

Научная статья / Research Article

УДК 004.89

DOI: 10.14258/SSI(2024)1-02.

О применимости тематического моделирования для определения акторов социально-политической мобилизации с использованием low-code аналитических платформ

Иван Юрьевич Степанов

Кемеровский государственный университет, Кемерово, Россия,
zextel1995@gmail.com, <https://orcid.org/0000-0002-7938-8049>

Елена Анатольевна Кранзеева

Кемеровский государственный университет, Кемерово, Россия,
elkranzeeva@mail.ru, <https://orcid.org/0000-0003-2577-9017>

Евгений Васильевич Головацкий

Кемеровский государственный университет, Кемерово, Россия,
xomaik@rambler.ru, <https://orcid.org/0000-0002-8485-5852>

Инна Вениаминовна Донова

Кемеровский государственный университет, Кемерово, Россия,
idonova@gmail.com, <https://orcid.org/0000-0001-9370-5402>

Анна Леонидовна Бурмакина

Кемеровский государственный университет, Кемерово, Россия,
anna-sidjakina@rambler.ru, <https://orcid.org/0000-0003-3087-4393>

Аннотация. Рассматривается современный подход к анализу социально-политических процессов, основанный на предположении, что тематическое моделирование, применяемое посредством low-code платформ, может повлиять на общее качество проводимого аналитиками исследования в целях выявления ключевых акторов и динамики социально-политических процессов. Авторы демонстрируют, как тематическое моделирование может выявлять взаимосвязи и тренды, недоступные при использовании традиционных методов анализа, и предлагают объединенную методологию, позволяющую исследователям социальных наук применять эти инструменты для более глубокого понимания механизмов социально-политической мобилизации. В работе представлен кейс-стади, демонстрирующий, как тематическое моделирование может выявить скрытые связи между различными акторами и их вклад в динамику мобилизации.

Ключевые слова: тематическое моделирование, кластеризация, социально-политическое взаимодействие, мобилизация, аналитические платформы, компьютерная социология

Финансирование: работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ (проект FZSR-2023-0006 «Сетевая социально-политическая мобилизация в регионах ресурсного типа Сибири: исследовательские возможности инструментов data-mining»).

Для цитирования: Степанов И.Ю., Кранзеева Е.А., Головацкий Е.В., Донова И.В., Бурмакина А.Л. О применимости тематического моделирования для определения акторов социально-политической мобилизации с использованием low-code аналитических платформ // Society and Security Insights. 2024. Т. 7, № 1. С. 27–39. doi: 10.14258/ssi(2024)1-02.

The Applicability of Topic Modeling to Identify Actors of Socio-Political Mobilization using Low-Code Analytical Platforms

Ivan Yu. Stepanov

Kemerovo State University, Kemerovo, Russia,
zextel1995@gmail.com, <https://orcid.org/0000-0002-7938-8049>

Elena A. Kranzeeva

Kemerovo State University, Kemerovo, Russia,
elkranzeeva@mail.ru, <https://orcid.org/0000-0003-2577-9017>

Evgenyi V. Golovatskyi

Kemerovo State University, Kemerovo, Russia,
xomaik@rambler.ru, <https://orcid.org/0000-0002-8485-5852>

Inna V. Donova

Kemerovo State University, Kemerovo, Russia,
idonova@gmail.com, <https://orcid.org/0000-0001-9370-5402>

Anna L. Burmakina

Kemerovo State University, Kemerovo, Russia,
anna-sidjakina@rambler.ru, <https://orcid.org/0000-0003-3087-4393>

Abstract. The article presents a contemporary perspective on the analysis of socio-political processes, grounded in the premise that the application of topic modeling through low-code platforms can substantially enhance the quality of research performed by analysts. This enhancement is particularly significant in identifying the pivotal actors and the evolving dynamics within socio-political processes. The authors argue that topic modeling, a relatively novel approach compared to traditional methods, is capable of uncovering relationships and trends that might otherwise remain obscured. In advocating for this approach, the paper proposes an integrated methodology. This methodology is designed to empower researchers in the social sciences, en-

abling them to effectively utilize these innovative tools. The objective is to deepen their comprehension of the underlying mechanisms that drive socio-political mobilization. To substantiate their argument, the authors present various case studies. These case studies demonstrate the effectiveness of topic modeling in revealing otherwise hidden connections among various actors. Additionally, they illustrate how topic modeling sheds light on the contributions of these actors to the dynamics of mobilization. This approach represents a significant advancement in the field, offering new insights and a more nuanced understanding of complex socio-political landscapes.

Keywords: topic modeling, clustering, socio-political interaction, mobilization, analytical platforms, computer sociology

Financing: The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation (project FZSR-2023-0006 «Network socio-political mobilization in the regions of the resource type of Siberia: research capabilities of data-mining tools»).

For citation: Stepanov, I.Yu., Kranzeeva, E.A., Golovatsky, E.V., Donova, I.V., Burmakina, A.L. (2024). The Applicability of Topic Modeling to Identify Actors of Socio-Political Mobilization using Low-Code Analytical Platforms. *Society and Security Insights*, 7(1), 27–39. (In Russ.). doi: 10.14258/ssi(2024)1-02.

Введение

Тематическое моделирование — это технология статистического анализа текстов для автоматического выявления тематики в больших коллекциях документов (Arishev, Vorontsov, 2020). Тематическое моделирование не претендует на понимание смысла текста, однако оно способно отвечать на вопросы «о чем этот текст» или «какие общие темы имеет эта пара текстов».

Одни из ранних моделей были ориентированы на использование дистрибутивной семантики как метода предсказания/выявления определенного контекста. Наибольшую популярность до сих пор сохранили такие модели, как Word2Vec и FastText. С развитием нейронных сетей популярность получали архитектуры, использующие рекуррентные нейроны, такие как RNN, GRU и LSTM, однако их вычислительная сложность остается достаточно высокой и по сей день (Greff et al., 2017).

Одним из ранних алгоритмов тематического моделирования является LSA — Latent Semantic Analysis. LSA был запатентован в 1988 г. Скоттом Дирвестером и др. (Deerwester et al, 1988). В 1999 г. Томас Хофманн предложил модель вероятностного латентного семантического анализа (PLSA) (Hofmann, 2000). Входными данными модели являются матрица документов-слов и количество кластеров (тем). На выходе же модель выдает распределение слов в темах и тем в документах. Улучшением предыдущего подхода стало латентное размещение Дирихле (Latent Dirichlet Allocation, LDA) — одна из самых известных и широко применяемых тематических моделей (Черкасов, 2020).

Появление альтернативного подхода — аддитивной регуляризации тематических моделей (Additive Regularization for Topic Modeling, ARTM) расширяет базовую модель PLSA общим механизмом регуляризации. Большим преимуще-

ством ARTM является его гибкость. Обучив пробную модель в виде стандартной LDA или ARTM без регуляризаторов, исследователь может сформулировать, чего не хватает в текущей модели (Сухарева, Воронцов, 2019).

Развитие интернет-технологий привело к тому, что в свободном доступе появилось большое количество цифровых следов пользователей: социально-демографические признаки, контактные данные, но главное — большое количество текстовой информации, позволяющей не только составить «портрет» конечного пользователя, но и решать большое количество различных задач.

Наибольшей популярностью в России пользуется социальная сеть VK («ВКонтакте»), она ежедневно генерирует посредством пользователей огромные массивы данных, большую часть из которых возможно собрать и проанализировать.

Однако проведение большинства state-of-the-art исследований предъявляет повышенные требования к исследователю, в частности — уровню владения исследовательским инструментарием, языками программирования и средствами и техниками визуализации данных для составления отчетной документации. Одним из решений, которое способно снизить «порог входа» исследователя, может являться использование low-code аналитических платформ.

Как правило, такие платформы позволяют в режиме визуального проектирования, при помощи встроенных «блоков», описать шаги, необходимые для проведения исследования.

Аналитические платформы как особый подвид инструментов проведения исследований стали легко доступны давно. В зависимости от сложности задачи и количества имеющихся в распоряжении исследователя (или лаборатории) ресурсов могут быть использованы такие решения, как Orange Data Mining, KNIME Analytical Platform или PolyAnalyst. Они являются универсальными с точки зрения как междисциплинарности, так и мультимодальности. Так, при помощи Orange Data Mining и его отдельных инструментов экономисты и финансисты изучают трансформацию лексики годовых отчетов корпораций (Михненко, 2022), исследуются недобросовестные практики на рынке ценных бумаг (Анашкина, Колесова, 2020), анализируется отраслевая специфика налоговых доходов российских регионов (Колбягина, 2022). К примеру, археологи классифицируют реконструкции и артефакты по фотографиям их силуэтов (инверсия информации) (Гончаренко, 2022). Платформа PolyAnalyst нашла применение в исследованиях социологов и политологов — с целью анализа информационного поля, коммуникаций, изучения реакции на события и тренды путем семантического анализа текстов из социальных сетей (Лебёдкина, Александрова, Орлова, 2021). Представители филологических и географических наук на материале сетевых текстов (включая как соцсети, так и RSS-каналы) изучают новостные дискурсы (Горчакова и др., 2021), дискурсы компьютерной безопасности (Исаева, 2022), а также анализ ментальных туристско-рекреационных пространств (Коньшев, 2022).

В условиях динамичной трансформации социально-политической ситуации в современной России требуются комплексные обновленные приемы изучения взаимодействия со стороны власти и населения, однако возможности современных ис-

следовательских методик ограничены (Мирошниченко, 2016). Недостаточно изученной стороной социальных и политических взаимодействий гражданских сообществ являются новые приемы и формы мобилизации по отношению к их акторам, что актуализирует использование гибридных методов анализа (Головацкий и др., 2018).

Социально-политическая мобилизация, рассматриваемая с учетом возможностей применения исследователями тематического моделирования, может быть представлена в формате уникальных слоев повседневной коммуникации сетевых участников, доступных для применения инструментов аналитических платформ. Предлагается составление каталогов событий, изначальные схемы (переменные отбора, анализа) могут быть изменены в процессе сбора и обработки информации в силу ряда причин, недоступности данных или специфики форм мобилизационной активности, ангажированности рассматриваемых участников и пр. В этой связи событийный анализ сетевого коммуникативного многообразия требует достаточной строгости в формулировках и четко следования «маршруту исследований», а с другой стороны — определенной гибкости в работе с данными (Семенов, 2018) и интерпретации полученных результатов.

Множество субъектов в условиях социальных взаимодействий подразумевают также наличие n -го множества следов деятельности, включая и цифровое пространство во всем своем многообразии каналов и форматов цифровых данных. Массивы big data в силу своей структурной сложности подразумевают присутствие в данных множествах различного рода ошибок (багов, ворнингов, ошибок сегментации и т.п.).

В этой связи исследователи отмечают необходимость учитывать три типа возможных систематических ошибок (bias): ошибка отбора информации (selection bias), связанная с избирательным информационным освещением событий; ошибка описания (description bias), связанная с неточной и/или неполной информацией о событиях в источниках; исследовательская ошибка (research bias), связанная с неточностью ручного или автоматического кодирования (подготовки данных) (Hutter, Porta, 2014).

Анализ сетевых участников (акторов) мобилизации позволяет обнаружить как минимум две структуры, позволяющие строить акторно-сетевые модели социальной и политической мобилизации в виртуальной среде. Во-первых, «реальные» акторы (субъекты физического пространства), коммуникативная деятельность которых способна мобилизовать гражданские и организационные инициативы, поддерживать актуальную повестку дня, использовать индивидуально-личностные ресурсы и пр. Во-вторых, речь может идти о сетевых акторах социальной и политической коммуникации (их вреда формирования и продуктивного взаимодействия в сетевом пространстве) (Володенков, Митева, 2016; Алексеев, 2019).

Методы исследования

В рамках данной статьи применяются смешанные методы исследования для анализа эффективности тематического моделирования в определении акторов социально-политической мобилизации при использовании low-code аналитических платформ. Исследование включает в себя качественный анализ, основанный

на систематическом обзоре литературы, и количественный анализ, использующий данные из различных социально-политических контекстов, представленных в виде слабоструктурированных текстовых корпусов с преобладающим русским языком и сообщениями на иных языках. Тематическое моделирование, реализованное в среде аналитической платформы PolyAnalyst, позволяет выявить ключевые паттерны и тенденции, связанные с акторами мобилизации, коим выступает круг пользователей VK, проявивших наибольшую активность в роли комментаторов и создателей постов. Для проверки надежности и валидности методов тематического моделирования используются различные наборы данных для использования в алгоритме кросс-валидации на датасетах с предварительно заданной тематикой. В качестве метрик качества моделей выступали когерентность и перплексия. Кроме того, анализируется способность этих методов идентифицировать скрытые взаимосвязи между акторами и влиять на динамику мобилизации. Такой подход позволяет оценить потенциал и ограничения использования low-code платформ в социально-политических исследованиях.

Результаты исследования

В качестве примера, демонстрирующего особенности и преимущества использования описываемых аналитических платформ, можно рассмотреть задачу тематического моделирования в среде аналитической платформы PolyAnalyst над массивом данных — сообщений, доступных в публичном доступе, среди тематических групп сообществ «ВКонтакте» в Тюменской и Кемеровской областях, как в регионах, схожих по ресурсному типу, территориальной общности возможностей и, возможно, притягательности проблем. Под тематическими группами подразумеваются сообщества, объединенные по территориальному признаку, в которых доступны функции комментирования и/или создания обсуждений пользователями, а в качестве векторной тематики выступает социальная сфера. Сбор производился с использованием публичного API VK методом веб-скраппинга, позволяющего сохранять не только текст сообщений, но и сопутствующие метаданные (время публикации, публичный id автора и др.). В качестве референса было проведено тематическое моделирование с использованием модели LDA из модуля gensim ввиду поддержки параллельных вычислений на многоядерных вычислительных машинах, результаты которого приведены в виде совокупности гистограмм (рис. 1).

В рамках аналитической платформы работа представляется не в виде «кода», а в виде визуальных блоков, которые комбинируются между собой, формируя pipeline исследования. Для описанной ранее задачи была сформирована следующая структура исследования в среде аналитической платформы PolyAnalyst (рис. 2).

В качестве удобного дополнения PolyAnalyst позволяет сформировать метатему, объединив некоторое множество тем по соединяющему их признаку. Например, под названием «Мусорная категория» скрываются все нерелевантные к социально-политической мобилизации темы. Тем не менее можно сделать предположение, что присутствие элементов этого «мусорного» массива информации способно выступать точками сборки/пересечения рассматриваемых сетевых сообществ (рис. 3).

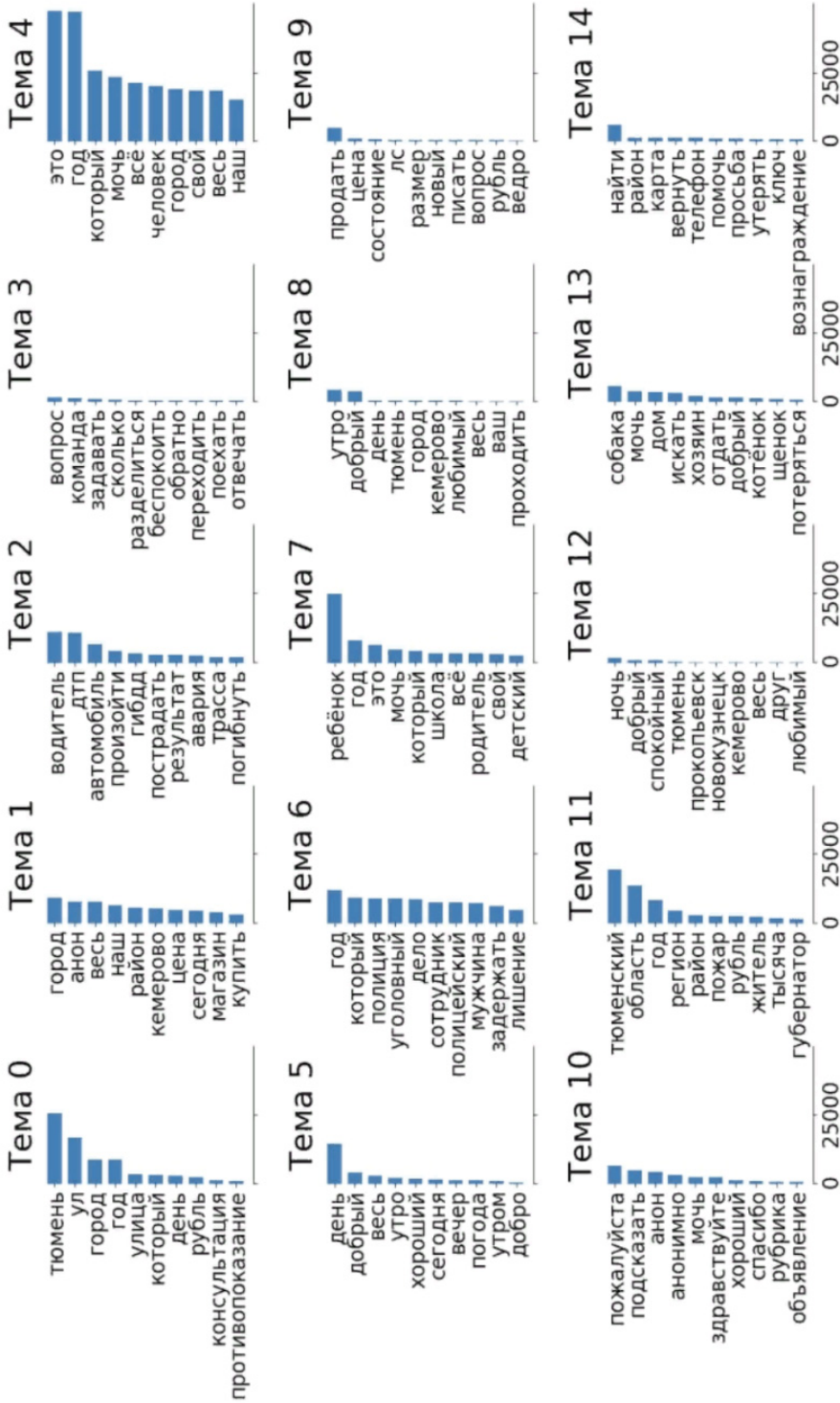


Рисунок 1 — Результаты тематического моделирования с использованием модели LADA (Кузбасс, Тюменская область, декабрь 2022 г., исходный массив ~2 млн строк, ручной отбор тематических групп).

Figure 1 — Results of thematic modeling using the LADA model (Kuzbass, Tyumen region, December 2022, initial array of ~2 million lines, manual selection of thematic groups).

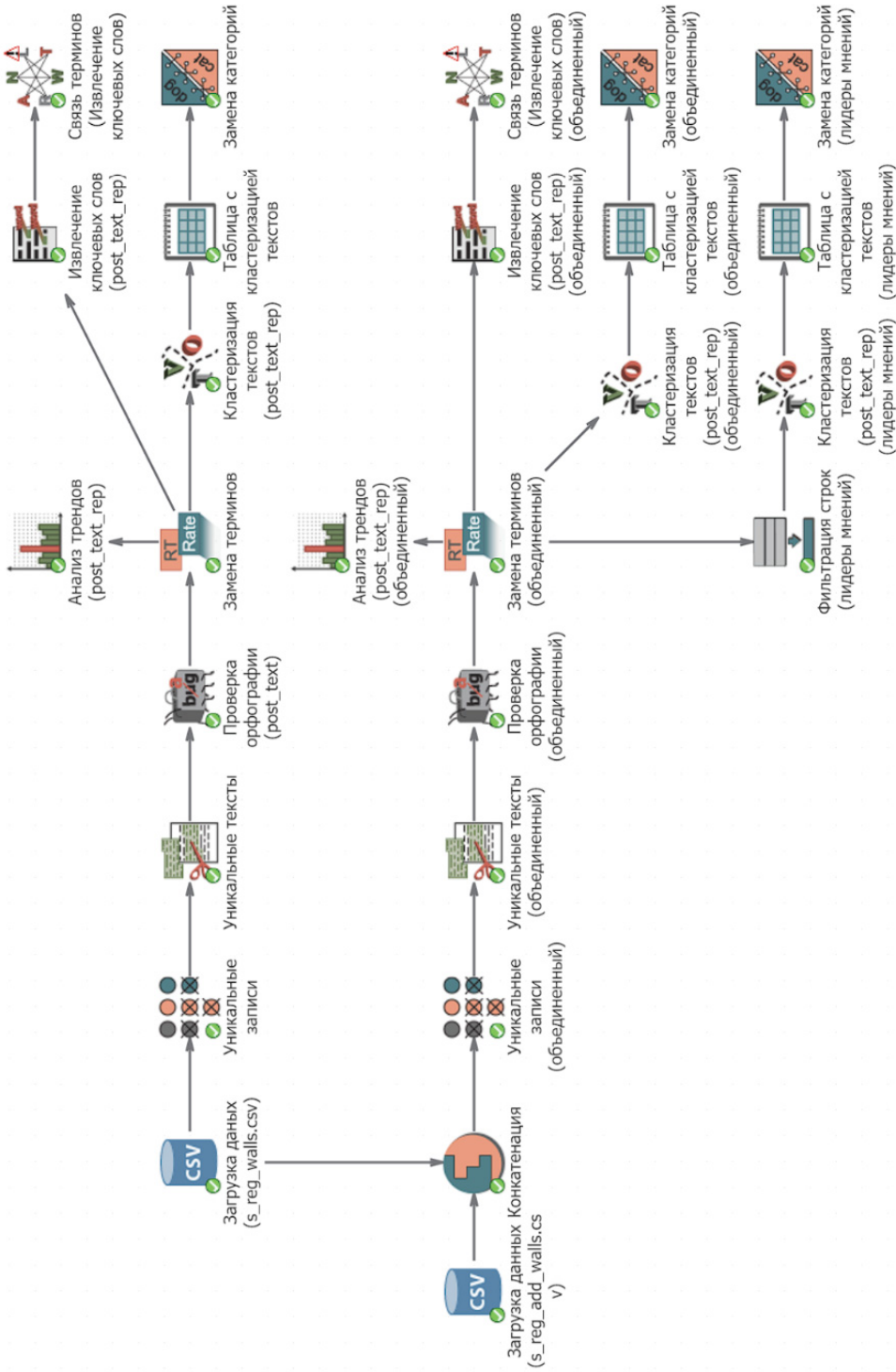


Рисунок 2 — Структура этапов сборки исследования в среде PolyAnalyst.
Figure 2 — The structure of the research assembly stages in the PolyAnalyst platform.



Рисунок 3 — Укрупненная тематическая группировка повестки сетевых сообществ в период с января по декабрь 2022 г. (Кузбасс, Тюменская область, ручная сортировка тематических групп).

Figure 3 — Enlarged thematic grouping of the agenda of network communities in the period from January to December 2022 (Kuzbass, Tyumen region, manual sorting of thematic groups).

Обсуждение и заключение

Инструменты тематического моделирования позволяют получать выделенные темы для анализа социальной и политической мобилизации сетевых участников по актуальной обсуждаемой в сообществах повестке. Инструменты тематического моделирования позволяют получать «длинные связи», неочевидные для исследователей при первичном рассмотрении.

Представленное исследование демонстрирует значительный потенциал тематического моделирования в анализе социально-политических процессов. Методы тематического моделирования, особенно в сочетании с low-code аналитическими платформами, открывают новые перспективы для исследователей в изучении сложных и динамичных сетевых сообществ. Примененные методы позволяют эффективно выявлять ключевые темы и акторов социально-политической мобилизации, предоставляя более глубокое понимание механизмов их взаимодействия.

Отдельно стоит подчеркнуть значимость интеграции тематического моделирования и low-code аналитических платформ в исследовательскую практику социальных наук. Это открывает новые возможности для глубокого и многоаспектного изучения социально-политических явлений, способствуя развитию методов анализа и понимания сложных социальных процессов.

СПИСОК ИСТОЧНИКОВ

- Алексеев М. С. Технологии девиантного поведения в сетевом пространстве // Молодой ученый. 2019. № 49 (287). С. 539–542.
- Анашкина М. В., Колесова И. В. Применение datamining в процессе выявления недобросовестных практик на рынке ценных бумаг (на основе кейса) // Финансовые исследования. 2020. №3 (68). С. 16–26.
- Володенков С. В., Митева В. В. Сетевые информационные войны в современных условиях: основные акторы и стратегии // PolitBook. 2016. №3.
- Гончаренко Ю. Д. К возможности группировки керамики в среде для анализа данных ORANGE (на основе материалов самосдельского городища) // Материалы Всероссийской (с международным участием) археологической студенческой конференции (Астрахань, 01–03 февраля 2022 г.). Астрахань, 2022. С. 185–187.
- Горчакова О. Ю., Ларионова А. В., Александрова Ю. К., Петров Е. Ю. Особенности новостного дискурса социально-политической тематики на примере региональных новостных пабликов в социальной сети «ВКонтакте» // Филология: научные исследования. 2021. № 3. С. 1-17. DOI: 10.7256/2454-0749.2021.3.35234
- Исаева Е. В. Тематическое моделирование в дискурсе компьютерной безопасности: исследование на примере публикаций информационных бюллетеней и новостных лент // Вестник Пермского университета. Российская и зарубежная филология. 2022. №14 (2). <https://doi.org/10.17072/2073-6681-2022-2-18-26>
- Колбягина Л. А. Кластерный анализ отраслевой структуры налоговых доходов на региональном уровне // Известия БГУ. 2022. № 3. С. 493–500.
- Коньшев Е. В. Методика изучения ментального туристско-рекреационного пространства по отзывам туристов (на примере Кировской области) // Вестник Московского университета. 2022. Сер. 5: География. № 5. С. 16–28.
- Кранзеева Е. А., Головацкий Е. В., Орлова А. В. Социальное и политическое взаимодействие местных сообществ региона в условиях реактивных отношений: кейсы благоустройства городского пространства // Вестник Томского государственного университета. 2021. № 464. С. 81–90. DOI: 10.17223/15617793/464/10
- Лебедкина Н. С., Александрова Ю. К., Орлова В. В. Анализ миграционных потоков молодежи на территории субъектов Российской Федерации // Векторы благополучия: экономика и социум. 2021. № 2 (41). С. 57–72. doi:10.18799/26584956/2021/3(42)/1089
- Мирошниченко И. В. Сетевая публичная политика и управление. М.: Аргмак-Медиа, 2016. 296 с.
- Михненко П. А. Data mining как инструмент мультимодальной бизнес-аналитики: трансформация лексики годовых отчетов госкорпорации «Ростех» // Вестник Российского экономического университета им. Г. В. Плеханова. 2022. № 6. С. 126–136. doi:10.21686/2413-2829-2022-6-126-136
- Семенов А. Событийный анализ протестов как инструмент изучения политической мобилизации // Социологическое обозрение. 2018. Т. 17, № 2. С. 317–341. doi: 10.17323/1728-192X-2018-2-317-341

Сухарева А. В., Воронцов К. В. Построение полного набора тем вероятностных тематических моделей // Интеллектуальные системы. Теория и приложения. 2019. Т. 23? № 4. С. 7–23.

Черкасов Е. И. Сравнение алгоритмов тематического моделирования при определении тематик постов людей в социальной сети «ВКонтакте» // Евразийский Союз Ученых. 2020. № 6–2 (75). С. 45–49.

Apishev M., Vorontsov K. Learning topic models with arbitrary loss // Proceedings of the XXth Conference of Open Innovations Association FRUCT. 2020. Vol. 26. P. 1–8. DOI: 10.23919/FRUCT48808.2020.9087559

Deerwester S., et al, Improving Information Retrieval with Latent Semantic Indexing, Proceedings of the 51st Annual Meeting of the American Society for Information Science 25, 1988. Pp. 36–40.

Golovatsky E., Kranzeeva E., Orlova A., Burmakina A. Social Practices of Mobilizing Population Initiatives: Prospects for Hybrid Methodology // International Conference on Communicative Strategies of Information Society (CSIS 2018). Advances in Social Science, Education and Humanities Research, 2018. Vol. 289. Pp. 8–13. doi: 10.2991/csis-18.2019.2

Greff K., Srivastava R. K., Koutník J., Steunebrink B. R., and J. Schmidhuber, «LSTM: A Search Space Odyssey,» in IEEE Transactions on Neural Networks and Learning Systems, Vol. 28, No 10. Pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.

Hofmann Th. Learning the Similarity of Documents: an information-geometric approach to document retrieval and categorization, Advances in Neural Information Processing Systems 12. MIT Press, 2000. Pp. 914–920.

Hutter S. 2014. Protest Event Analysis and Its Offspring // della Porta D. (ed.). Methodological Practices in Social Movement Research. Oxford: Oxford University Press. Pp. 335–367.

REFERENCES

Alekseev, M.S. (2019). Technologies of deviant behavior in the network space. *Young scientist*, 49(287), 539–542. (In Russ.).

Anashkina, M. V., & Kolesova, I. V. (2020). Application of datamining in the process of identifying unfair practices in the securities market (based on the case). *Finansovye issledovaniya*, 3(68), 16–26. (In Russ.).

Volodenkov S. V., Miteva V. V. (2016). Network information wars in modern conditions: main actors and strategies. *PolitBook*.3. (In Russ.).

Goncharenko, Yu. D. (2022). To the possibility of grouping ceramics in the environment for analysis of ORANGE data (based on the materials of the fort). In: D. V. Vasil'ev (Ed.), *Materials of the All-Russian (with international participation) archaeological student conference (Astrakhan, February 01–03, 2022)*. (pp. 185–187). Astrahan. (In Russ.).

Gorchakova, O. Yu., Larionova, A. V., Aleksandrova, Yu. K., & Petrov, E. Yu. (2021). Features of the social and political discourse on the example of regional news publics in the social network «VKontakte». *Filologiya: nauchnye issledovaniya*, 3, 1–17. doi:10.7256/2454-0749.2021.3.35234 (In Russ.).

- Isaeva, E. V. (2022). Topic modelling in the computer security discourse: case study on newsletters and news feeds. *Vestnik Permskogo universiteta. Rossijskaya i zarubezhnaya filologiya*, 14(2). doi:10.17072/2073-6681-2022-2-18-26 (In Russ.).
- Kolbyagina, L. A. (2022). Cluster analysis of the sectoral structure of tax revenues at the regional level. *Izvestiya BGU*, 3, 493–500. (In Russ.).
- Konyshov, E. V. (2022). Methods of studying the mental tourist-recreational space according to the reviews of tourists (on the example of Kirov region). *Vestnik Moskovskogo universiteta. Ser. 5: Geografiya*, 5, 16–28. (In Russ.).
- Kranzeeva, E. A., Golovackij, E. V., & Orlova, A. V. (2021). Social and political interaction of local communities of the region in reactive relations: cases of urban improvement. *Vestnik Tomskogo gosudarstvennogo universiteta*, 464, 81–90. doi:10.17223/15617793/464/10.
- Lebedkina, N. S., Aleksandrova, Yu. K., & Orlova, V. V. (2021). Analysis of migration flows of youth in the territory of the constituent entities of the Russian Federation. *Vektory blagopoluchiya: ekonomika i socium*, 2(41), 57–72. doi:10.18799/26584956/2021/3(42)/1089 (In Russ.).
- Miroshnichenko, I. V. (2016). *Network public policy and management*. Moscow : Argamak-Media. (In Russ.).
- Mihnenko, P. A. (2022). Data mining as a tool of multimodal business analytics: transformation of vocabulary of annual reports of the state corporation «RosTech». *Vestnik Rossijskogo ekonomicheskogo universiteta im. G. V. Plehanova*, 6, 126–136. doi:10.21686/2413-2829-2022-6-126-136 (In Russ.).
- Semenov, A. (2018). Event analysis of protests as a tool to study political mobilization. *Sociologicheskoe obozrenie*, 17(2), 317–341. doi:10.17323/1728-192X-2018-2-317-341 (In Russ.).
- Suhareva, A. V., & Voroncov, K. V. (2019) Building a complete set of probabilistic thematic model topics. *Intellektual'nye sistemy. Teoriya i prilozheniya*, 23(4), 7–23. (In Russ.).
- Cherkasov, E. I. (2020). Comparison of algorithms of thematic modeling in determining subjects of posts of people in the social network «VKontakte». *Evrazijskij Soюз Uchenyh*, 6–2(75), 45–49. (In Russ.).
- Apishev, M., & Vorontsov, K. (2020). Learning topic models with arbitrary loss. In *Proceedings of the XXth Conference of Open Innovations Association FRUCT*, Vol. 26 (pp. 1–8). DOI: 10.23919/FRUCT48808.2020.9087559
- Deerwester, S., et al. (1988). Improving Information Retrieval with Latent Semantic Indexing. *Proceedings of the 51st Annual Meeting of the American Society for Information Science*, 25, 36–40.
- Golovatsky, E., Kranzeeva, E., Orlova, A., & Burmakina, A. (2018). Social Practices of Mobilizing Population Initiatives: Prospects for Hybrid Methodology. In: *International Conference on Communicative Strategies of Information Society (CSIS 2018). Advances in Social Science, Education and Humanities Research*, 289 (pp. 8–13). <https://doi.org/10.2991/csis-18.2019.2>
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning*

Systems, 28(10), 2222-2232. doi:10.1109/TNNLS.2016.2582924

Hofmann, T. (2000). Learning the similarity of documents: an information-geometric approach to document retrieval and categorization. In: *Advances in Neural Information Processing Systems 12* (pp. 914–920). MIT Press.

Hutter, S. (2014). Protest event analysis and its offspring. In D. della Porta (Ed.), *Methodological Practices in Social Movement Research* (pp. 335–367). Oxford : Oxford University Press.

СВЕДЕНИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Иван Юрьевич Степанов — ассистент кафедры цифровых технологий института Цифры, Кемеровский государственный университет, г. Кемерово, Россия.

Ivan Yu. Stepanov — Assistant Professor, Department of Digital Technologies, Institute of Digitalization, Kemerovo State University, Kemerovo, Russia.

Елена Анатольевна Кранзеева — д-р социол. наук, доцент, зав. кафедрой социологических наук, Кемеровский государственный университет, г. Кемерово, Россия.

Elena A. Kranzeeva — Dr. Sci. (Sociology), Head of the Department of Sociological Sciences, Kemerovo State University, Kemerovo, Russia.

Евгений Васильевич Головацкий — д-р социол. наук, доцент, профессор кафедры социологических наук, Кемеровский государственный университет, г. Кемерово, Россия.

Evgeny V. Golovatsky — Dr. Sci. (Sociology), Professor of the Department of Sociological Sciences, Kemerovo State University, Kemerovo, Russia.

Инна Вениаминовна Донова — канд. экон. наук, доцент кафедры менеджмента имени И.П. Поварича Института экономики и управления, Кемеровский государственный университет, г. Кемерово, Россия.

Inna V. Donova — Cand. Sci. (Economics), Associate Professor, I.P. Povarich Department of Management, Institute of Economics and Management, Kemerovo State University, Kemerovo, Russia.

Анна Леонидовна Бурмакина — старший преподаватель кафедры социологических наук, Кемеровский государственный университет, г. Кемерово, Россия.

Anna L. Burmakina — Senior Lecturer, Department of Sociological Sciences, Kemerovo State University, Kemerovo, Russia.

Статья поступила в редакцию 12.01.2024;
одобрена после рецензирования 15.02.2024;
принята к публикации 20.02.2024.

The article was submitted 12.01.2024;
approved after reviewing 15.02.2024;
accepted for publication 20.02.2024.