

Репрезентация России в официальном китайском медиадискурсе: корпусно-дискурсивный анализ публикаций «Жэньминь жибао» (2023–2025 гг.)

Вакулина А. Д., Иванченко Н. В.

The Representation of Russia in Official Chinese Media Discourse: A Corpus-Assisted Discourse Analysis of People's Daily Publications (2023–2025)

Vakulina A. D., Ivanchenko N. V.

ДЛЯ ЦИТИРОВАНИЯ: Вакулина А. Д., Иванченко Н. В. Репрезентация России в официальном китайском медиадискурсе: корпусно-дискурсивный анализ публикаций «Жэньминь жибао» (2023–2025 гг.) // Digital Orientalia (Цифровое востоковедение), 2026, Т. 6, № 2, Ст. 468209, С. 1–24, DOI [10.31696/DO.27824012-468209](https://doi.org/10.31696/DO.27824012-468209)

CITE THIS PAPER: Vakulina A. D., Ivanchenko N. V. The Representation of Russia in Official Chinese Media Discourse: A Corpus-Assisted Discourse Analysis of People's Daily Publications (2023–2025) // Digital Orientalia, 2026, Vol. 6, No. 2, Art. id. 468209, P. 1–24, DOI [10.31696/DO.27824012-468209](https://doi.org/10.31696/DO.27824012-468209)

Вакулина Алёна Дмитриевна — Национальный исследовательский университет «Высшая школа экономики» (101000, Россия, Москва, ул. Мясницкая, 20); Vakulina Alena Dmitrievna — National Research University Higher School of Economics (Myasnitskaya Ulitsa, 20, Moscow, 101000); ✉ advakulina_1@edu.hse.ru

Иванченко Надежда Вадимовна — кандидат филологических наук, Национальный исследовательский университет «Высшая школа экономики» (101000, Россия, Москва, ул. Мясницкая, 20), Школа востоковедения, Институт востоковедения Российской академии наук (107031, Москва, Рождественка, 12), Лаборатория цифровых исследований современного Востока ИВ РАН; Ivanchenko Nadezhda Vadimovna — Candidate of Sciences (Linguistics), National Research University Higher School of Economics (Myasnitskaya Ulitsa, 20, Moscow, 101000), School of Asian Studies, Institute of Oriental Studies of Russian Academy of Sciences (Rozhdestvenka Str., 12, Moscow, 107031), Digital Research Laboratory of Contemporary East; ✉ Nivanchenko@hse.ru

Статья получена 16 июня 2026 г., отредактирована 5 июля 2026 г., принята 7 августа 2026 г.

Аннотация: В статье рассматривается репрезентация России в официальном китайском медиадискурсе на материале публикаций газеты «Жэньминь жибао» за 2023–2025 гг. Исследование выполнено в рамках комплексного корпусно-дискурсивного подхода и направлено на выявление ключевых языковых средств, тематических моделей и дискурсивных стратегий, посредством которых формируется образ России в китайских государственных СМИ. Эмпирическую базу составил специализированный корпус из 2835 публикаций, собранных методом автоматизированного парсинга. В работе применяются методы тематического моделирования LDA, коллокационного анализа, анализа модальности и дискурсивных конструкций. Установлено, что Россия в китайском официальном медиадискурсе репрезентируется преимущественно как стратегический партнер Китая, легитимный участник международных отношений и значимый актор глобальной политики. Выявлены устойчивые тематические кластеры, связанные с дипломатическим сотрудничеством, энергетикой, идеологией, культурно-гуманитарным взаимодействием и исторической памятью. Показано, что негативные конструкции встречаются сравнительно редко и в большинстве случаев связаны не с критикой России, а с внешним воздействием на нее. Полученные результаты демонстрируют системный и институционально обусловленный характер репрезентации России в китайском медиадискурсе.

Ключевые слова: корпусная лингвистика, медиадискурс, Жэньминь жибао, Россия, Китай, дискурсивный анализ, тематическое моделирование, коллокационный анализ

Abstract: The article examines the representation of Russia in official Chinese media discourse based on publications in People's Daily from 2023 to 2025. The study is conducted within a comprehensive corpus-assisted discourse approach and aims to identify the key linguistic means, thematic patterns, and discursive strategies through which the image of Russia is constructed in Chinese state media. The empirical basis of the study consists of a specialized corpus of 2835 publications collected through automated web

scraping. The study employs LDA topic modeling, collocation analysis, modality analysis, and analysis of discursive constructions. The findings indicate that Russia is predominantly represented in official Chinese media discourse as a strategic partner of China, a legitimate participant in international relations, and a significant actor in global politics. Stable thematic clusters related to diplomatic cooperation, energy, ideology, cultural and humanitarian interaction, and historical memory are identified. The analysis shows that negative constructions occur relatively infrequently and, in most cases, are associated not with criticism of Russia but with external influences exerted upon it. The findings demonstrate the systemic and institutionally conditioned nature of Russia's representation in Chinese media discourse.

Keywords: corpus linguistics, media discourse, People's Daily, Russia, China, discourse analysis, topic modeling, collocation analysis

Введение

В современных условиях медиадискурс функционирует как ключевое пространство конструирования политически значимых смыслов, при этом в значительной мере именно через него формируется общественное восприятие государств и характера межгосударственных отношений. На фоне углубления российско-китайского стратегического партнёрства исследовательский интерес концентрируется на том, как Россия репрезентируется в официальном китайском медиадискурсе. Для корпусного анализа тексты китайских партийных СМИ представляют особую удобность, поскольку для них характерны высокая степень стилевой унифицированности, активное использование политической лексики и терминологическая стандартизированность. Благодаря этим особенностям становится возможным выявление типовых стратегий репрезентации. «До сих пор в китайских учебниках журналистской теории подчеркивается такая функция журналистики: газеты, радио, телевидение представляют собой рупор партии, правительства и народа страны» (10 Цзя, 2013: с. 53). В этом контексте газета «Жэньминь жибао», является главным печатным органом КПК и инструментом трансляции ключевой повестки.

Цель исследования заключается в выявлении ключевых языковых средств и основных способов репрезентации России в официальном китайском медиадискурсе, при этом эмпирическую базу составляют публикации газеты «Жэньминь жибао» за период 2023–2025 гг., а теоретико-методологической основой выступает корпусно-дискурсивный подход. В центре исследования находятся лингвистические механизмы, посредством которых конструируется образ страны в медиатексте. Для обеспечения большей объективности получаемых результатов был задействован комплексный корпусный анализ, в рамках которого осуществлялось сопоставление данных, полученных различными методами, между собой.

«Корпус не содержит новой информации о языке, но программное обеспечение предлагает нам новый взгляд на знакомое» (² Hunston, 2002: с. 1). Посредством аналитических инструментов корпус может быть структурирован таким образом, что становятся различимыми категории и языковые явления, которые при сплошном чтении обычно остаются незамеченными. В настоящем исследовании применяется специализированный корпус, под которым, понимается «собрание текстов определенного типа, способное репрезентативно отражать характерные черты выбранного языкового варианта» (² Hunston, 2002: с. 14).

В исследовании дискурс рассматривается как интегративный механизм, посредством которого текстовая ткань соединяется с экстралингвистическими параметрами. К их числу относятся идеология, система властных отношений, институциональный контекст и условия производства материала. Материальной основой реализации данного дискурса выступает газета «Жэньминь жибао». Вследствие институциональной подчиненности партийной верхушке каждый публикуемый материал о Российской Федерации включается в процесс формирования ее образа. В исследовании связь между медиа и дискурсом интерпретируется как соотношение инструмента и площадки реализации, где языковая норма функционирует как механизм поддержания властных отношений. Корпусно-ориентированный дискурсивный анализ активно применяется в исследованиях китайского медиадискурса. Так, Ц. Пэй и Л. Чэн на материале китайской и британской прессы показали, что корпусные методы позволяют выявлять устойчивые модели репрезентации социальных и поли-

тических объектов, скрытые за регулярными лексическими паттернами и коллокациями. Как отмечают исследователи: «Методологически корпусно-ориентированный критический дискурс-анализ не только помогает обнаружить наиболее значимые статистические языковые закономерности, которые наиболее четко раскрывают предпочтения двух корпусов [текстов]... но и помогает описывать, интерпретировать и объяснять эти языковые закономерности в более широком социально-политическом контексте» (6 Pei, Cheng, 2024). Более позднее исследование Т. Ли и Ф. Пань, выполненное на материале регулярных пресс-конференций МИД КНР, расширяет данный подход применительно к официальному дипломатическому дискурсу. Авторы показывают, что корпусный анализ позволяет сопоставлять кооперативные и полемические стратегии; по их формулировке, высказывания пресс-секретарей «формулируют позицию Китая и инициативы международного сотрудничества» (4 Li, Pan, 2025).

«Эволюция саморегулирования журналистской деятельности в Китае за исследуемый период прошла путь от жесткой этатистской модели к более смягченному варианту» (7 Ван, 2023: с. 58). Доминирующее положение в информационном пространстве сохраняется за государственными структурами и КПК. Условное деление китайских средств массовой информации включает партийные, государственные, коммерческие и региональные издания. В контексте данного исследования приоритетное значение отводится партийной прессе. Регулирование медиасферы выстраивается на основе сочетания формально провозглашенной свободы выражения мнений и разветвленной системы ограничительных механизмов, вследствие чего в редакционной практике закрепляются формы самоконтроля. Подобная роль официальных СМИ соответствует классической концепции формирования повестки дня, согласно которой пресса «может быть не слишком успешной в том, чтобы говорить людям, что думать, однако она поразительно успешна в том, чтобы говорить своим читателям, о чём думать» (6 Pei, Cheng, 2024: с. 177). Тем самым средства массовой информации оказывают влияние не столько на содержание индивидуальных убеждений, сколько на определение круга общественно значимых тем и вопросов. В рамках настоящего исследования данная специфика рассматривается как определяющая: материалы государственных китайских

СМИ интерпретируются как элемент институционально организованного официального дискурса, а не как совокупность индивидуальных журналистских высказываний.

Формирование корпуса. Результаты корпусного анализа

Современная корпусная лингвистика предоставляет возможность многоаспектного анализа текстов, в рамках которого существенную роль играют частотные исследования, анализ устойчивых словосочетаний и сопоставление распределения языковых единиц. В качестве ключевого инструмента используется конкорданс, представляющий собой «упорядоченные списки контекстов употребления определенной леммы или фразы» (²Hunston, 2002: с. 9). Полученные контексты анализировались с точки зрения типичного окружения и моделей лексической сочетаемости. Коллокация при этом понимается как «статистически значимая тенденция слов к совместной встречаемости» (³Jurafsky, Martin, 2000: с. 596). В настоящем исследовании конкордансный и коллокационный анализ единиц 俄罗斯 «Россия» и 俄国 «т.ж.» сочетался с тематическим, модальным и дискурсивно-конструкционным анализом. Технические параметры предобработки и отдельных процедур приведены в приложении.

В качестве одного из основных аналитических инструментов применялось тематическое моделирование на основе алгоритма LDA¹. «Цель тематического моделирования – выявление скрытых семантических структур – тем, или топиков» (⁹Кирина, 2022: с. 1). Каждая тема в рамках данного подхода интерпретируется как распределение лексических единиц, тогда как каждый документ описывается как комбинация нескольких тем с различной степенью вероятности их представленности. В отличие от простого частотного представления LDA позволяет выявлять статистические связи между документами и леммами (¹Blei, Ng, Jordan, 2003). Вместе с тем модель использовалась как разведочный инструмент: наименования тем присваивались на основе наиболее весомых лексем и не рассматривались как самостоятельное доказательство дискурсивной иерархии.

¹Latent Dirichlet Allocation

Эмпирическая основа данного исследования сформирована корпусом публикаций газеты «Жэньминь жибао» за период 2023–2025 гг. Материалы извлекались автоматизированно по лексемам 俄罗斯 «Россия» и 俄国 «т.ж.». По первому запросу было получено 3330 публикаций, по второму – 468; после удаления пустых строк, дубликатов и текстов недостаточной длины итоговый корпус составил 2835 публикаций. Такая поисковая стратегия обеспечивает воспроизводимый отбор текстов с прямым названием России, однако может не включать материалы, в которых страна обозначается косвенно, например через упоминания российского руководства, Москвы или двусторонних отношений без лексем 俄罗斯 и 俄国. Поэтому сформированная выборка отражает значительный сегмент исследуемого медиадискурса, но не претендует на исчерпывающий охват всей российской тематики в китайских СМИ. Техническая процедура сбора и очистки данных приведена в приложении.

Тематическое моделирование (LDA)

Для тематического моделирования проводились сегментация китайского текста, удаление нерелевантных символов и стоп-слов, а также фильтрация по частям речи. После очистки из 2835 публикаций в аналитическую выборку вошло 2476 текстов. При формировании словаря учитывались устойчивые биграммы и триграммы, исключались редкие и чрезмерно частотные единицы; итоговый лексический массив включил 7202 единицы. Точные пороговые значения, перечень примененных библиотек и параметры представления корпуса вынесены в приложение.

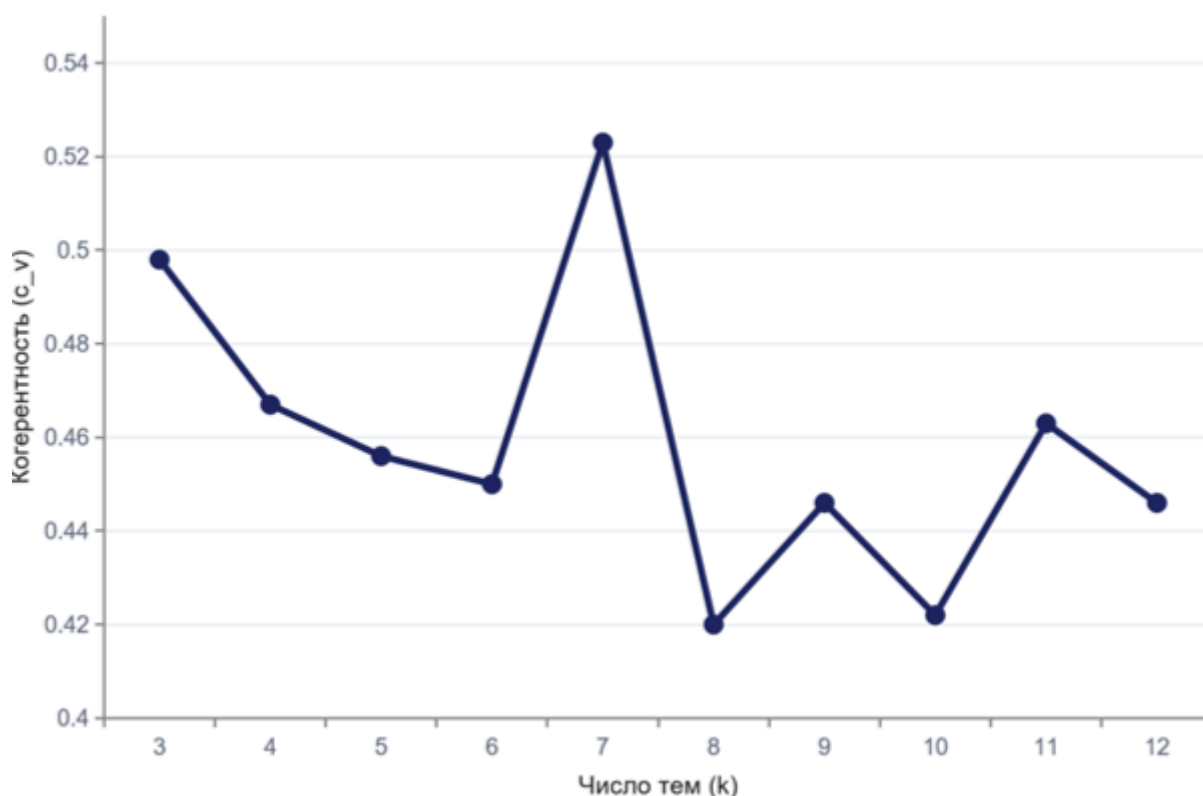


Рис. 1. Подбор оптимального числа тем для LDA-модели: ось X – число тем (от 3 до 12), ось Y – значение когерентности. Максимум достигается при $k = 7$

Определение оптимального числа тем осуществлялось с опорой на метрику когерентности: программно были последовательно сгенерированы модели с количеством тем от 3 до 12. Наиболее высокое значение когерентности (0,5230) зафиксировано при $k = 7$ (см. Рис. 1), что обусловило выбор данного значения в качестве итогового параметра модели. После превышения семи тем наблюдаются колебания показателя и его общее снижение. В рамках моделирования каждый документ был представлен как комбинация тем с различными вероятностными весами; дополнительно ему присваивалась доминирующая тематическая категория. Полученная конфигурация рассматривается как аналитически устойчивый вариант структурирования данного корпуса, но не как исчерпывающая классификация его содержания.

В состав наиболее частотных лексем Темы 0 («Военно-политический контекст и персонализация политического дискурса») входят: 普京 «Путин» (0,0062), 商品 «товар» (0,0032), 美国 «США» (0,0029), 特朗普 «Трамп» (0,0026). Несколько семантических полей, пересекаясь, формируют внутреннюю структуру данной темы. К ним относятся персонализированное

политическое позиционирование (普京 «Путин»), международное окружение Российской Федерации (美国 «США», 特朗普 «Трамп», 日本 «Япония», 中俄关系 «китайско-российские отношения»), а также военная повестка (乌克兰_危机 «украинский кризис», 战俘 «военнопленные»). Лексема 商品 «товар», несмотря на высокий статистический вес, не образует автономного экономического кластера ввиду изоляции от релевантной терминологии. В пределах данного тематического блока присутствие имени 普京 «Путин» рядом с политическими и конфликтными маркерами указывает на персонализацию образа государства через фигуру его руководителя.

Событиями на украинской территории ограничен фокус Темы 1 («Военно-политический конфликт»). Особого внимания среди ключевых единиц заслуживают: 乌克兰 «Украина» (0,0117), 美国 «США» (0,0072), 乌军 «украинская армия» (0,0057), 人员 «люди» (0,0046). Доминирование украинской проблематики верифицируется высокой частотностью лексем 乌克兰 «Украина» и 乌军 «украинская армия», которые прямо указывают на стороны противостояния. Интеграция США в медиаполе в статусе важнейшего внешнего участника выражается маркером 美国 «США», а введение лексемы 特朗普 «Трамп» нацелено на интенсификацию индивидуализации политического дискурса. Происходит расширение географических рамок за пределы регионального кризиса посредством лексемы 伊朗 «Иран». Что касается непосредственно военного и гуманитарного измерений, то они транслируются через единицы 导弹 «ракеты» и 人员 «люди», маркирующие средства ведения боевых действий и кадровые потери.

Российско-китайское стратегическое взаимодействие выступает ядром Темы 2 («Сотрудничество и историческая память»). Перечень ведущих лексем включает в себя: 金砖 «БРИКС» (0,0126), 普京 «Путин» (0,0076), 中俄关系 «китайско-российские отношения» (0,0061), 中俄 «Китай и Россия» (0,0060). Вес объединения БРИКС в пространстве китайских массмедиа детерминирует лидирующую позицию лексемы 金砖 «БРИКС». Системный характер двусторонних контактов фиксируется единицами 中俄关系 «китайско-российские отношения» и 中俄 «Китай и Россия», на что ранее указывалось в исследовании Сунь Миньцина. Проекция

современного взаимодействия на фундамент памяти о Второй мировой войне обеспечивается историческими маркерами 胜利_庆典 «празднование победы» и 国事访问_纪念_苏联_卫国战争 «государственный визит_память_СССР_Великая Отечественная война». Таким образом, тема объединяет институциональную повестку и разделяемый исторический нарратив.

Тема 3 (“Локальные события и новостная хроника”), внутри которой тематическая структура формируется преимущественно на уровне разнородных фрагментов. В составе ключевых лексем представлены: 波兰 «Польша» (0,0056), 事故 «авария» (0,0039), 白俄罗斯 «Беларусь» (0,0033), 米舒斯_白俄罗斯_戈洛夫 «Мишустин_Беларусь_Головченко» (0,0033), 牡丹 «пион» (0,0033), 毛宁 «Мао Нин» (0,0033), 火灾 «пожар» (0,0030), 事件 «событие» (0,0030). По сравнению с предыдущими тематическими блоками данная конфигурация отличается меньшей внутренней связностью, поскольку элементы объединяются скорее статистически, нежели за счет единого смыслового ядра. Отдельные происшествия фиксируются посредством лексем 事故 «авария», 火灾 «пожар» и 事件 «событие», при этом их представление в корпусе осуществляется без развернутой интерпретации. Географическая рамка задается единицами 波兰 «Польша» и 白俄罗斯 «Беларусь», которые обозначают приграничные по отношению к России государства и формируют региональный фон. Регулярные упоминания взаимодействия между российским и белорусским руководством отражены в составной лексеме 米舒斯_白俄罗斯_戈洛夫 «Мишустин_Беларусь_Головченко», через которую фиксируются институциональные контакты. Отклоняющиеся от доминирующей новостной логики элементы также обнаруживаются в рамках анализируемого тематического ряда. Включение 毛宁 «Мао Нин» в официальный китайский медиадискурс обусловлено ее статусом представителя пресс-службы МИД. Культурно-символическое поле маркируется лексемой 牡丹 «пион»: согласно существующим исследованиям, «пион в китайской культурной традиции функционирует не только как ботанический объект, но и как символ государственности и идентичности, активно используемый в дипломатических и культурных практиках» (8 Го, Ян, 2020: с. 36). Через систему подобных репрезентаций периферийные компоненты вовлекаются

в конструирование комплементарного фонового образа международных связей.

Специфика Темы 4 («Культурное и гуманитарное сотрудничество») заключается в репрезентации пространства гуманитарного и культурного взаимодействия, где фиксируется формирование устойчивого кластера, ориентированного на белорусский вектор. В перечне наиболее репрезентативных единиц фигурируют: 白俄罗斯 «Беларусь» (0,0063), 文化 «культура» (0,0059), 中白 «китайско-белорусские» (0,0045), 白方 «белорусская сторона» (0,0041). Смысловое поле темы синтезирует белорусский контекст и культурную терминологию, представленную лексемами 文化 «культура», 文学 «литература», 艺术 «искусство» и сопутствующими понятиями, что позволяет трактовать РФ и Республику Беларусь как общее гуманитарное пространство. Отсылкой к ШОС или указанием на проведение совместных двусторонних мероприятий может выступать присутствие маркера 上海 «Шанхай». Транспортно-организационное измерение добавляется лексемой 客机 «пассажирский самолет», размыкающей исходные границы темы.

Состав лексем Темы 5 позволяет условно обозначить ее как «Идеологический дискурс». Ядро исследуемого фрагмента составляют: 白俄罗斯 «Беларусь» (0,0043), 马克思主义 «марксизм» (0,0033), 普京 «Путин» (0,0032), 思想 «идея» (0,0028). Теоретико-идеологическая платформа политической коммуникации задается лексемами 马克思主义 «марксизм», 思想 «идея», 理论 «теория» и рядом других доктринальных понятий. Практическая реализация данных установок концептуализируется единицей 建设 «строительство», тогда как маркер 经济 «экономика» вводит социально-экономическое измерение. Единицы 白俄罗斯 «Беларусь», 普京 «Путин» и 庆典 «празднование» связывают доктринальную лексику с конкретным политическим и событийным контекстом.

Прагматическая сторона российско-китайского диалога, сопряженная со стратегическим партнерством и энергетической кооперацией, транслируется в Теме 6 («Стратегическое партнерство и энергетическое сотрудничество»). Ведущими лексемами здесь признаются: 能源 «энергия» (0,0065), 普京 «Путин» (0,0061), 中俄 «Китай и Россия» (0,0047), 中俄关系 «китайско-российские отношения» (0,0046). Ключевое значение

энергетики для данного тематического блока фиксируется весом лексемы 能源 «энергия». Единицы 中俄 «Китай и Россия», 中俄关系 «китайско-российские отношения», 战略 «стратегия», 伙伴关系 «партнерство» и 全面_战略_协作 «всеобъемлющее стратегическое сотрудничество» воспроизводят канон дипломатической риторики, а лексема 和平 «мир» включает взаимодействие в повестку политического равновесия. Присутствие 普京 «Путин» в списке наиболее весомых единиц учитывается как общая особенность официально-политического корпуса и не образует отдельного вывода.



Рис. 2. Круговая диаграмма распределения документов по темам

Наибольшая доля исследуемого корпуса, составляющая 26,5% от общего объема текстового массива, приходится на Тему 5, что отражено на круговой диаграмме (Рис. 2). В рамках принятой интерпретации это указывает на заметную представленность идеологической лексики наряду с практическими сюжетами. Однако распределение документов по темам LDA само по себе не позволяет утверждать, что идеологический дискурс доминирует.

нирует над прагматической повесткой: для такого вывода необходима отдельная качественная проверка репрезентативной части корпуса. Поэтому результат рассматривается как тенденция, выявленная конкретной моделью. Конфликтная повестка (Тема 1) также остается заметной, но не исчерпывает образ России. Присутствие Беларуси в Теме 4 и 5 позволяет выделить устойчивый белорусский компонент культурно-гуманитарного и идеологического контекста, не распространяя этот вывод за пределы исследованной выборки.

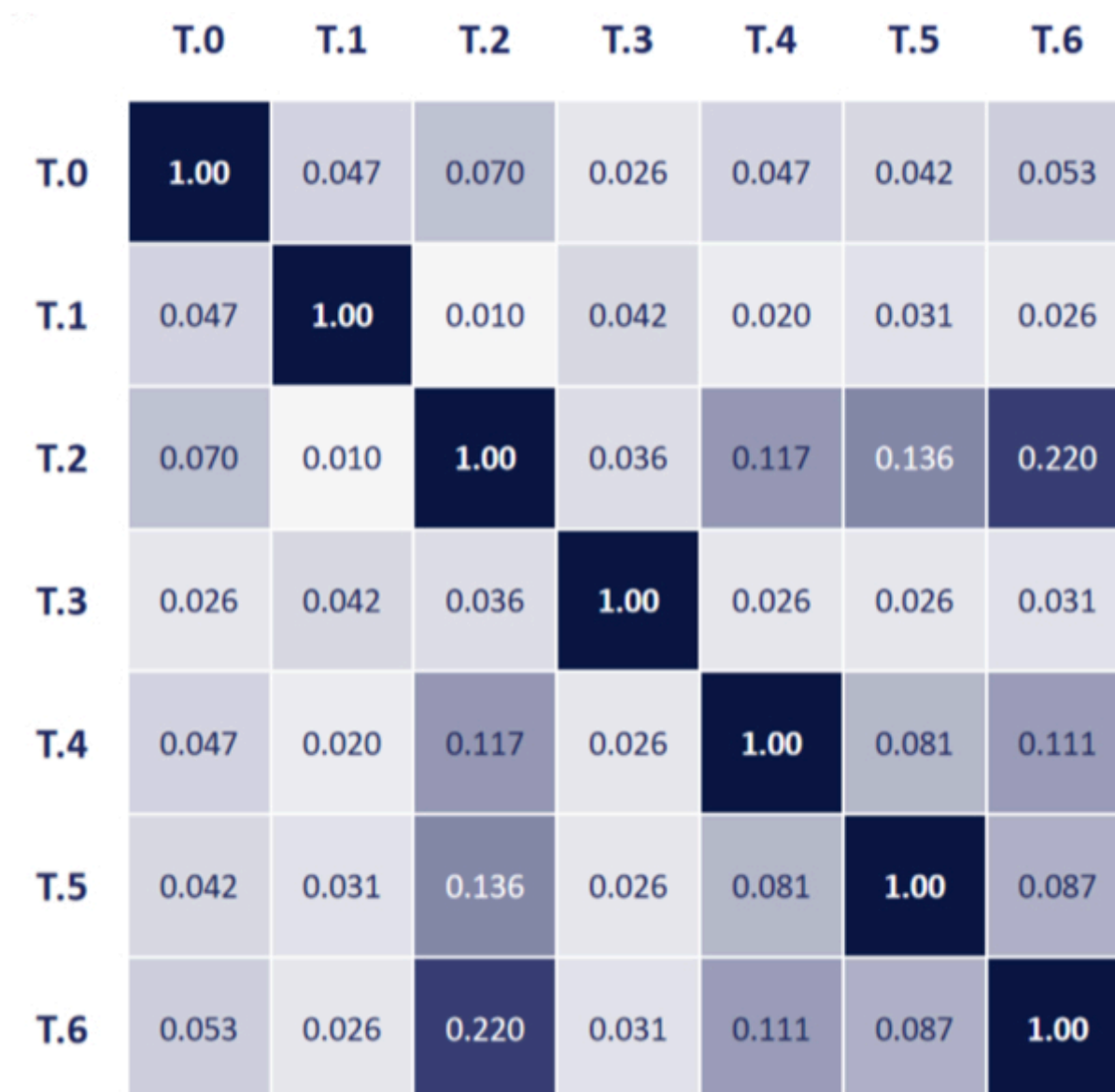


Рис. 3. Тепловая карта корреляций тем (матрица пересечений)

Степень статистической взаимосвязи между выделенными темами, рассчитанная на основе их совместной представленности в документах, иллюстрируется тепловой картой корреляций (Рис. 3). Полученные коэф-

фициенты варьируются от 0 до 0,220. Максимальный показатель зафиксирован для Темы 2 («Сотрудничество и историческая память») и Темы 6 («Стратегическое партнерство и энергетическое сотрудничество») - 0,220. Более умеренная связь обнаруживается между Темами 2 и 5 («Идеологический дискурс») - 0,136, а также Темами 2 и 4 («Культурное и гуманитарное сотрудничество») - 0,117. Тема 1 («Военно-политический конфликт») демонстрирует минимальные связи с остальными блоками, что в пределах модели может указывать на относительную обособленность конфликтной повестки.

Совместная представленность Тем 2 и 6 связывает экономическое и энергетическое сотрудничество с институциональным и историческим контекстом: упоминаниями БРИКС, двусторонних отношений и дипломатического взаимодействия. Корреляции между Темами 2 и 5, а также 4 и 6 показывают, что институциональная, идеологическая, культурная и экономическая линии не являются полностью изолированными. Промежуточное положение занимает Тема 0 («Военно-политический контекст и персонализация политического дискурса»), для которой характерны слабые, но устойчивые корреляции с большинством остальных тем. Корреляционные показатели описывают структуру модели и не рассматриваются как самостоятельное доказательство причинной связи или дискурсивной легитимации.

Коллокационный анализ

Коллокационный анализ использовался для выявления статистически значимых связей между лексемами в корпусе. Центральным объектом стала единица 俄罗斯 «Россия», для которой извлекалось локальное окружение в пределах ± 5 токенов. Устойчивость связей проверялась с учетом частотности, T-score и MI-score; дополнительно анализировались биграммы и типичные модели левого и правого окружения. Программные средства и последовательность расчетов приведены в приложении.

Высокие показатели частотности и статистической значимости демонстрируют лексемы 中国 «Китай», 莫斯科 «Москва», 总统 «президент», 普京 «Путин», 主席 «председатель», 国家 «государство». Полученные

результаты дополняют тематическое моделирование и показывают преобладание официально-политического контекста. Дипломатическая деятельность, межгосударственные отношения и обозначения политических акторов выступают главными маркерами локального окружения России. Роль имен собственных в персонализации дискурса уже рассмотрена в рамках Темы 0; здесь они учитываются как элемент типичной коллокационной модели.

Китайская медиаперспектива и новостной формат подачи материалов эксплицируются левым окружением лексемы 俄罗斯 «Россия», где наиболее регулярно фиксируются единицы 习近平 «Си Цзиньпин», 主席 «председатель», 中国 «Китай», 新华社 «агентство Синьхуа», 记者 «корреспондент», 莫斯科 «Москва». Слова 总统 «президент», 普京 «Путин», 中国 «Китай», 国家 «государство», а также глаголы 进行 «проводить», 举行 «провести», 出席 «присутствовать», отражающие официальные мероприятия, занимают преимущественно правую позицию.

Корпус характеризуется широким присутствием клишированных конструкций официального дискурса, среди которых выделяются: 俄罗斯总统 «президент России», 俄罗斯总理 «премьер-министр России», 会见俄罗斯 «встреча с Россией», 访问俄罗斯 «визит в Россию». Их регулярное воспроизведение свидетельствует о высокой степени стандартизации языка официальных СМИ. Официальный регистр поддерживается за счет повторяющихся обозначений должностей, государств, встреч, визитов и публичных мероприятий. В целом контекстное распределение демонстрирует устойчивую структуру: слева от 俄罗斯 «Россия» обычно располагается источник сообщения или политический субъект, справа действие, событие или его конкретизация. Образ России формируется преимущественно через участие в дипломатических и межгосударственных процессах, без выраженной эмоциональной оценки. Существенную роль играет китайская перспектива подачи материала, поскольку Россия чаще всего упоминается в непосредственной связи с Китаем и двусторонним взаимодействием.

Анализ модальности

Методика исследования была ориентирована на анализ модальных конструкций и определение синтаксической роли упоминаний России в окружении модальных глаголов. В корпусе фиксировалось совместное употребление наименований России с 28 модальными словами, после чего на основе набора синтаксических паттернов и порядка слов определялась субъектная или объектная позиция страны. Учитывались речевые глаголы, маркеры намерения, объектные показатели и лексика давления, ограничения и побуждения. Подробная схема классификации вынесена в приложение.

Для предложений, не соответствовавших установленным паттернам, использовалось правило порядка слов; неклассифицируемые примеры выделялись в категорию *unknown*. Поскольку метод основан на эвристических правилах, приближенных к упрощенному синтаксическому парсингу, его результаты требуют осторожной интерпретации и качественной проверки неоднозначных контекстов.

Таблица 1. Частотность модальных слов и доля субъектных употреблений

№	Модальное слово	Перевод	Частотность	Доля субъективных употреблений, %
1	会	мочь, уметь; будет (в будущем)	3034	53,3
2	要	хотеть; нужно; собираться	1509	58,1
3	能	мочь, быть способным	982	56,4
4	得	нужно, следует; необходимо	945	62,5
5	应	следует, должно	833	37,7
6	希望	надеяться, желать	190	43,2
7	计划	планировать, план ²	154	65,6
8	可能	возможно, вероятно	133	55,6
9	需	требоваться, необходимо	126	68,3
10	准备	готовиться, намереваться	95	46,3
11	可以	можно, иметь возможность	74	52,7

²слово 计划 грамматически не является модальным глаголом, но семантически связано с модальностью намерения через значение планируемости или намерения

№	Модальное слово	Перевод	Частотность	Доля субъективных употреблений, %
12	须	необходимо, следует	67	53,7
13	能够	быть в состоянии, иметь возможность	59	69,5
14	必须	должен, обязан; необходимо	56	51,8
15	需要	нуждаться; необходимо	51	58,8
16	试图	пытаться, стремиться	47	55,3
17	不能	не мочь; нельзя	25	28,0
18	应该	следует, должно быть	20	65,0
19	打算	намереваться, планировать	19	52,6
20	只能	можно только; быть вынужденным лишь	12	66,7

Статистические данные о частотности модальных слов и распределении ролей субъект/объект представлены выше (Таблица 1). Согласно общей выборке, в 54,4% случаев Россия выступает как субъект модального действия, то есть рассматривается как актер, способный или обязанный совершать действие, тогда как в 45,6% контекстов ей отводится объектная роль воздействия со стороны других участников коммуникации.

С лингвистической точки зрения репрезентация России определяется преимущественно синтаксической организацией высказывания, а не только набором используемых лексем. Существенное значение приобретает распределение модальных конструкций, через которые формируется образ страны в медиадискурсе. Наиболее частотным модальным словом выступает 会 «мочь, уметь; будет (в будущем)» (3034 употребления), что свидетельствует о преобладании перспективы будущего и ориентации китайских СМИ на прогнозируемые события, ожидаемые процессы, дипломатические сценарии и предполагаемые изменения. Вторую позицию по частотности занимает 要 «хотеть; нужно; собираться» (1509 употреблений). Особенно заметны различия при сопоставлении показателей субъектности. Модальные глаголы, связанные с возможностью, намерением и потенциальным действием, такие как 得 «нужно, следует; необходимо» (62,5% субъектных употреблений), 计划 «планировать, план» (65,6%) и 能够 «быть в состоянии, иметь возможность» (69,5%), преимущественно помещают Россию в позицию активного актора. В подобных конструкци-

ях страна репрезентируется как самостоятельный субъект, обладающий способностью к действию, стратегическому планированию и политической инициативе. Иная картина наблюдается в контекстах модальности долженствования. Для глагола 应该 «следует, должно» доля субъектных употреблений составляет лишь 37,7%, вследствие чего в большинстве случаев Россия оказывается включенной в конструкции внешнего требования, ожидания или обязательства. Сходная тенденция фиксируется и для слова 不能 «не мочь; нельзя»: субъектная позиция представлена только в 28% примеров, что указывает на репрезентацию России как актора, ограниченного внешними обстоятельствами, политическими рамками или международными условиями. Таким образом, различные типы модальности напрямую коррелируют с распределением субъектной активности в медиадискурсе. Конструкции, выражающие возможность, намерение и потенциальность, тяготеют к субъектной позиции России, тогда как модальность запрета, ограничения и долженствования чаще связывается с объектной ролью.

С дискурсивной точки зрения выявленная модель обладает принципиальной значимостью для интерпретации образа России в медиатекстах. Преобладание субъектной модальности, особенно характерное для конструкций с 得 «нужно, следует; необходимо», 计划 «планировать, план» и 能够 «быть в состоянии, иметь возможность», способствует формированию представления о России как о суверенном акторе, самостоятельно определяющем политические решения, реализующем намерения и обладающем способностью к действию. Одновременно заметная доля объектных конструкций, составляющая 45,6% и особенно выраженная в контекстах модальности долженствования (应该 «следует, должно») и запрета (不能 «не мочь; нельзя»), отражает включенность России в систему внешних ожиданий, нормативных ограничений, международных требований и регулирующих механизмов.

В результате в официальном китайском медиадискурсе складывается функционально сбалансированный образ России. Страна представляется активным участником международных процессов, чья субъектность сохраняется в рамках существующей системы внешних регуляций и политических ограничений. Подобная репрезентация носит гибридный

характер: модальность возможности и намерения закрепляет за Россией роль субъекта действия, тогда как конструкции долженствования и запрета соотносят её с позицией объекта внешнего воздействия.

Дискурсивный анализ китайского текста

Методика была ориентирована на выявление устойчивых дискурсивных паттернов и объединяла контекстный, лексический, конструкционный и статистический уровни анализа. Сначала извлекались контексты ключевой леммы в окне по пять слов с обеих сторон, затем рассчитывалась частотность биграмм и триграмм с упоминанием России. На следующем этапе применялся сформированный по результатам предыдущих процедур словарь позитивных, негативных и нейтральных рамок. Такой подход позволил сопоставить регулярные конструкции, но ограничил анализ заранее формализованными шаблонами. Полный перечень моделей и техническая процедура приведены в приложении.

В корпусе объёмом около 1,3 млн токенов было выявлено 8434 дискурсивно значимых конструкции, что свидетельствует о высокой плотности оценочно маркированных паттернов. В среднем на одну статью приходится около трёх устойчивых рамок, связанных с упоминанием России. Полученные данные подтверждают тезис С. Ханстон о способности корпусного анализа выявлять регулярности, которые остаются мало заметными при последовательном чтении текстов.

Наиболее показательным результатом оказалось распределение оценочных рамок. Доля позитивных конструкций составила 56,6%, нейтральных – 35,8%, тогда как негативные рамки охватывают 7,7% случаев. В пределах заданного словаря конструкций кооперативные и нейтральные модели количественно преобладают над критическими; этот вывод относится к формализованным паттернам, а не ко всему многообразию оценочных контекстов. Из 646 негативных случаев доминирующей моделью выступает конструкция «санкции против России» (518 употреблений), в которой Россия представлена объектом воздействия третьих участников. Сочетание 谴责俄罗斯 «осудить Россию» зафиксировано 38 раз. Таким

образом, негативная рамка чаще связана с внешним воздействием на Россию, чем с ее прямым осуждением в исследованном корпусе.

Выявленные на лексическом уровне тенденции находят свое подтверждение в результатах коллокационного и n-граммного анализа. Высокая степень сцепленности двух элементов дискурса и репрезентация Российской Федерации преимущественно в связке с КНР фиксируются через доминирование биграммы 中俄 «Китай и Россия» (2268 употреблений). Эффект взаимной ориентированности на кооперацию и формального равенства сторон создается за счет практически симметричного распределения конструкций 中方愿同俄方 «китайская сторона готова сотрудничать с российской стороной» и 俄方愿同中方 «российская сторона готова сотрудничать с китайской стороной» (245 и 233 случая соответственно). При описании двустороннего взаимодействия регулярно используются лексемы 合作 «сотрудничество», 战略协作 «стратегическое сотрудничество», 伙伴关系 «партнерство» и сопутствующие термины, выступающие устойчивыми маркерами официального политического дискурса КНР.

Функцию легитимации через предоставление российской стороне права голоса выполняет преобладание нейтральной рамки «Россия заявляет / сообщает», представленной 1877 случаями. Выступать в качестве самостоятельного источника информации без выраженной редакторской оценки России позволяют конструкции 俄罗斯表示 «Россия заявила», 俄方称 «российская сторона сообщила», 据俄罗斯 «по данным России». Это свидетельствует о восприятии государства как авторитетного и легитимного участника международной коммуникации.

Сочетание контекстного, n-граммного и конструкционного анализа позволило перейти от отдельных словоупотреблений к сопоставлению более общих дискурсивных тенденций. Вместе с тем анализ охватывал только формализуемые конструкции и строился на заранее заданных шаблонах регулярных выражений, поэтому его результаты не заменяют качественного чтения контекстов. Согласованность с тематическим и коллокационным анализом повышает устойчивость интерпретации, но не является независимой верификацией, поскольку все процедуры применялись к одной и той же выборке.

Заключение

Рассмотреть образ России в газете «Жэньминь жибао» на тематическом, лексико-синтагматическом, модальном и дискурсивно-конструкционном уровнях позволило применение комплексного корпусно-дискурсивного подхода. Результаты отдельных процедур сопоставлялись между собой и с теоретическими положениями первой части работы. Повторяемость некоторых признаков на разных уровнях повышает устойчивость интерпретации, однако не делает ее абсолютно объективной: методы основаны на общем корпусе и зависят от параметров отбора, разметки и модели.

Содержательно значимым выводом работы является модульный характер партнерского нарратива. В тематической модели он не образует единого смыслового блока, а распределяется между институционально-историческим, энергетическим и культурно-гуманитарным сегментами. Модальный и конструкционный анализы выявляют еще одну особенность: Россия представлена одновременно самостоятельным актором и объектом внешних воздействий. Негативные рамки в большинстве случаев связаны с действиями третьих сторон, что разделяет кооперативную и конфликтную линии внутри исследованного корпуса.

В нескольких тематических блоках фиксируется также устойчивое присутствие Беларуси. Это позволяет рассматривать связку «Россия + Беларусь» как самостоятельный дискурсивный паттерн, который в существующих исследованиях описан недостаточно. Представление о сугубо комплементарной или нейтральной тональности китайского официального медиадискурса частично корректируется выявлением регулярной негативной тематики. При этом внешние действия в отношении России, а не сама страна, чаще всего выступают мишенью для деструктивных конструкций.

К числу ограничений исследования относятся формирование выборки по двум прямым наименованиям России, которое могло исключить косвенные упоминания; ограниченность корпуса материалами одной газеты за трехлетний период; использование LDA-модели с «мешком слов» без отдельной качественной верификации всего тематического распре-

деления; частичная формализация модального анализа и эвристический характер дискурсивной разметки. Эти ограничения не отменяют выявленных закономерностей, но задают границы их обобщения.

В целом репрезентация России в «Жэньминь жибао» представляет собой структурированную дискурсивную систему, в которой взаимодействуют институциональные, персонализированные и оценочные механизмы. Образ России формируется не как прямое отражение реальности, а как результат дискурсивного конструирования, выявляемого через комплекс лингвистических методов.

Библиография

1. **Blei, D. M., Ng, A. Y., Jordan, M. I.** Latent Dirichlet Allocation. // Journal of Machine Learning Research. – 2003. – Vol. 3. – P. 993–1022.
2. **Hunston, S.** Corpora in Applied Linguistics. – Cambridge University Press, 2002. – 360 p. – ISBN 978-0-521-80171-3. – <https://doi.org/10.1017/CBO9781139524773>.
3. **Jurafsky, D., Martin, J. H.** Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. / ed. Russell, S., Norvig, P. – Englewood Cliffs, N.J. : Prentice Hall, 2000. – 934 p.
4. **Li, T., Pan, F.** Constructing diplomatic discourse: a corpus-driven analysis of the discursive strategies by the spokespersons of China's Ministry of Foreign Affairs during a public health crisis. // Frontiers in Psychology. – 2025. – Vol. 16. – URL: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2025.1635767/full> (accessed: 03.08.2026). – Online. – <https://doi.org/10.3389/fpsyg.2025.1635767>.
5. **McCombs, M. E., Shaw, D. L.** The Agenda-Setting Function of Mass Media. // Public Opinion Quarterly. – 1972. – Vol. 36. – № 2. – P. 176–187.
6. **Pei, J., Cheng, L.** Representations of 5G in the Chinese and British Press: A Corpus-Assisted Critical Discourse Analysis. // Humanities & Social Sciences Communications. – 2024. – Vol. 11. – URL: <https://www.nature.com/articles/s41599-024-02896-8> (accessed: 06.05.2026). – Online. – <https://doi.org/10.1057/s41599-024-02896-8>.
7. **Ван, Ф.** Особенности саморегулирования журналистики в России и Китае в период 1991–2023 гг. // Вестник Московского университета. Серия 10: Журналистика. – 2023. – № 4. – С. 42–65. – <https://doi.org/10.30547/vestnik.journ.4.2023.4265>.

8. **Го, Ж., Ян, Х.** Цветочная символика в китайской культуре на примере образа пиона. // International Journal of Humanities and Natural Sciences. – 2020. – № 3-1 (42). – С. 36–38. – <https://doi.org/10.24411/2500-1000-2020-10199>.
9. **Кирина, М. А.** Сравнение тематических моделей на основе LDA, STM и NMF для качественного анализа русской художественной прозы малой формы. // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. – 2022. – Т. 20. – № 2. – С. 93–109. – <https://doi.org/10.25205/1818-7935-2022-20-2-93-109>.
10. **Цзя, Л.** Особенности правового регулирования СМИ в Китае. // СМИ в условиях информационной глобализации: материалы Междунар. науч.-практ. конф. (Екатеринбург, 25–26 апреля 2013 г.) / сост. Д. Л. Стровский, О. В. Ильина – Екатеринбург : Урал. фед. ун-т, 2013. – С. 50–54 – URL: https://elar.urfu.ru/bitstream/10995/21688/1/sbornik_2013.pdf (дата обращения: 29.05.2026).

References

1. **Blei, D. M., Ng, A. Y., Jordan, M. I.** Latent Dirichlet Allocation. // Journal of Machine Learning Research. – 2003. – Vol. 3. – P. 993–1022.
2. **Hunston, S.** Corpora in Applied Linguistics. – Cambridge University Press, 2002. – 360 p. – ISBN 978-0-521-80171-3. – <https://doi.org/10.1017/CBO9781139524773>.
3. **Jurafsky, D., Martin, J. H.** Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. / ed. Russell, S., Norvig, P. – Englewood Cliffs, N.J. : Prentice Hall, 2000. – 934 p.
4. **Li, T., Pan, F.** Constructing diplomatic discourse: a corpus-driven analysis of the discursive strategies by the spokespersons of China's Ministry of Foreign Affairs during a public health crisis. // Frontiers in Psychology. – 2025. – Vol. 16. – URL: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2025.1635767/full> (accessed: 03.08.2026). – Online. – <https://doi.org/10.3389/fpsyg.2025.1635767>.
5. **McCombs, M. E., Shaw, D. L.** The Agenda-Setting Function of Mass Media. // Public Opinion Quarterly. – 1972. – Vol. 36. – № 2. – P. 176–187.
6. **Pei, J., Cheng, L.** Representations of 5G in the Chinese and British Press: A Corpus-Assisted Critical Discourse Analysis. // Humanities & Social Sciences Communications. – 2024. – Vol. 11. – URL: <https://www.nature.com/articles/s41599-024-02896-8> (accessed: 06.05.2026). – Online. – <https://doi.org/10.1057/s41599-024-02896-8>.

7. **Wang, F.** Features of Journalism Self-Regulation in Russia and China in the Period 1991–2023. // Moscow University Bulletin. Series 10: Journalism. – 2023. – № 4. – P. 42–65. – <https://doi.org/10.30547/vestnik.journ.4.2023.4265>.
8. **Guo, Z., Yang, H.** Floral Symbolism in Chinese Culture: The Case of the Peony Image. // International Journal of Humanities and Natural Sciences. – 2020. – № 3-1 (42). – P. 36–38. – <https://doi.org/10.24411/2500-1000-2020-10199>.
9. **Kirina, M. A.** Comparison of LDA, STM, and NMF-Based Topic Models for Qualitative Analysis of Russian Short Fiction. // Novosibirsk State University Bulletin. Series: Linguistics and Intercultural Communication. – 2022. – Vol. 20. – № 2. – P. 93–109. – <https://doi.org/10.25205/1818-7935-2022-20-2-93-109>.
10. **Jia, L.** Features of Legal Regulation of Mass Media in China. // Mass Media in the Context of Informational Globalization: Proceedings of the International Scientific and Practical Conference (Yekaterinburg, April 25–26, 2013) / comps. D. L. Strovsky, O. V. Ilyina – Yekaterinburg : Ural Federal University, 2013. – P. 50–54 – URL: https://elar.urfu.ru/bitstream/10995/21688/1/sbornik_2013.pdf (accessed: 29.05.2026).