P. 10-13. — DOI: 10.2991/assehr.k.220208.002.

2. Ahmet'janova N.A. Sovremennye pechatnye izdaniya Respubliki Bashkortostan: uchebnoe posobie [Modern Printed Publications of the Republic of Bashkortostan: Study Guide] / N.A. Ahmet'janova. — Ufa: Republic of Bashkortostan Publishing House, 2018. — 113 p. [in Russian]
3. Berezovskaja A.A. «Vot-vot vojdet v Scopus»: kontent-analiz nauchnyh jelektronnyh zhurnalov ["About to enter Scopus": content analysis of scientific electronic journals] / A.A. Berezovskaja // Aktual'nye problemy lingvistiki i literaturovedeniya: Sbornik materialov VI (XX) Mezhdunarodnoj konferencii molodyh uchenykh, Tomsk, 18-19 aprelja 2019 goda [Current Problems of Linguistics and Literary Studies: Collection of Materials of the VI (XX) International Conference of Young Scientists, Tomsk, April 18-19, 2019] / Ed. by E.O. Tret'jakov. — Tomsk: STT, 2020. — P. 334-335. [in Russian]
4. Bodrov A.A. Filosofija bol'shih dannyh v paradigme mezhdisciplinarnykh issledovanij [Philosophy of Big Data in the Paradigm of Interdisciplinary Research] / A.A. Bodrov // Vestnik Samarskogo gosudarstvennogo tehničeskogo universiteta. Serija: Filosofija [Bulletin of Samara State Technical University. Series: Philosophy]. — 2020. — № 1(3). — P. 14-21. [in Russian]
5. Borisova U.S. Imidzh mjera Jakutska: kontent-analiz internet SMI [Image of Yakutsk Mayor: Content Analysis of Internet Media] / U.S. Borisova, N.R. Popov // Sovremennye nauchnye issledovanija: aktual'nye voprosy, dostizhenija i innovacii: sbornik statej XIV Mezhdunarodnoj nauchno-praktičeskoj konferencii, Penza, 5 oktjabrja 2020 goda [Modern Scientific Research: Topical Issues, Achievements and Innovations: Collection of Articles of XIV International Scientific and Practical Conference, Penza, October 5, 2020]. — Penza: Science and Enlightenment, 2020. — P. 211-213. [in Russian]
6. Demicheva A.S. Analiz bol'shih dannyh kak jempiricheskiy metod nauchnogo poznaniya [A Big Data Analysis as an Empirical Method of Scientific Knowledge] / A.S. Demicheva // Pjataja zimnjaja shkola po gumanitarnoj informatike: sbornik dokladov, Kaliningrad, 16-18 dekabrja 2021 goda [Fifth Winter School on Humanitarian Informatics: Collection of Papers, Kaliningrad, December 16-18, 2021]. — Kaliningrad: Baltic Federal University named after Immanuel Kant, 2021. — P. 3-7. [in Russian]
7. Maranchak A.G. Social'nye seti kak novyj političeskij instrument: proekt kontent-analiza seti Instagram [Social Media as a New Political Tool: Instagram Content Analysis Project] / A.G. Maranchak // Molodezh' XXI veka: obraz budushhego: Materialy Vserossijskoj nauchnoj konferencii s mezhdunarodnym uchastiem, Sankt-Peterburg, 14-16 nojabrja 2019 goda [Youth of the XXI Century: Image of the Future: Proceedings of the All-Russian Scientific Conference with International Participation, St. Petersburg, November 14-16, 2019] / Chief ed. N.G. Skvorcov, Ju.V. Asochakov. — Saint Petersburg: Skifia-Print, 2019. — P. 203-204. [in Russian]
8. Nekrylova V.P. Kontent-analiz profilej Instagram s uchebnymi materialami po RKI i nemeckomu jazyku [A Content Analysis of Instagram Profiles with RFL and German Teaching Materials] / V.P. Nekrylova // Lingvistika i vyzovy sovremennoj paradigmy obshhestvennykh otnoshenij: mezhdisciplinarnoe, mezhkul'turnoe, mezh#jazykovoe vzaimodejstvie: materialy mezhdunarodnogo lingvističeskogo foruma, Voronezh, 24-28 nojabrja 2021 goda. Chast' 2 [Linguistics and Challenges of Modern Paradigm of Social Relations: Interdisciplinary, Intercultural, Interlingual Interaction: Materials of the International Linguistic Forum, Voronezh, November 24-28, 2021. Part 2]. — Voronezh: Voronezh State University, 2022. — P. 96-101. [in Russian]
9. Nikol'skaja A.V. Primenenie metoda kontent-analiza v izuchenii otnoshenija pol'zovatelej social'nykh setej k sovremennoj rossijskoj medicine [Application of Content Analysis Method in the Study of Social Media Users' Attitudes towards Modern Russian Medicine] / A.V. Nikol'skaja, A.A. Kostrigin // Juzhno-rossijskij zhurnal social'nykh nauk [South Russian Journal of Social Sciences]. — 2019. — Vol. 20. — № 1. — P. 72-90. — DOI: 10.31429/26190567-20-1-72-90. [in Russian]
10. Obshhestvenno-političeskaja gazeta «Vechernjaja Ufa». Novosti ["Vechernyaya Ufa" social-political newspaper. News]. — URL: <http://vechufa.ru/news/> (accessed: 06.08.2022). [in Russian]
11. Pashinjan I.A. Kontent-analiz kak metod issledovanija: dostoinstva i ogranichenija [A Content Analysis as a Research Method: Strengths and Limitations] / I.A. Pashinjan // Nauchnaja periodika: problemy i reshenija [Scientific Periodicals: Problems and Solutions]. — 2012. — № 3. — P. 13-18. — URL: <https://cyberleninka.ru/article/n/kontent-analiz-kak-metod-issledovaniya-dostoinstva-i-ogranicheniya> (accessed: 08.05.2023). [in Russian]
12. Poljakova E.V. Massmedijnij diskurs: specifika i osobennost' [Mass-media Discourse: Specificity and Peculiarity] / E.V. Poljakova, N.A. Sadykova // Filologičeskie nauki: sostojanie, perspektivy, novye paradigmy issledovanij: Materialy IV Mezhdunarodnoj nauchnoj konferencii [Philological Sciences: State, Prospects, New Paradigms of Research: Proceedings of the IV International Scientific Conference] / chief ed. A.R. Muhtarullina. — Ufa: Publishing Center of BashSU, 2021. — P. 118-121. — DOI: 10.33184/fnspni-2021-11-18.17. [in Russian]
13. MaxQDA 2022 User Manual // MAXQDA. — URL: <https://www.maxqda.com/help-mx22/welcome> (accessed: 06.05.2023). [in Russian]
14. Studencov O.R. Kontent-analiz onlajn-versij informacionnykh programm na chuvashskom jazyke gosudarstvennoj teleradiokompanii «Chuvashija» (obzor) [A Content Analysis of Online Versions of News Programmes in Chuvash Language of the State TV and Radio Company "Chuvashia" (review)] / O.R. Studencov, G.V. Horas'kina, L.A. Vasil'eva // Mediasreda [Media Environment]. — 2018. — № 13. — P. 212-216. [in Russian]