

В. КОНЧА*

**СКВОЗЬ ПРИЗМУ *BERT*:
ТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
ТРАНСФОРМАЦИИ КОЛЛЕКТИВНЫХ ИДЕНТИЧНОСТЕЙ
УЧАСТНИКОВ ДВИЖЕНИЯ *BLACK LIVES MATTER*
ПОД ВЛИЯНИЕМ КОНТРПРОТЕСТА¹**

Аннотация. Контрпротесты стали распространенным ответом на протестные движения по всему миру. Понимание этого процесса часто основывается на коллективных идентичностях, которые определяются в противовес аутгруппе – будь то власти или конкурирующим политическим силам. В контексте движения *Black Lives Matter (BLM)* первоначальный образ «другого» в основном приписывался властям. Однако появление контрпротестного движения *All Lives Matter (ALM)* усложнило эту картину – теперь образ «другого» приписывается также и участникам контрпротестного движения, внося новый слой сложности в динамику идентичностей и порождая вопросы о том, как эти идентичности трансформируются при столкновении с контрпротестом. Применяя новый подход в тематическом моделировании *BERT*, автор анализирует тематические сдвиги в постах участников протестного движения *BLM* в период с 2013 по 2014 г. в социальной сети *Twitter (X)*. Применение тематического моделирования с помощью *BERT (Bidirectional Encoder Representations from Transformers)* позволило провести детальный анализ онлайн-сообщений, улавливая контекстные зависимости и раскрывая сложности трансформации коллективной идентичности в ответ на внешний вызов – появление контрпротестного движения *ALM*. Настоящее исследование заполняет две важные лакуны в литературе. Во-первых, оно предлагает мето-

* **Конча Валерия**, аспирант, преподаватель Департамента политики и управления, младший научный сотрудник Международного центра изучения институтов и развития, Национальный исследовательский университет «Высшая школа экономики» (Москва, Россия), e-mail: vkoncha@hse.ru

¹ Статья подготовлена в результате проведения исследования в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ)

логическое улучшение, внедряя *BERT* для анализа динамической трансформации тем, что является шагом вперед по сравнению с доминирующим использованием *LDA (Latent Dirichlet allocation)*, который не улавливает контекстуальных нюансов. Во-вторых, оно решает содержательную задачу: исследует, как изменяются коллективные идентичности участников протестного движения *BLM* под влиянием контрпротестов, что является пробелом в исследовательской литературе. *BERT*-тематическое моделирование показало себя как мощный инструмент для изучения трансформации коллективных идентичностей, предлагая перспективную методологическую основу для дальнейших исследований на пересечении проблематики протеста, контрпротеста и коллективных идентичностей в цифровую эпоху.

Ключевые слова: BlackLivesMatter; AllLivesMatter; коллективная идентичность; модель *BERT*; протест; контрпротест; онлайн-сообщения.

Для цитирования: Конча В. Сквозь призму *BERT*: тематическое моделирование трансформации коллективных идентичностей участников движения Black Lives Matter под влиянием контрпротеста // Политическая наука. – 2025. – № 1. – С. 219–239. – DOI: <http://www.doi.org/10.31249/poln/2025.01.10>

Введение

Коллективная идентичность служит важным инструментом для мобилизации участников протестных движений, укрепления их солидарности и выражения протестных настроений. А. Мелуччи определяет коллективную идентичность как постоянно развивающееся чувство принадлежности к группе, сформированное общим опытом и коллективными действиями [Melucci, 1996]. В протестных движениях эта идентичность часто формируется в противовес внешнему образу «другого» или аутгруппе [Гулевич, 2019], такому, как, например, власть. Однако возникновение контрпротестных движений вносит дополнительные уровни сложности в процесс трансформирования идентичности участников протестного движения. Под контрпротестом следует понимать в широком смысле реакцию на протестное движение; это протестная активность, которая возникает в непосредственной близости от идеологически противоположного протеста и основана на коллективной идентичности для обеспечения статус-кво [Wood, 2020; Inata, 2021]. Зальд и Усим описывают взаимодействие протеста и контрпротеста как «танго мобилизации и демобилизации», подразумевая моменты времени, когда эти движения сталкиваются под давлением определенных событий [Zald, Useem, 1987]. Постоянная динамика взаимодействия протеста-контрпротеста заставляет движения не только отстаивать свои цели, но и реагировать на

противостоящие идеологии, тем самым трансформируя коллективную идентичность. Такое явление наблюдается во многих протестных движениях, основой которых выступают острые социальные и политические проблемы, которые зачастую поляризуют общество.

Известными примерами такого взаимодействия протестов и контрпротестов являются *Black Lives Matter (BLM)* и *All Lives Matter (ALM)*. Движение *BLM*, возникшее в 2013 г. для борьбы с «системным расизмом»¹ и насилием со стороны полиции в отношении чернокожих, изначально определяло свою коллективную идентичность как оппозицию властям, и, в частности, полиции относительно расовой дискриминации. Однако рост *ALM* как контрдвижения в 2014 г. привел к появлению новой антагонистической силы для *BLM*. *ALM* активно пропагандирует универсалистское послание, отвергающее расовую специфику заявлений *BLM* о том, что «жизни чернокожих важны» – оно выступает за то, что «все жизни» должны быть в равной степени приоритетными. Этот сдвиг заставил участников *BLM* переопределить свою коллективную идентичность не только с точки зрения их борьбы с «системным расизмом», но и в противовес идеологической позиции *ALM*. Таким образом, в результате появления контрпротестного движения для участников протеста *BLM* сформировалось две аутгруппы, относительно которых они отстраивают свою коллективную идентичность – полиция и участники контрпротестного движения. Так коллективная идентичность не сводится к противопоставлению другой группе, а является отражением «соотношения» с другими группами [Малинова, 2015, с. 166–167]. Более того, коллективные идентичности участников протеста не статичны и постоянно трансформируются [Tarrow, 2011].

Коллективная идентичность представляет собой динамическую конструкцию, которая формируется через дискурсивные процессы, включая использование и распространение определенных смысловых паттернов, или тем. Темы выступают как инструмент артикуляции ценностей, целей и границ группы, помогая ее участникам осознавать свою принадлежность и отстраивать свою идентичность относительно образа «другог». В контексте протестных и контрпротестных движений смысловые паттерны становятся

¹ В контексте движения *BLM* под системным расизмом понимается то, что чернокожее население хуже интегрировано в общественные институты в целом по сравнению с остальным населением.

ключевыми элементами, через которые коллективная идентичность не только выражается, но и трансформируется в ответ на внешние вызовы. Анализ выявленных тем позволяет не просто описывать характер дискурса, но и проследить, как они способствуют трансформации коллективных идентичностей, усиливают солидарность внутри группы и задают рамки ее взаимодействия с внешними акторами.

Изучение трансформации коллективной идентичности опирается на различные методы, от традиционных до современных, от качественных до количественных.

Арсенал качественных методов изучения трансформации коллективной идентичности достаточно богат. Холлан и др., например, использовали этнографический метод для изучения формирования и эволюции коллективной идентичности на примерах трех протестных движений в Шотландии, Непале и Канаде [Holland et al., 2008]. Штюрмер и Саймон использовали метод телефонного интервью, чтобы запечатлеть субъективный опыт участников протеста в Германии и исследовать, как изменилась их коллективная идентичность в течение трех лет [Stürmer, Simon, 2004]. Другая группа исследователей сосредоточила свое внимание на изучении трансформации коллективной идентичности через анализ дискурса [Koller, 2012; Shahin, Ng, 2021]. Исследователи анализируют речи, манифесты, сообщения в социальных сетях, лозунги и другие формы коммуникации, чтобы понять, как артикулируется такая идентичность и как она развивается с течением времени. Анализ дискурса особенно полезен для изучения трансформации идентичности, поскольку он позволяет исследователям отслеживать, как язык движения меняется в ответ на внешние вызовы или внутренние события.

Нарративный анализ – это еще один качественный метод, используемый для изучения того, как коллективные идентичности фреймируются и трансформируются в протестных движениях. Этот подход фокусируется на историях участников протестного движения и их опыте, исследуя, как эти нарративы способствуют построению общей идентичности. Поллетта утверждала, что нарративы играют решающую роль в формировании коллективной идентичности, поскольку они предоставляют связную историю, которая объясняет цель, ценности и задачи движения [Polletta, 2006]. Нарративный анализ позволяет исследователям изучать, как идентичность конструируется посредством повествования и как эти истории развиваются по мере развития движения. Например,

Бенфорд и Сноу в своей работе исследуют, как социальные движения используют фреймы и нарративы для создания коллективной идентичности [Benford, Snow, 2000]. Они утверждают, что фреймирование (как форма нарративного анализа) играет ключевую роль в конструировании и поддержании коллективной идентичности участников протестов, а также в мобилизации ресурсов и легитимации протестных действий, выделяя такие фреймы как диагностические, прогностические и мотивационные. В работе они рассматривают, как движения создают и изменяют нарративы в зависимости от внешних вызовов, что позволяет движениям адаптироваться и продолжать функционировать.

Хотя качественные методы, такие как анализ дискурса, анализ фреймов и нарративный анализ, дают более глубокое понимание того, как конструируются и трансформируются коллективные идентичности, они не лишены ограничений. Одной из ключевых проблем является масштабируемость – их крайне тяжело применить к крупномасштабным цифровым движениям, которые генерируют огромные объемы данных. Кроме того, качественные методы в значительной степени зависят от интерпретации исследователя, что может привести к субъективности и предвзятости в анализе.

В контексте современных протестных движений, которые все чаще организуются и мобилируются через цифровые платформы, растет потребность в методах, которые могут обрабатывать большие наборы данных, сохраняя при этом глубину анализа, предлагаемую качественными подходами. Это привело к разработке смешанных подходов, которые сочетают качественные методы с вычислительными методами, такими как тематическое моделирование и анализ настроений (англ. *sentiment analysis*). Интегрируя эти методы, исследователи могут анализировать трансформацию идентичности в масштабе, при этом по-прежнему фиксируя богатый, нюансированный опыт участников протеста.

Анализ социальных сетей, с другой стороны, применялся для изучения структурных связей между участниками цифровых движений. Беннетт и Сегерберг исследовали, как децентрализованные протестные движения без четкого лидера, такие как *Occupy Wall Street* и «арабская весна», формировали гибкие, индивидуализированные коллективные идентичности через социальные сети [Bennett, Segerberg, 2012]. Однако, хотя этот подход отображает связи участников, он не углубляется в тематическое содержание, формирующее их идентичности.

В последнее время методы машинного обучения (в частности, анализ настроений) стали применяться для изучения трансформации идентичности в онлайн – социальных движениях. Например, группа исследователей провела анализ настроений участников протеста *BLM* на корпусе онлайн сообщений из *Twitter** [Peng et al., 2022]. Авторы выявили ряд тем и событий, которые были связаны с положительными общественными настроениями, такими как социальная справедливость, которые были связаны с негативными настроениями, такими как жестокость полиции.

Латентное распределение Дирихле (*LDA*) стало одним из наиболее широко используемых методов тематического моделирования в социальных науках для изучения больших текстовых наборов данных. *LDA* особенно полезно для выявления скрытых тем в протестных движениях, позволяя исследователям отслеживать, как дискурс развивается с течением времени. В контексте трансформации коллективной идентичности *LDA* использовалось для выявления меняющихся тем и нарративов, которые определяют самоощущение и позицию движения. Таким образом, исследователи использовали *LDA* для анализа данных *Twitter* движения *BLM*, продемонстрировав, как коллективная идентичность конструировалась посредством обсуждений расовой несправедливости, жестокости полиции и солидарности сообщества [Giorgi et al., 2022].

Однако опора *LDA* на представления «мешка слов» (англ. *bag-of-words*), которые обрабатывают слова независимо от его позиции в предложении, иногда может скрывать более глубокие, зависящие от контекста значения терминов, критически важные для понимания нюансов трансформации коллективной идентичности. Хотя *LDA* эффективен в определении широких тематических категорий, он с трудом улавливает развивающуюся семантику дискурса, связанного с идентичностью. Это ограничение заставило исследователей искать более контекстуально-чувствительные модели, которые лучше отражают сложность современных протестных движений и их динамические идентичности.

В последние годы на смену традиционным моделям пришли более современные подходы, такие как *BERT* (*Bidirectional Encoder Representations from Transformers*). *BERT* отличается тем, что способен анализировать текст справа налево и слева направо, изучая взаимосвязь между всеми словами в предложении [Grootendorst, 2022], что делает его особенно эффективным для моделирования

* Здесь и далее: заблокирован на территории Российской Федерации.

дискурса. Применение *BERT* в *BERTopic* позволяет не только выявить темы, но и отразить их динамическое развитие в зависимости от контекста, что особенно важно при исследовании протестных движений, таких как *BLM*, где семантика и тональность могут значительно изменяться в ответ на внешние вызовы.

Таким образом, настоящее исследование заполняет две важные лакуны в литературе. Во-первых, оно предлагает методологическое улучшение, внедряя *BERT* для анализа динамической трансформации тем, что является шагом вперед по сравнению с доминирующим использованием *LDA*, который не улавливает контекстуальных нюансов. Во-вторых, оно решает содержательную задачу: исследует, как изменяются коллективные идентичности участников протестного движения *BLM* под влиянием контрпротестов, что является пробелом в исследовательской литературе.

Данные

В качестве данных в исследовании используются онлайн-сообщения (посты) участников протестных движений *BLM* в социальной сети *X (Twitter)*. Обоснованием такого выбора служит то, что, во-первых, исследование фокусируется на изучении трансформации коллективных идентичностей самих участников протеста в ответ на контрпротест. Во-вторых, *BLM* и *ALM* не являются организациями со стабильными лидерами и протестными манифестами – организация протестного движения чаще всего происходит на локальном уровне, а очередная волна протестной активности и вовсе начинается спонтанно с поста одного из участников такого протеста.

Временной рамкой исследования является период с 12 июля 2013 г. по 31 декабря 2014 г. Начальная граница обусловлена первым использованием хештега *#BlackLivesMatter* в *Twitter*. В 2014 г. в ответ на протестное движение возникает контрпротестное под хештегом *#AllLivesMatter*. Поэтому для отслеживания трансформации коллективных идентичностей участников протестного движения имеет смысл отследить и сравнить тематики сообщений участников *BLM* до и после столкновения с *ALM*.

Данные были выгружены с помощью клиент-платформы веб-скраппинга *Apify*, так как традиционный доступ к выгрузке сообщений через *Twitter API* был ограничен для индивидуального пользования. Ввиду неравномерности количества сообщений за

выбранный период до и после 2014 г., было решено случайным образом выбрать 50 тыс. постов так, чтобы каждый пост имел одинаковую вероятность быть выбранным. Посты отбирались не только по хештегам *#BlackLivesMatter* и *#AllLivesMatter* (только для 2014 г.), но и по топу ассоциированных хештегов с главными хештегами. В выбранный временной период такими хештегами стали *#icantbreathe* *#ferguson* *#ericgarner* *#blacktwitter* *#mikebrown* *#tamirrice* *#shutitdown* (*Pew Research Center*). Как можно заметить, многие из них содержат имена чернокожих жертв полицейской жестокости в период 2013–2014 гг.

Описание метода

В этой статье применяется *BERTopic*, подход к нейронному тематическому моделированию, который использует метод встраивания (англ. *embeddings*) *BERT* для выявления скрытых тем в текстовых данных большого объема. В отличие от обычных моделей, которые полагаются на «мешки слов», *BERTopic* использует предварительно обученные модели на основе трансформатора для генерации плотных векторных встраиваний онлайн-сообщений.

BERTopic начинается с преобразования входных онлайн-сообщений в векторные представления. Хотя существует множество способов, с помощью которых это можно сделать, в данном исследовании использовался преобразователь предложений («*all-MiniLM-L6-v2*»), поскольку они вполне способны улавливать семантическое сходство между документами.

Далее для оптимизации процесса кластеризации размерность встраивания документов уменьшается с помощью *UMAP* (*Uniform Manifold Approximation and Projection*), что сохраняет локальную и глобальную структуру данных. Затем сокращенные вложения кластеризуются с помощью *HDBSCAN* (*Hierarchical Density-Based Spatial Clustering of Applications with Noise*). *HDBSCAN* особенно подходит для этой задачи, поскольку он может идентифицировать кластеры различной плотности, что позволяет обнаруживать сложные закономерности в данных. Однако одним из ограничений является производительность алгоритма кластеризации *HDBSCAN*, который часто классифицирует большинство документов как выбросы, исключая их из дальнейшего анализа. В контексте текущей задачи это не является большой проблемой, так как механизм работы алгоритма *HDBSCAN* позволяет акцентировать внимание на

особенно значительных и плотных кластерах, тем самым более четко задавая темы. После формирования кластеров применяется процедура *TF-IDF* (*Term Frequency-Inverse Document Frequency*) на основе специальных классов для извлечения наиболее репрезентативных слов для каждой темы. Эта процедура рассматривает все документы в кластере как единое целое и вычисляет важность каждого слова для темы на основе его частоты в кластере и его относительной частоты во всем корпусе. Это позволяет генерировать согласованные представления тем, которые основаны как на локальном, так и на глобальном распределении слов.

Важно отметить, что для данной модели совсем не нужно заниматься подготовкой текста (например, удаление стоп-слов, пунктуации и т.д.). *BERTTopic*, использующий встраивания на основе BERT и методы кластеризации, предлагает более гибкий и тонкий подход к моделированию тем по сравнению с традиционными моделями, такими как *LDA*, учитывая контекст и значение слов, которые играют решающую роль для понимания основных тем.

Результаты исследования

В результате было построено две модели тематического моделирования *BERT*: для 2013 и 2014 гг. соответственно. Таблица 1 содержит результаты выделения самых популярных семи тем для первого временного периода, а рисунок 1 иллюстрирует вклад топ-5 слов в каждую из топ-7 тем.

Таблица 1

Результаты выделения тем и примеры, репрезентирующие онлайн-сообщения участников протестного движения *BLM* в 2013 г.

Тема	N	<i>BERT</i> -Название	Репрезентация слов	Пример репрезентирующего онлайн-сообщения
1	2	3	4	5
1	1015	0_call_police_dey_policebrutality	call, police, dey, policebrutality, g_bilz, officers, report, brutality, sandford_police, calling	SO angry abt this #policebrutality. Plz do something. I have a family & am scared of our police...
2	918	1_racism_racist_casual_reverse	racism, racist, casual, reverse, manderson12369, jackmacfb, beartrapstrong, exists, rothmus	I guess racism is not as clear cut to the oppressor. Hard to see from where you are standing.

Продолжение табл. 1

1	2	3	4	5
3	678	2_brutality_cops_police_video	brutality, cops, police, video, videos, backfires, fight, astonishing, caught, handled	THAT'S POLICE BRUTALITY! http://t.co/AyKlVnWnqK
4	556	3_blacks_blues_wear_pack	blacks, blues, wear, pack, double, greys, ones, reds, buy, greens	Blacks!!!
5	555	4_icantbreathe_itsjustcelle_lmfaowizzzyjr	icantbreathe, itsjustcelle, lmfaowizzzyjr, dying, omg, tink_tink25, dead, aairsimon-singing, fsmith_87	@FightMate Black Lives Matter 👊 #icantbreathe
6	398	5_suspect_robbery_burglary_arrested	suspect, robbery, burglary, arrested, shooting, seek, charged, armed, arrest, searching	«[Video] Two black suspects arrested after robbery spree across San Antonio, police say Source: KENS 5 Eyewitness News https://t.co/xx9tAA048C »
7	381	6_blacks_huffman_mynameisface_bikinibabes_sa	blacks, huffman, mynameisface, bikinibabes_sa, breh, 89bornboss, aej_blacks, sammy_blacks, fearmoho, stillgray	Blacks are everywhere....

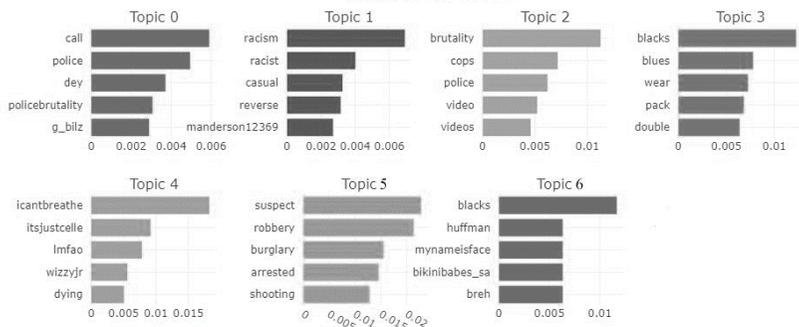


Рис. 1.

Вклад топ-5 слов в семь самых популярных тем в 2013 г.

Первая тема репрезентирует жестокость полиции и правоохранительных органов по отношению к участникам протестного движения *BLM*. Наиболее часто встречающиеся термины в этой теме – «*call*», «*police*», «*dey*», «*police brutality*» и пользовательский псевдоним «*g_bilz*». Наличие таких слов, как «*police*» и «*police brutality*», указывает на то, что эта тема сосредоточена на призывах к действию и на главном фокусе недовольства *BLM* в 2013 г.:

насилии полиции в отношении чернокожих. Повторяющееся упоминание полиции и связанных с ней тем показывает, что участники активно занимались проблемой жестокости полиции, сообщали об инцидентах, требовали внимания и отстраивали коллективные идентичности относительно данной аутгруппы. Частое упоминание призывов к действию и прямые ссылки на полицию говорят о том, что участники были сосредоточены на мобилизации вокруг конкретных инцидентов и повышении осведомленности.

Вторая тема явным образом подчеркивает проблему расовой дискриминации, с которой сталкивались участники протестного движения. Ключевые слова этой темы включают «*racism*», «*racist*», «*casual*» (повседневное дискриминационное поведение), «*reverse*» (применительно к обратному расизму) и «*manderson12369*» (пользовательский ник, который активно был включен в обсуждение этой темы). Тема охватывает обсуждения вокруг расизма как в явной, так и в скрытой форме. Наличие слова «*reverse*» предполагает, что участники участвовали в дебатах или отпорах заявлениям об «обратном расизме», в то время как такие термины, как «расистский» и «случайный», подчеркивают системный и повседневный характер расовой дискриминации.

Третья тема посвящена видеороликам, на которых запечатлены факты жестокости полиции. Основными словами здесь являются «*brutality*», «*cops*», «полиция», «видео». Акцент на «видео» предполагает, что участники в 2013 г. активно обсуждали визуальные доказательства жестокости полиции, что способствовало широкому обсуждению инцидентов, связанных с насилием со стороны полиции. Эта тема подчеркивает роль социальных сетей и вирусных видеороликов в документировании и распространении информации о жестокости полиции. Повторное использование «видео» предполагает, что участники использовали визуальный контент как мощный инструмент для разоблачения насилия полиции и мобилизации поддержки движения.

Следующая тема подчеркивает коллективный символизм движения *BLM*. Ключевые термины в этой теме – «*blacks*», «*blues*», «*wear*», «*pack*», «*double*». Эта тема предполагает символическое обрамление движения, где маркеры идентичности, такие как цвет кожи и цвет одежды (синий для полиции), становятся важными в построении идентичности движения и подчеркивает враждебные отношения между участниками протестного движения *BLM* и правоохранительными органами.

Тема 5 посвящена усилению социальных сетей и хештегов, ассоциированных с движением *BLM*. В этой теме представлены такие термины, как «*icantbreathe*», «*itsjustcelle*», «*lmfao*», «*wizzyjr*» и «*dying*». Фраза «*I can't breath*» является культовым лозунгом (и позже, хештегом), связанным со смертью Эрика Гарнера (чернокожий мужчина, который был убит при задержании полицейским путем удушения), и стала ключевым призывом для движения *BLM*. «*Itsjustcelle*» и «*wizzyjr*» являются пользователями, которые усилили распространение подобных сообщений. Эта тема отражает активизм хештегов и роль отдельных аккаунтов в социальных сетях в усилении сообщения *BLM*.

Тема 6 иллюстрирует преступность и правоохранные органы через такие ключевые слова, как «*suspect*», «*robbery*», «*burglary*», «*arrested*», «*shooting*». Эта тема отражает разговоры о криминализации чернокожих и непропорциональном акценте на преступности. Участники протестного движения заявляли, что чернокожих часто выставляют подозреваемыми или преступниками, даже когда они являются жертвами насилия.

И, наконец, последняя тема в топе самых популярных касается объективации чернокожего населения. Ключевые слова здесь – «*blacks*», «*huffman*», «*mynameisface*», «*bikinibabes_sa*» и «*breh*». Упоминание таких терминов, как «*bikinibabes*» и «*mynameisface*», предполагает разговор об объективации черных тел, возможно, связанный с тем, как черные люди, особенно женщины, изображаются в СМИ или на социальных платформах. Эти обсуждения могут касаться того, как конструируются и эксплуатируются идентичности чернокожего населения, способствуя более широкому разговору о системном расизме и репрезентации.

Таким образом, темы 2013 г. иллюстрируют, как коллективная идентичность *BLM* формировалась посредством множественных взаимосвязанных дискурсов: мобилизация против жестокости полиции, продолжающаяся борьба с расизмом во всех его формах, символическое напряжение между чернокожими общинами и правоохранными органами и использование социальных сетей в качестве важнейшего инструмента для выражения идентичности и усиления протеста.

Аналогичная модель была построена для массива 2014 г. Таблица 2 отображает результаты выделения топа-7 популярных тем, топ ключевых слов по каждой теме, а также пример одного из постов, который репрезентирует тему.

Таблица 2

**Результаты выделения тем и примеры,
репрезентирующие онлайн-сообщения участников
протестного движения *VLM* после взаимодействия
с контрпротестным движением *ALM* в 2014 г.**

Тема	N	BERT-Название	Репрезентация слов	Пример репрезентирующего онлайн-сообщения
1	3089	0_racism_prejudice_racist_alive	racism, prejudice, racist, alive, definition, exist, still, race, racists, it	Racism still alive, man racism is still alive... People are so fucking stupid the only reason prejudice and racism exists is cause people look for reasons to hate the people around them
2	817	1_responsible_torture_party	responsible, torture, party, no, for, and, is, brutality, of	@WhiteHouse The "Party of No" is responsible for police brutality and torture. http://t.co/7IP6a6YI3g
3	700	2_cops_icantbreathe_law_ticket	cops, icantbreathe, law, ticket, unarmed, killing, policestate, heroic, iambig7, kill	#icantbreathe YOU CAN HELP with this fatal police encounter database @fatalencounters http://t.co/O6IILCvMnm ,
4	689	3_mate_alllivesmatter_johnfugelsang_thank	mate, alllivesmatter, johnfugelsang, thank, joesantagato, true, agree, billybatts6670, carlquintanilla, dumb	@AnahitaSeda mate #AllLivesMatter, @Saffronian mate #AllLivesMatter, @KatiePavlich mate #AllLivesMatter
5	636	4_brutality_police_cases_victims	brutality, police, cases, victims, also, sallykohn, problem, oldbaldgooner, issue, concern	@OldBaldGooner The idea that I can just obey the law to prevent police brutality is not only insulting but an out right
6	613	5_matter_lives_all	matter, lives, all, peter-sagal, deenanicolemtv, agree, rapzie, youranonews, unclerush, scandalmonium	Don't say that all lives matter! Blacks are faced with everyday challenge, you aren't!
7	600	6_supportng_bencarson_insults_roots	supportng, bencarson, insults, roots, 5fxftieolu, x7tqgczdyj, towards, african, israel, fxnwkoem5r	BenCarson insults African roots by supportng Israel racism towards Blacks http://t.co/5fxFTieolu @cher23145 : BenCarson insults African roots by supportng Israel racism towards Blacks http://t.co/5fxFTieolu

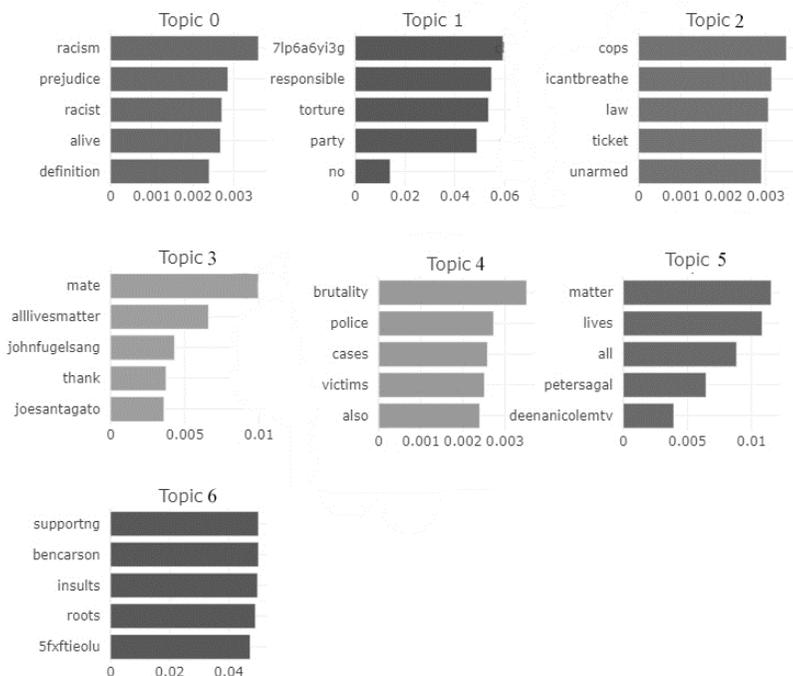


Рис. 2.

Вклад топ-5 слов в семь самых популярных тем в 2014 г.

Первая тема ярко иллюстрирует проблему расизма и ассоциированные с ними предрассудки: используются такие ключевые слова как «*racism*», «*prejudice*», «*racist*», «*alive*», «*definition*». Эта тема сосредоточена на обсуждениях системного расизма и на продолжающейся борьбе с расовыми предрассудками. Термин «*alive*» предполагает, что расизм сохраняется в современном обществе, несмотря на заявления о том, что был достигнут прогресс. Также данная тема отражает центральный компонент идентичности *BLM* в ее борьбе с расизмом. Онлайн-сообщения сосредоточены на признании продолжающегося присутствия расовых предрассудков и оспаривании заблуждений об искоренении расизма.

Вторая тема сосредоточена на вопросе ответственности. Ключевыми словами в определении темы стали «*responsible*»,

«torture», «party», «no». Данная тема связана с вопросами подотчетности, направленными против политических субъектов (Белый дом, в частности) или правоохранительных органов за их роль в системном насилии. Тема фокусируется на привлечении институтов к ответственности за их роль в расовой несправедливости, в частности, правоохранительных органов и политических организаций. Онлайн-сообщения иллюстрируют, что участники выступали за большую ответственность и реформы для решения проблемы продолжающегося насилия и дискриминации.

Как и в случае 2013 г., тема полиции и правового насилия и угнетения осталась актуальной и в 2014 г. Она подчеркивает роль правоохранительных органов в сохранении насилия в отношении чернокожих, особенно в случаях, когда жертвами являются невооруженные люди. Постоянное использование фразы «*I Can't Breathe*» или одноименного хештега указывает на сосредоточенность движения на системной несправедливости и необходимости реформ.

Четвертая тема ассоциируется с противопоставлением себя как участника протеста *BLM* с участниками контрпротеста *All Lives Matter*. Эта тема отражает контрнарратив *All Lives Matter (ALM)* и напряженность между *ALM* и *BLM*. Присутствие *ALM* в обсуждении показывает, что участники *BLM* взаимодействовали с этой оппозицией и реагировали на нее, а *ALM* все еще был на стадии формирования, как отдельный самостоятельный протест-реакция. В частности, участники *BLM* пытались оспорить «разбавление» своего сообщения более универсалистской риторикой *ALM*, что подчеркивает идеологическую борьбу между двумя движениями.

Следующая тема, как и в 2013 г., сосредоточена на жестокости полиции. Однако обсуждение жертв встречается в среднем чаще, чем ранее. Эта тема вращается вокруг конкретных случаев жестокости полиции и жертв, пострадавших от рук правоохранительных органов. Она отражает дискуссии о привлечении полиции к ответственности за эти действия, необходимости системной реформы, и подчеркивает акцент *BLM* на поиске справедливости для жертв полицейского насилия и выделении громких случаев жестокости. Участники протеста призывают сосредоточиться на рассмотрении вопроса ответственности и системных изменений.

Тема 6 подчеркивает, что именно жизни чернокожих важны; в онлайн-сообщениях подчеркивается, что движение нацелено на борьбу за признание в обществе, которое исторически обесценивало их жизни. Включение влиятельных лиц и имен пользователей

указывает на роль деятелей социальных сетей в усилении этого послания.

Тема 7 фокусируется на обсуждениях поддержки чернокожих сообществ и признании культурных корней системного расизма. Упоминание Бена Карсона, видного политического деятеля, предполагает, что некоторые сообщения включали критику его позиции по расовым вопросам (стоит упомянуть, что Бен Карсон сам является афроамериканцем, который призвал активистов *BLM* направлять свой гнев на другие цели, а не на полицию и участников контрпротеста).

Темы, извлеченные из модели *BERT* для дискурса *BLM*, раскрывают несколько ключевых аспектов коллективных идентичностей участников протестного движения. Центральными в этих дискуссиях являются темы расизма и системных предрассудков, где участники протеста активно оспаривали представление о том, что расизма больше не существует. Также особое внимание уделяется жестокости полиции (темы 2 и 5), при этом устойчивое влияние движения «Я не могу дышать» служит мощным символом системного насилия, с которым сталкиваются сообщества чернокожих.

Модель также охватывает взаимодействие между *BLM* и контрдвижениями, такими как *All Lives Matter* (темы 3 и 4), где участники взаимодействовали и реагировали на контрнарративы, которые пытались подорвать или переориентировать *BLM* с их точки зрения. Эти темы иллюстрируют идеологическую борьбу между двумя движениями, при этом участники *BLM* отстаивали необходимость особого внимания к жизням чернокожих.

Кроме того, обсуждения образа полицейского и правосудия для правоохранительных органов (тема 4) показывают, как участники справлялись с ответной реакцией со стороны движений за правоохранительные органы, отражая сложности поддержания единого сообщения перед лицом оппозиции.

Наконец, сообщения о поддержке чернокожих сообществ и понимании корней расизма (тема 7) подчеркивают солидарность внутри движения и признание более глубоких культурных и исторических факторов, способствующих системной расовой несправедливости. Включение влиятельных лиц в социальных сетях подчеркивает роль цифровых платформ в распространении и формировании сообщения *BLM*. В целом эти темы раскрывают многогранную коллективную идентичность, сформированную как внутренними целями, так и внешними вызовами.

Заключение

Исследование трансформации коллективной идентичности в контексте протестных и контрпротестных движений показало динамичную и многослойную природу коллективных идентичностей участников протестного движения *Black Lives Matter*. Применение *BERT*-тематического моделирования позволило выявить эволюцию дискурса движения в ответ на появление контрпротеста *All Lives Matter*.

Выявленные темы напрямую связаны с формированием и трансформацией коллективной идентичности, так как они служат ключевыми инструментами артикуляции целей, ценностей и границ группы. Темы создают общее пространство смыслов, которое позволяет участникам осознавать свою принадлежность и выстраивать солидарность. Через темы группа формирует свое понимание «мы» в противопоставлении «другим», определяет приоритеты борьбы и адаптируется к внешним вызовам. Эти темы становятся не только выражением существующих проблем, но и активным процессом переосмысления идентичности, который позволяет группе сохранять единство и актуальность в изменяющемся контексте.

Исследование трансформации коллективной идентичности участников движения *Black Lives Matter* (BLM) продемонстрировало две ключевые стадии ее развития под влиянием протестов и контрпротестов. На первом этапе идентичность BLM была четко направлена на борьбу с полицейским насилием и системным расизмом. Движение формировало себя как оппозицию властным структурам, которые воспринимались основным источником несправедливости. Коллективная идентичность основывалась на объединении участников вокруг борьбы с этими системными проблемами.

С появлением контрдвижения *ALM* начался второй этап, который характеризовался усложнением и расширением рамок коллективной идентичности. Контрнарратив *ALM* поставил под сомнение уникальность борьбы *BLM* за расовую справедливость, что вынудило движение переосмыслить свою идентичность. Акценты в дискурсе *BLM* начали смещаться: с появлением *ALM* «другим» стало не только государство, но и участники контрпротеста, которые оспаривают их цели и ценности. В ответ BLM расширило свое видение, включив в рамки идентичности элементы культурного сопротивления и критики идеологических основ контрпротеста.

Применение BERT-тематического моделирования позволило глубже понять динамику этих изменений. Современная модель анализа языковых данных выявила контекстуальные зависимости, показывая, как слова и фразы в дискурсе BLM менялись под влиянием внешних вызовов. Эти результаты подтверждают, что коллективная идентичность протестных движений является не статичной, а адаптивной структурой, отражающей их способность реагировать на внутренние и внешние вызовы.

Данное исследование вносит вклад в существующую литературу, заполняя пробелы как в методологической, так и в содержательной части. Во-первых, оно использует *BERT*-тематическое моделирование для анализа коллективных идентичностей, что является шагом вперед по сравнению с традиционными моделями, такими как *LDA*. *BERT* позволяет учитывать контекстуальные зависимости между словами, что особенно важно для анализа трансформации коллективной идентичности, так как слова и фразы, используемые участниками, меняются в зависимости от внешних вызовов. Это исследование демонстрирует преимущества использования более современных моделей для анализа больших объемов данных социальных сетей.

Во-вторых, на содержательном уровне исследование проливает свет на то, как контрпротестные движения, такие как *ALM*, могут влиять на трансформацию коллективной идентичности протестующих, заставляя их пересматривать свои рамки и дискурс. В литературе, посвященной коллективной идентичности, редко учитывается влияние контрдвижений на идентичность самих протестующих, что делает это исследование особенно актуальным для понимания динамики взаимодействия между протестом и контрпротестом.

Несмотря на успешное применение *BERT*, данное исследование имеет свои ограничения. Стоит отметить, что работа с *BERT* требует значительных вычислительных ресурсов, а результаты могут зависеть от настроек модели и удаления выбросов.

Будущие исследования могут сосредоточиться на более длительном временном отрезке (не одно столкновение, а несколько) для отслеживания эволюции коллективных идентичностей. Также полезным будет включение эмоционального анализа наряду с тематическим моделированием, чтобы исследовать, как изменения в тональности и эмоциональной окраске соответствуют тематическим сдвигам.

В целом же *BERT*-тематическое моделирование показало себя как мощный инструмент для изучения трансформации коллективных идентичностей, предлагая перспективную методологическую основу для дальнейших исследований на пересечении проблематик протеста, контрпротеста и коллективных идентичностей в цифровую эпоху.

V. Koncha*

BERT in focus: topic modeling the transformation of collective identities of the participants of the Black Lives Matter movement in response to the counter-protest¹

Abstract. Counter-protests have become a common response to protest movements around the world. Understanding this process is often based on collective identities formed in opposition to an out-group. In the context of the Black Lives Matter (BLM) movement, the initial image of the “other” was mainly attributed to the authorities. However, the emergence of the All Lives Matter (ALM) counter-protest movement has changed the way how collective identities are constructed – now the image of the “other” is also attributed to participants in the counter-protest movement, introducing a new layer of complexity into the dynamics of identities and raising questions about how these identities are transformed when faced with counter-protest. Applying a new approach in BERT topic modeling, the author analyzes thematic shifts in the posts of BLM protesters between 2013 and 2014 on the social network Twitter (X). The application of the BERT topic modeling model allowed to conduct a detailed analysis of online messages, capturing contextual dependencies and revealing the complexities of the transformation of collective identity in response to the emergence of the ALM counter-protest movement. The study contributes to the existing literature by filling two gaps: demonstrating the methodological advantage of using BERT on empirics and exploring how the collective identities of protest movement participants are transformed in the context of protest and counter-protest dynamics. BERT topic modeling is a powerful tool for studying the transformation of collective identities, offering a promising methodological basis for further research at the intersection of protest, counter-protest, and collective identities in the digital age.

Keywords: BlackLivesMatter; AllLivesMatter; collective identity; BERT model; protest; counter-protest; online messages

For citation: Koncha V. BERT in focus: topic modeling the transformation of collective identities of the participants of the Black Lives Matter movement in response to the counter-protest. *Political science (RU)*. 2025, N 1, P. 219–239. DOI: <http://www.doi.org/10.31249/poln/2025.01.10>

* Koncha Valeriya, HSE University (Moscow, Russia), e-mail: vkoncha@hse.ru

¹ The article was prepared within the framework of the HSE University Basic Research Program

References

- Benford R.D., Snow A.D. Framing processes and social movements: an overview and assessment. *Annual review of sociology*. 2000, N 26, P. 611–639.
- Bennett W.L., Segerberg A. The logic of connective action: digital media and the personalization of contentious politics. *Information, communication & society*. 2012, N 15 (5), P. 739–768. DOI: <https://doi.org/10.1080/1369118X.2012.670661>
- Giorgi S., Guntuku S. C., Himelein-Wachowiak M., Kwarteng A., Hwang S., Rahman M., Curtis B. Twitter corpus of the #BlackLivesMatter movement and counter protests: 2013 to 2021. *Proceedings of the 16-th International AAAI conference on web and social media*. 2022, N 16, P. 1228–1235. DOI: <https://doi.org/10.1609/icwsm.v16i1.19373>
- Grootendorst M., BERTopic: neural topic modeling with a class-based TF-IDF procedure. Preprint N 2203.05794, 2022. DOI: <https://doi.org/10.48550/arXiv.2203.05794>
- Gulevich O. *Psychology of interpersonal relations*. Moscow: Yurayt, 2019, 432 p. (In Russ.)
- Holland D., Fox G., Daro V. Social movements and collective identity: a decentered, dialogic view. *Anthropological quarterly – ANTHROPOL QUART*. 2008, N 81, P 95–126. DOI: <https://doi.org/10.1353/anq.2008.0001>
- Inata K. Protest, counter-protest and organizational diversification of protest groups. *Conflict management and peace science*. 2021, N 38 (4), P. 434–456. DOI: <https://doi.org/10.1177/0738894219869916>
- Koller V. How to analyse collective identity in discourse – textual and contextual parameters. *Critical approaches to discourse analysis across disciplines*. 2012, N 5 (2), P. 19–38.
- Malinova O.Yu. The concept of “Other” in identity studies: analysis of modern discussions. *Political science (RU)*. 2015, N 4, P. 154–169. (In Russ.)
- Melucci A. *The playing self: person and meaning in the planetary society*. New York, Cambridge: Cambridge university press, 1996, 188 p.
- Peng J., Fung J., Murtaza M., Rahman A., Walia P., Obande D., Verma A. A sentiment analysis of the Black Lives Matter movement using Twitter. *STEM fellowship journal*. 2022, N 8, P. 1–11. DOI: <https://doi.org/10.17975/sfj-2022-015>
- Polletta F. *It was like a fever: storytelling in protest and politics*. Chicago, London: The University of Chicago press, 2006, 258 p. DOI: <https://doi.org/10.7208/chicago/9780226673776.001.0001>
- Shahin S., Ng Y.M.M. Connective action or collective inertia? Emotion, cognition, and the limits of digitally networked resistance. *Social movement studies*. 2021, N 21 (4), P. 530–548. DOI: <https://doi.org/10.1080/14742837.2021.1928485>
- Stürmer S., Simon B. The role of collective identification in social movement participations: A panel study in the context of the German gay movement. *Personality and social psychology bulletin*. 2004, N 30 (3), P. 263–277. DOI: <https://doi.org/10.1177/0146167203256690>
- Tarrow S. *Power in movement: Social movements and contentious politics*. New York, Cambridge: Cambridge university press, 2011, 332 p.
- Wood L. Policing counter-protest. *Sociology compass*. 2020, N 14 (11), P. 1–10. DOI: <https://doi.org/10.1111/soc4.12833>

Zald M.N., Useem B. Movement and countermovement interaction: mobilization, tactics, and state involvement. In: Zald M.N., McCarthy J.D. (eds). *Social movements in an organizational society*. New Brunswick: Transaction, 1987, P. 247–272.

Литература на русском языке

Гулевич О. Психология межличностных отношений / под ред. Д.И. Фельдштейн. – М.: Юрайт, 2019. – 432 с.

Малинова О.Ю. Концепт «Другого» в исследованиях идентичности: анализ современных дискуссий // Политическая наука. – 2015. – № 4. – С. 154–169.