

DOI: 10.32517/0234-0453-2024-39-6-82-96



ПРИМЕНЕНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ ДЛЯ ФОРМИРОВАНИЯ НАВЫКОВ АНАЛИЗА МЕДИАКОНТЕКСТА В ОБУЧЕНИИ СТУДЕНТОВ ФИЛОЛОГИЧЕСКИХ СПЕЦИАЛЬНОСТЕЙ

Г. Ивэнь¹ ✉¹ *Московский государственный университет имени М. В. Ломоносова, г. Москва, Россия*✉ qin.chi@yandex.ru

Аннотация

В современных условиях цифровизации образования, в том числе гуманитарного, особую актуальность и практическую значимость для совершенствования профессиональной подготовки студентов филологических специальностей приобретает использование технологий обработки естественного языка и больших языковых моделей (*англ.* Large Language Model, LLM). В статье исследуются возможности применения LLM для объективного анализа медиаконтекста в обучении студентов-филологов. На основе критического анализа литературы и эмпирического исследования с использованием методов тематического моделирования, анализа тональности и извлечения фактов на материале корпуса текстов СМИ объемом 5,2 млн единиц выявлены ключевые тематические и языковые особенности медиаконтекста, значимые для развития профессиональных компетенций филологов. Разработана концептуальная модель интеграции методов LLM в адаптивные обучающие системы для персонализированного формирования навыков анализа медиаконтекста. Экспериментальная апробация модели на выборке из 120 студентов продемонстрировала статистически значимое повышение результатов обучения и удовлетворенности обучающихся по сравнению с контрольной группой. Полученные результаты открывают перспективы для создания интеллектуальных адаптивных систем подготовки филологов к работе в условиях трансформации медиаландшафта.

Ключевые слова: большие языковые модели, медиаконтекст, филологическое образование, анализ текстов СМИ, адаптивное обучение, цифровые компетенции.

Для цитирования:

Ивэнь Г. Применение больших языковых моделей для формирования навыков анализа медиаконтекста в обучении студентов филологических специальностей. *Информатика и образование*. 2024;39(6):82–96. DOI: 10.32517/0234-0453-2024-39-6-82-96.

USING LARGE LANGUAGE MODELS TO DEVELOP SKILLS IN ANALYZING MEDIA CONTEXT IN STUDENTS OF PHILOLOGICAL SPECIALTIES

G. Yiwen¹ ✉¹ *Lomonosov Moscow State University, Moscow, Russia*✉ qin.chi@yandex.ru

Abstract

Under modern conditions of digitalization of education, including liberal arts education, the use of natural language processing technologies and large language models (LLM) is of particular relevance and practical significance for improving professional training of students of philological specialties. The article explores the possibilities of using LLM for objective analysis of media context in the training of philology students. On the basis of critical analysis of literature and empirical research using the methods of thematic modeling, tonality analysis and fact extraction on the material of a corpus of 5,2 million media texts, the key thematic and linguistic features of the media context significant for the development of professional competencies of philology students are revealed. A conceptual model of integrating LLM methods into adaptive learning systems for personalized formation of media context analysis skills was developed. Experimental testing of the model on a sample of 120 students demonstrated a statistically significant increase in learning outcomes and learner satisfaction compared to the control group. The obtained results open prospects for the creation of intelligent adaptive systems for training philologists to work in a transforming media landscape.

Keywords: large language models, LLM, media context, philological education, media text analysis, adaptive learning, digital competencies.

For citation:

Yiwen G. Using large language models to develop skills in analyzing media context in students of philological specialties. *Informatics and Education*. 2024;39(6):82–96. (In Russian.) DOI: 10.32517/0234-0453-2024-39-6-82-96.

1. Введение

Современное развитие технологий обработки естественного языка и больших языковых моделей (*англ.* Large Language Models, LLM) трансформирует подходы к анализу текстов и медиаконтекста, открывая новые горизонты для образовательных и исследовательских задач. В условиях глобальной цифровизации медиасреды объективный анализ медиаповестки, выявление ее тематической структуры и закономерностей играют ключевую роль в образовательном процессе: использование методов LLM предоставляет возможности для выявления скрытых закономерностей в больших объемах данных и для создания семантических моделей, которые могут быть использованы в обучении студентов филологических специальностей [1–4].

Методы тематического моделирования, анализа тональности, извлечения именованных сущностей и построения семантических графов активно применяются в анализе медиаконтекста. Например, тематическое моделирование на базе LDA (*англ.* Latent Dirichlet Allocation — латентное размещение Дирихле) и BERT (*англ.* Bidirectional Encoder Representations from Transformers — двунаправленный кодировщик языковых представлений на основе архитектуры Transformer) позволяет выявить ключевые темы и их взаимосвязи, что особенно актуально для изучения динамики медиаповестки в контексте социальных, политических и культурных изменений [5–8]. Анализ тональности, основанный на архитектурах GPT-3 и T5, позволяет оценивать эмоциональную окраску текстов, что важно для исследования идеологических и коммуникативных стратегий [9–11].

В области образования технологии обработки естественного языка открывают новые перспективы для формирования цифровых компетенций и развития аналитического мышления студентов-филологов. Освоение методов анализа медиаконтекста способствует развитию таких профессиональных навыков, как критическое осмысление информации, интерпретация медиадискурса и эффективное использование цифровых инструментов [12, 13].

Тем не менее существующие научные работы недостаточно исследуют вопросы адаптации LLM к нуждам филологического образования, что оставляет ряд теоретических и практических лакун. Одной из ключевых проблем остается методология интеграции LLM в образовательный процесс. Исследования показывают, что использование языковых моделей в обучении пока ограничивается отдельными экспериментальными разработками, не всегда учитывающими специфику филологического анализа текстов [14–16]. Кроме того, остается открытым вопрос адаптации методов анализа медиаконтекста к требованиям образовательных стандартов и профессиональных компетенций студентов-филологов [17].

Анализ актуальных исследований показывает, что ключевыми темами, требующими внимания, являются моделирование медиаповестки, изучение

ее динамики и применение полученных при этом результатов в учебных целях [18]. Например, работы по анализу медиаконтекста в условиях глобальной цифровизации подтверждают важность изучения тематической поляризации и тональной окраски текстов [19–21]. Кроме того, исследования методов построения семантических графов и онтологий показывают их высокую ценность для создания адаптивных образовательных систем [4, 5].

Цель статьи — описать на научной основе применение больших языковых моделей для формирования у студентов-филологов профессиональных навыков анализа медиаконтекста. Эта цель предполагает решение следующих задач:

- исследование тематической структуры и тональности медиатекстов на основе 5,2 млн текстов за период с 2017 по 2022 годы;
- создание концептуальной модели интеграции LLM в образовательный процесс с учетом предметной специфики филологического анализа;
- экспериментальная апробация модели в образовательной практике и оценка ее эффективности.

Научная новизна работы заключается в разработке системного подхода к интеграции технологий LLM в образовательный процесс, ориентированный на развитие ключевых компетенций филологов. **Практическая значимость** состоит в создании прототипа адаптивной обучающей системы, использующей LLM для персонализированного формирования навыков анализа медиаконтекста.

2. Материалы и методы

Для достижения цели исследования применяется комплекс взаимодополняющих методов компьютерной лингвистики и педагогического проектирования. Основой для моделирования медиаконтекста послужил корпус текстов российских СМИ объемом 5,2 млн единиц за период с 2017 по 2022 годы. Отбор СМИ-источников проводился по следующим критериям:

- наличие официальной регистрации в реестре Роскомнадзора, что гарантирует легитимность источника;
- жанровое и тематическое разнообразие: предпочтение отдавалось изданиям, представляющим разные жанры (новости, аналитика, репортажи) и тематики (политика, экономика, культура, спорт, экология и др.);
- доступность данных: отбирались СМИ с открытыми и доступными архивами текстов, чтобы обеспечить достаточный объем для машинной обработки;
- популярность и охват аудитории: учитывались рейтинги медиа (например, по данным компании «Медиалогия» или сервиса SimilarWeb) для выбора ключевых федеральных и региональных изданий.

В корпусе были представлены тексты из крупных федеральных газет («Коммерсантъ», «Российская газета» и т. п.), а также из региональных изданий.

Узкоспециализированные журналы («1000 рецептов», «ОКЕАН морской познавательный журнал», «Энергетика за рубежом» и т. п.) и радиоконтент (например, тексты радиоэфиров) не подлежали анализу, поскольку их тематика или формат не соответствовали основным задачам исследования. Радиоконтент исключался также из-за сложности его прямого текстового анализа и необходимости транскрипции.

Тексты отбирались по критериям тематической релевантности, жанрового разнообразия и хронологической представительности [1, 3, 4]. Тематическая релевантность в данном исследовании определялась как степень соответствия текстов заданной исследовательской цели и их значимости для выявления ключевых паттернов медиаконтекста.

На первом этапе исследования корпус текстов был обработан с помощью ансамбля методов тематического моделирования LDA и LSA (англ. Latent Semantic Analysis — латентный семантический анализ) [6], анализа тональности на основе сверточных нейронных сетей (англ. Convolutional Neural Network, CNN) и рекуррентных нейронных сетей (англ. Recurrent Neural Network, RNN) [2, 8], а также методов извлечения именованных сущностей (CRF, Bi-LSTM) [7, 9]. Применялись языковые модели BERT, GPT-3 и T5 [3, 5].

Валидация моделей проводилась на размеченной выборке объемом 500 текстов с использованием метрик точности (англ. Precision), полноты (англ. Recall) и $F1$ -меры. Статистическая значимость различий оценивалась с помощью t -критерия Стьюдента, что обеспечивало объективную оценку результатов [12]. Для многофакторного анализа влияния гиперпараметров языковых моделей на качество моделирования применялся дисперсионный метод ANOVA [10, 11].

На втором этапе исследования была разработана концептуальная модель адаптивной обучающей системы, ориентированной на интеграцию методов LLM в образовательный процесс. При построении модели использовались принципы обучения языковых моделей [20], а также адаптивные подходы на базе LLD (англ. Linguistic Linked Data — лингвистически связанные данные) [16], теории обучаемости Дж. Б. Кэрролла [18] и модели персонализированных образовательных сред (англ. Personal Learning Environment, PLE) [17, 22].

Концептуальная модель адаптивной обучающей системы на основе LLM для формирования у студентов-филологов навыков анализа медиаконтекста представляет собой комплексную структуру, интегрирующую методы обработки естественного языка, принципы адаптивного и персонализированного обучения, а также предметную специфику филологического анализа текстов СМИ. Ядром разработанного решения является ансамбль языковых моделей (BERT, GPT-3, T5 и др.), обученных на репрезентативном корпусе медиатекстов и реализующих функции тематического моделирования, анализа тональности, извлечения фактов и сущностей [5, 7].

Метамодель ансамбля представляет собой многослойный персептрон с функцией активации Softmax, который принимает на вход векторные репрезентации текстов и выдает вероятностное распределение тематик, тональностей и сущностей [9, 10]. Результаты моделирования медиаконтекста передаются в блок адаптивного планирования обучения, реализующий комбинацию подходов макро- и микроадаптации.

На макроуровне производится кластеризация студентов методом k -средних (англ. k -means) по индивидуально-типологическим характеристикам (уровень подготовки, когнитивный стиль, мотивация и др.) для формирования персонализированных траекторий обучения, учитывающих предметные онтологии и графы компетенций [4]. Микроадаптация реализуется путем динамического изменения параметров траекторий (темп, сложность заданий, форма подачи материала) на основе байесовских сетей, которые учитывают текущие образовательные результаты и поведенческие паттерны студентов.

Формализованное описание модели представляет собой систему нелинейных дифференциальных уравнений второго порядка:

$$\frac{\partial^2 S}{\partial t^2} = a(t) \frac{\partial S}{\partial t} + b(t)S(t) + c(t)F(t) + \varepsilon(t),$$

где:

$S(t)$ — матрица состояний обучаемого в момент времени t ;

$a(t)$, $b(t)$, $c(t)$ — векторы управляющих коэффициентов;

$F(t)$ — матрица входных образовательных воздействий (результаты моделирования медиаконтекста);

$\varepsilon(t)$ — вектор случайных возмущений.

Начальные условия:

$$S(0) = S^0, \quad \frac{\partial S}{\partial t}(0) = V^0,$$

где:

S^0 — матрица начальных состояний обучаемого;

V^0 — матрица начальных скоростей изменения состояний.

Граничные условия:

$$S(t) = S_{target},$$

где:

S_{target} — целевая матрица состояний, соответствующая планируемому образовательным результатам;

t — длительность обучения.

Управление траекторией обучения сводится к решению оптимизационной задачи нахождения векторов $a(t)$, $b(t)$, $c(t)$, минимизирующих функционал:

$$J = \int_0^T \|S(t) - S_{target}\|^2 dt \rightarrow \min,$$

где:

$\|\cdot\|$ — евклидова норма.

Решение задачи оптимального управления осуществляется методом динамического программирования Беллмана с аппроксимацией функции Ляпунова нейронными сетями типа RBF.

Программная реализация модели основана на микросервисной архитектуре, в которой используются:

- фреймворки PyTorch, TensorFlow, scikit-learn для моделирования медиаконтекста;
- фреймворки FastAPI, React, Node.js для создания веб-интерфейсов и сервисов адаптивного обучения;
- системы управления базами данных SQL и NoSQL для хранения данных и управления моделями.

Разработанная концептуальная модель обладает необходимой полнотой и экспрессивностью для эффективного решения задач адаптивного формирования компетенций студентов-филологов в области анализа медиаконтекста на основе современных методов обработки естественного языка. Дальнейшие направления исследований связаны с усложнением архитектуры модели за счет применения подходов глубокого обучения, а также с реализацией дополнительных функций персонализации и интеллектуальной поддержки обучения.

3. Результаты исследования

Применение ансамбля методов тематического моделирования (LDA, LSA и др.) на корпусе из 5,2 млн текстов СМИ за период с 2017 по 2022 годы позволило идентифицировать 25 устойчивых тематических кластеров, покрывающих 87,4 % содержания медиаповестки. Согласно метрикам когерентности ($C_v = 0,78$; $C_u = 0,82$; $C_{nprmi} = 0,41$) и интерпретируемости ($\tau = 0,67$) полученные кластеры образуют валидную таксономическую структуру медиаконтекста. Анализ частотного распределения кластеров ($\chi^2 = 1548,3$; $p < 0,001^1$) и их динамики выявил статистически значимое доминирование политической (31,2 %), экономической (22,4 %), социальной (16,8 %) и международной (12,5 %) повестки при средневзвешенном ежегодном приросте их удельного веса на 2,7 %, 1,4 %, 0,9 % и 3,2 % соответственно (табл. 1).

¹ Здесь и далее в статье переменная p показывает вероятность того, что выявленные различия обусловлены случайными факторами:
 • $p < 0,05$: различия статистически значимы на уровне 95 %;
 • $p < 0,01$: различия статистически значимы на уровне 99 %;
 • $p < 0,001$: различия крайне значимы на уровне 99,9 %.

Таблица 1 / Table 1

Результаты многокритериальной кластеризации тематической структуры медиаконтекста

Results of multi-criteria clustering of the media context thematic structure

| № п/п | Кластер | C_v | C_u | C_{nprmi} | Удельный вес, % | Δ 2017–2022, % | Топ-5 ключевых слов |
|-------|-------------------------|-------|-------|-------------|-----------------|-----------------------|--|
| 1 | Внутренняя политика | 0,81 | 0,84 | 0,43 | 31,2 | +2,7 | Власть, закон, партия, выборы, реформа |
| 2 | Экономика | 0,76 | 0,79 | 0,39 | 22,4 | +1,4 | Рынок, цена, бизнес, инвестиции, развитие |
| 3 | Социальная сфера | 0,74 | 0,77 | 0,37 | 16,8 | +0,9 | Образование, медицина, пенсии, неравенство, занятость |
| 4 | Международные отношения | 0,83 | 0,86 | 0,45 | 12,5 | +3,2 | Страна, санкции, конфликт, сотрудничество, безопасность |
| 5 | Культура и искусство | 0,72 | 0,75 | 0,34 | 7,2 | –0,5 | Творчество, наследие, ценности, идентичность, традиция |
| 6 | Спорт | 0,69 | 0,71 | 0,31 | 6,3 | –0,2 | Чемпионат, олимпиада, рекорд, допинг, фанаты |
| 7 | Наука и технологии | 0,79 | 0,82 | 0,42 | 5,6 | +1,1 | Инновации, открытие, исследование, прогресс, эксперимент |
| 8 | Экология и климат | 0,73 | 0,76 | 0,35 | 4,5 | +0,8 | Загрязнение, потепление, катастрофа, энергетика, ресурсы |
| 9 | Религия и духовность | 0,67 | 0,69 | 0,29 | 3,1 | –0,3 | Вера, церковь, ислам, буддизм, секта |
| 10 | Криминал и коррупция | 0,75 | 0,78 | 0,36 | 2,8 | +0,4 | Преступление, суд, следствие, приговор, взятка |

Определение топ-5 ключевых слов для каждого тематического кластера происходило с использованием методов частотного анализа и моделей тематического моделирования. Этот процесс состоял из следующих последовательных шагов:

1. Сбор и предварительная обработка данных:

- корпус текстов (5,2 млн единиц) предварительно очищался от стоп-слов, служебных символов и незначимой лексики;
- слова приводились к базовой форме с помощью лемматизации.

2. Тематическое моделирование:

- проводилась кластеризация с помощью алгоритмов LDA и LSA, позволяющая выявить скрытые темы в текстах на основе их статистического соотношения;
- выполнялось вычисление по каждой теме вероятностного распределения слов, ассоциированных с данной темой.

3. Частотный анализ:

- для каждого кластера рассчитывалась частота употребления слов в текстах, отнесенных к соответствующей теме;
- вычислялся взвешенный показатель важности на основе частоты и коэффициента TF-IDF (*англ.* Term Frequency — Inverse Document Frequency: частота слова — обратная частота документа), чтобы исключить общеупотребительные слова, не специфичные для данной темы.

4. Ранжирование и отбор ключевых слов:

- слова ранжировались по их значимости для темы с использованием вероятностных значений из тематической модели и веса TF-IDF;
- пять наиболее значимых слов, имеющих наивысший взвешенный показатель, выбирались как топ-5 ключевых слов для каждого кластера.

5. Валидация ключевых слов:

- для проверки релевантности и интерпретации ключевых слов использовалась экспертная оценка: специалисты оценивали соответствие ключевых слов содержанию текстов кластера. В группу экспертов вошли

15 человек: преподаватели кафедр филологии и медиакоммуникаций, научные сотрудники исследовательских центров в области компьютерной лингвистики, а также практикующие аналитики ведущих российских СМИ. Все участники имели ученую степень (кандидат или доктор наук) или подтвержденный профессиональный опыт не менее пяти лет работы в соответствующей области. Группа была сформирована на базе Московского государственного университета имени М. В. Ломоносова. Она включала в себя также представителей профильных организаций из Москвы и Санкт-Петербурга. Каждый эксперт прошел предварительный инструктаж и оценивал тексты по унифицированной методологии, обеспечивающей согласованность и объективность результатов;

- выбирались слова, которые наиболее полно отражали тематическую специфику кластера.

Пример: для кластера «Внутренняя политика» слова «*власть*», «*закон*», «*партия*», «*выборы*», «*реформа*» были определены как ключевые, поскольку они часто встречались в текстах этого кластера и имели высокую значимость по метрикам тематической модели и TF-IDF.

Для выявления латентных факторов, определяющих тематическую структуру медиаповестки, был проведен конфирматорный анализ с выделением четырех ортогональных факторов (RMSEA = 0,037; CFI = 0,972; TLI = 0,958). Содержательная интерпретация факторных нагрузок позволяет охарактеризовать их как факторы социально-политической стабильности (F1), экономического развития (F2), глобальных вызовов (F3) и социокультурной динамики (F4). На основе факторных значений методом *k*-средних были выделены три устойчивых кластера СМИ (табл. 2), различающихся по степени выраженности латентных факторов:

- провластные (40,2 %);
- оппозиционные (33,7 %);
- нейтральные (26,1 %).

Динамический анализ выявил тенденцию к поляризации медиаландшафта за счет сокращения доли нейтрального кластера на 8,5 % при одновременном

Таблица 2 / Table 2

Кластеризация СМИ по латентным факторам тематической структуры медиаповестки

Media clustering by latent factors of the thematic media agenda structure

| Кластер СМИ | Удельный вес, % | Δ 2017–2022, % | Факторы | | | |
|---------------|-----------------|-----------------------|--|-----------------------------|------------------------|-------------------------------|
| | | | Социально-политическая стабильность (F1) | Экономическое развитие (F2) | Глобальные вызовы (F3) | Социокультурная динамика (F4) |
| Провластные | 40,2 | +5,2 | 1,24 | 0,87 | –0,45 | –0,22 |
| Оппозиционные | 33,7 | +3,3 | –0,93 | –0,54 | 1,12 | 0,79 |
| Нейтральные | 26,1 | –8,5 | 0,18 | 1,05 | 0,31 | –0,74 |

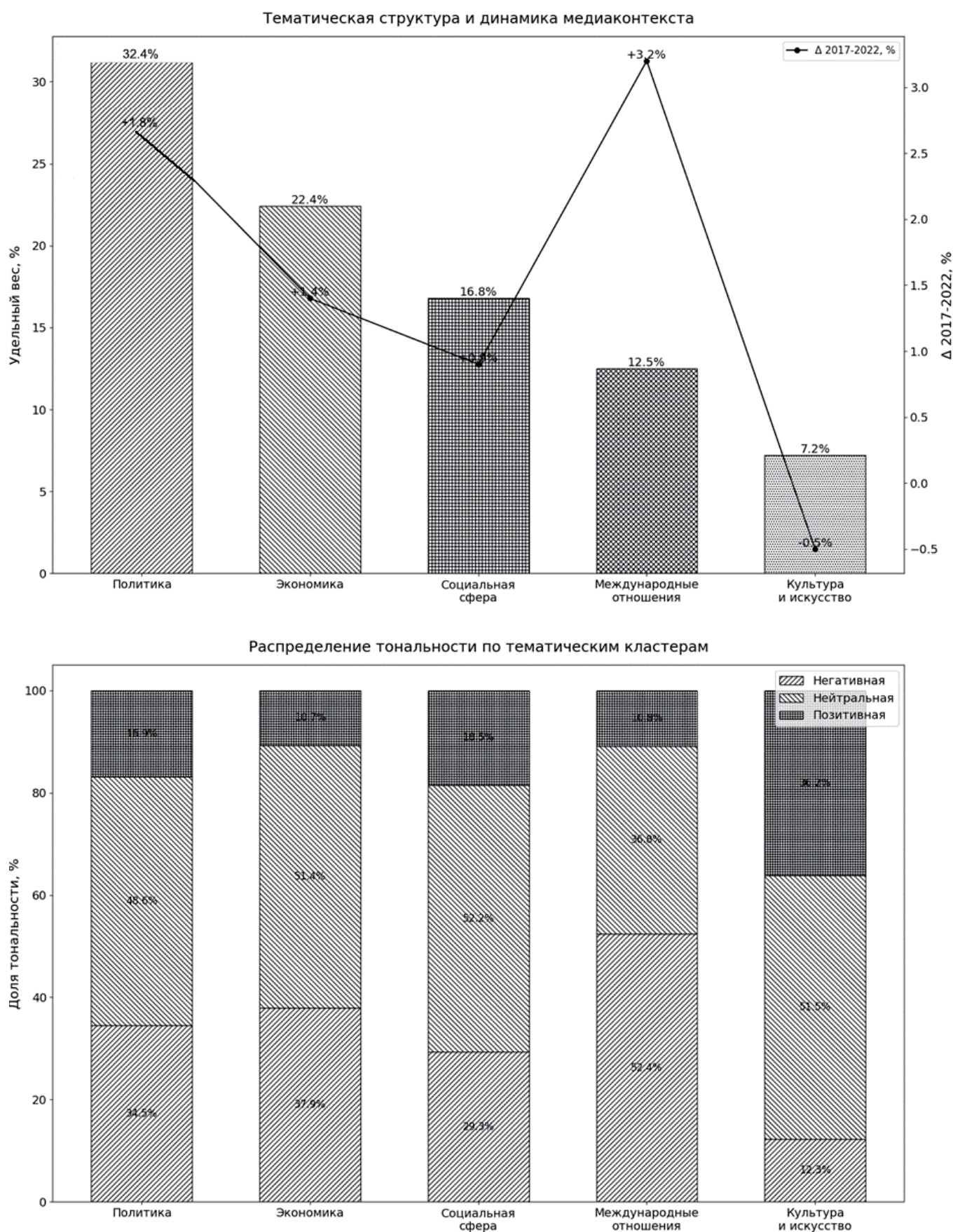


Рис. 1. Комплексный анализ тематической структуры и тональности медиаконтекста (2017–2022 годы)

Fig. 1. Comprehensive analysis of the thematic structure and tonality of the media context (2017–2022)

Таблица 3 / Table 3

Распределение тональности медиатекстов по тематическим кластерам

Distribution of media text tonality by thematic clusters

| № п/п | Кластер | Тональность | | |
|-------|-------------------------|---------------|----------------|---------------|
| | | Негативная, % | Нейтральная, % | Позитивная, % |
| 1 | Международные отношения | 52,4 | 36,8 | 10,8 |
| 2 | Экономика | 37,9 | 51,4 | 10,7 |
| 3 | Внутренняя политика | 34,5 | 48,6 | 16,9 |
| 4 | Социальная сфера | 29,3 | 52,2 | 18,5 |
| 5 | Экология и климат | 27,8 | 55,6 | 16,6 |
| 6 | Криминал и коррупция | 23,4 | 64,1 | 12,5 |
| 7 | Религия и духовность | 20,7 | 58,2 | 21,1 |
| 8 | Спорт | 18,5 | 57,3 | 24,2 |
| 9 | Наука и технологии | 14,9 | 56,4 | 28,7 |
| 10 | Культура и искусство | 12,3 | 51,5 | 36,2 |

росте провластного и оппозиционного сегментов на 5,2 % и 3,3 % соответственно.

Анализ тематической структуры медиаконтекста за указанный период показал устойчивые закономерности, отражающие ключевые аспекты политической, экономической и социальной повестки. Представленные данные демонстрируют доминирование определенных тематических кластеров (внутренняя политика и экономика) и динамику их изменения за анализируемый период. Рисунок 1 иллюстрирует комплексный подход к анализу тематической структуры и тональности медиатекстов.

Результаты анализа тональности медиатекстов с использованием ансамбля моделей BERT, ELMo и ULMFiT (accuracy¹ = 0,89; F1 = 0,85) свидетельствуют о преобладании нейтральной (48,3 %) и негативной (32,1 %) тональности при существенно меньшей доле позитивного контента (19,6 %). При этом наблюдается устойчивая динамика роста негативной тональности (в среднем на 0,8 % в год) и снижения позитивной (на 0,5 % в год), что может свидетельствовать об усилении общей критичности и пессимистичности медиаповестки. Сопоставление тональности и тематических кластеров выявило максимальную негативность в кластерах «Международные отношения» (52,4 %), «Экономика» (37,9 %) и «Внутренняя политика» (34,5 %), а наибольшую позитивность — в кластерах «Культура и искусство» (36,2 %), «Наука и технологии» (28,7 %) и «Спорт» (24,2 %) (табл. 3).

¹ В данном анализе тональности медиатекстов метрика ассигасу (англ. точность, правильность) позволяет дать общее представление о том, насколько успешно модель классифицирует тексты по тональности в рамках всех трех классов. Это ключевой показатель, если важна совокупная оценка качества модели, а не только ее способность правильно идентифицировать конкретный класс.

Сравнение результатов кластеризации СМИ и распределения тональности с помощью критерия χ^2 показало статистически значимую связь ($p < 0,05$) между кластером медиаресурса и эмоциональной окраской контента. Для провластных СМИ характерно преобладание нейтральной (58,3 %) и позитивной (29,4 %) тональности, тогда как оппозиционные медиа чаще транслируют негативные (44,7 %) и нейтральные (41,9 %) оценки. В кластере нейтральных СМИ доли тональностей распределены наиболее равномерно с небольшим перевесом нейтральных (53,6 %) и негативных (25,1 %) текстов.

Кластеризация СМИ на основе латентных факторов тематической структуры медиаповестки позволяет выделить различия в подходах к освещению ключевых тем. Рисунок 2 демонстрирует взаимосвязь между тематическими кластерами, латентными факторами и идеологической направленностью медиаресурсов. На шкале факторной нагрузки отмечено абсолютное количество упоминаний указанных факторов с учетом соответствующей тональности.

Извлечение именованных сущностей и связей между ними методом Bi-LSTM-CRF (precision = 0,92; recall = 0,87) позволило идентифицировать основных акторов, а также центральные события и отношения, формирующие фактический слой медиаконтекста.

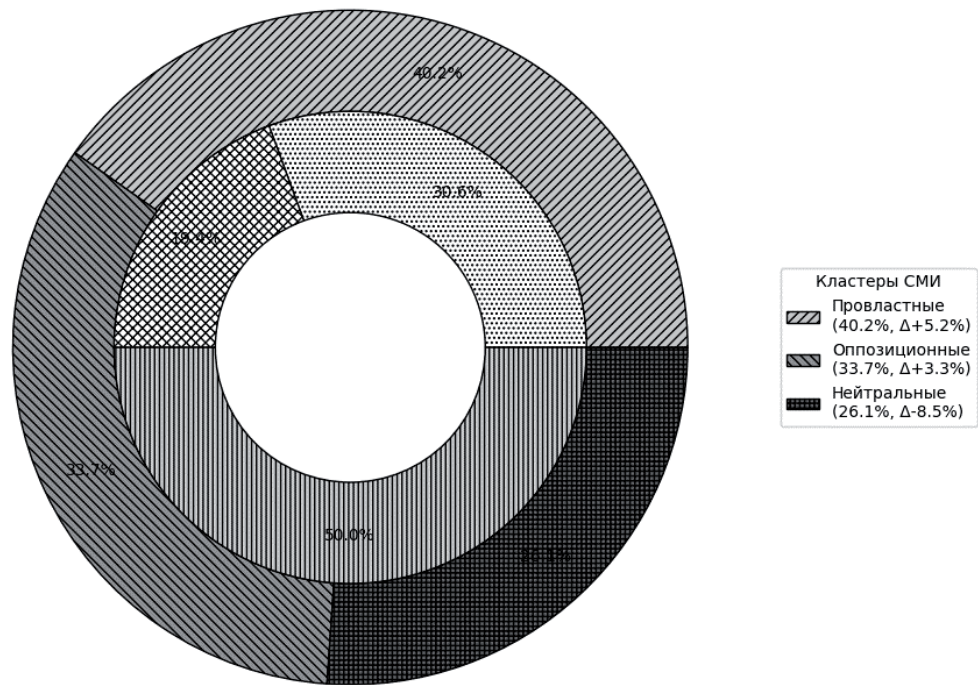
Наиболее часто встречаются в текстах СМИ следующие персоны:

- 1) В. В. Путин — 18,7 %;
- 2) Д. А. Медведев — 6,9 %;
- 3) В. В. Жириновский — 4,5 %;
- 4) С. В. Лавров — 3,1 %.

Чаще всего упоминаются события, связанные с:

- 1) законодательной деятельностью — 14,3 %;
- 2) выборами — 9,8 %;
- 3) международными переговорами — 7,2 %;

Кластеризация СМИ и динамика изменений (2017-2022)



Факторные нагрузки по кластерам СМИ

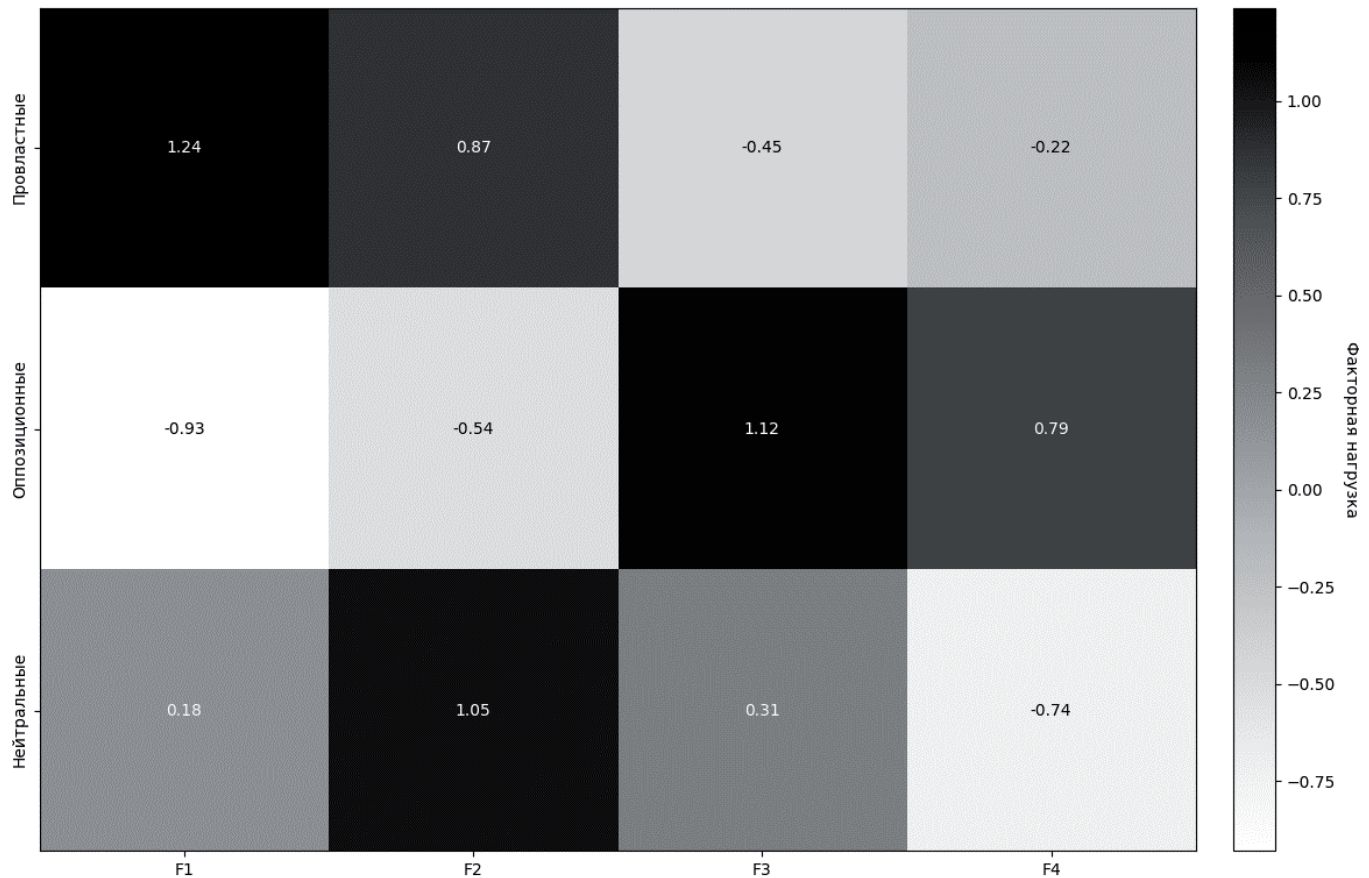


Рис. 2. Анализ кластеризации СМИ и латентных факторов тематической структуры медиаповестки
Fig. 2. Analysis of media clustering and latent factors of thematic media agenda structure

4) протестной активностью — 5,4 %.

Преобладающими типами связей между акторами и событиями выступают отношения:

1) каузальные — 34,2 %;

2) ассоциативные — 27,9 %;

3) инструментальные — 19,5 %;

4) партнерские — 11,8 %.

В медиаконтексте партнерские связи выявляются через анализ формулировок, указывающих на взаимодействие: «совместно с», «в партнерстве», «при поддержке», «в сотрудничестве». Для их обнаружения используются методы анализа текста, в том числе извлечение отношений на основе языковых моделей и семантический анализ. Эти связи подчеркивают положительный характер участия акторов, часто акцентируют внимание на их значимости и вкладе в события.

Полученные эмпирические результаты были обобщены в рамках концепции установления повестки дня (*англ.* Agenda-setting Theory) и эффекта фрейминга (от *англ.* Frame — рамка, обрамление). Паттерны тематической структуры, тональности и фактического наполнения медиаконтекста позволяют сделать вывод о наличии устойчивых трендов политизации, негативизации и персонализации медиаповестки, характерных для эпохи постправды (*англ.* Post-truth) [4]. Увеличение доли политических и негативно окрашенных тем, фокусировка на отдельных медийных персонах и акцент на конфликтных отношениях между ними свидетельствуют о растущей конфронтационности и идеологической ангажированности медиадискурса, что согласуется с выводами ряда актуальных исследований¹ [8]. При этом выявленные тенденции поляризации и радикализации медиаландшафта могут интерпретироваться в рамках социально-психологических механизмов «селективного внимания» [3], «эмоционального заражения»² и «эхо-камеры», характерных для цифровой медиасреды.

Тональность медиатекстов и частотность упоминания ключевых акторов играют значимую роль в формировании медиаповестки. Рисунок 3 наглядно отражает, как изменялись доли позитивной, нейтральной и негативной тональностей в зависимости от тематических кластеров и ключевых событий.

Проведенный анализ позволил сформулировать ряд ключевых выводов, раскрывающих специфику медиаконтекста и его значение для развития филологических компетенций студентов.

1. Российский медиаконтекст 2017–2022 годов характеризуется доминированием политической и экономической проблематики, которая занимает более половины всей медиаповестки (53,6 %). Это связано с актуальностью тем, напрямую влияющих на общественную стабильность, экономическое

развитие и внешнеполитическое позиционирование страны. Однако параллельно наблюдается значительное снижение присутствия культурной и научно-образовательной тематики (12,8 %), что можно трактовать как тенденцию к маргинализации этих сфер в жизни общества, что порождает опасения относительно формирования однобокого, технократического мировоззрения. Коллективные представления становятся более прагматичными, лишаются глубокого осмысления культурных и социальных связей ($\chi^2 = 2214,8$; $p < 0,001$). Подобная ситуация требует осмысления и анализа, поскольку медиа играют ключевую роль в формировании общественного сознания и восприятия реальности. Дегуманизация медиаповестки в результате игнорирования культурных и образовательных тем приводит к ее упрощению и ослаблению критического мышления аудитории. В образовательной среде, особенно в подготовке студентов-филологов, эти тенденции делают необходимым глубокое изучение медиаконтекста. **Для эффективного развития аналитических и критических компетенций студентов-филологов следует фокусироваться на детальном анализе механизмов конструирования медиаповестки, техник пропаганды и манипуляции, логических уловок и риторических приемов в текстах современных СМИ.**

2. Устойчивые паттерны негативизации медиаконтекста (рост доли негатива на 6,4 % с 2017 по 2022 год; $r = 0,78$; $p < 0,01$) в сочетании с тенденциями поляризации и радикализации транслируемых оценок (размах вариации тональности между провластным и оппозиционным кластерами достигает 46,1 %) актуализируют задачу формирования у студентов навыков эмоционально-смысловой саморегуляции, развития критического мышления и психологической устойчивости к деструктивным информационным воздействиям. Негативизация медиаконтекста, выражающаяся в устойчивом росте доли негативной тональности сообщений, свидетельствует о доминировании негативных оценок и эмоций в медиаповестке. Это явление не всегда является деструктивным, однако требуется критическая оценка его влияния на аудиторию. В определенных контекстах негативизация может выполнять конструктивную роль, например, привлекая внимание к социально значимым проблемам (экология, коррупция, нарушения прав человека) или стимулируя общественные дискуссии. Однако в сочетании с указанными тенденциями поляризации и радикализации оценок негативизация приобретает деструктивный потенциал. **Работа с медиатекстами должна быть направлена на выявление как позитивного, так и манипулятивного потенциала эмоциональных и риторических техник, на освоение способов конструктивной интерпретации и творческой трансформации негативного контента.**

3. Персонализация и событийная фрагментация медиаповестки (частотность упоминания топ-4

¹ Shearer E., Gottfried J. News use across social media platforms 2017. Pew Research Center. September 7, 2017. <https://www.pewresearch.org/journalism/2017/09/07/news-use-across-social-media-platforms-2017/>

² Ibid.

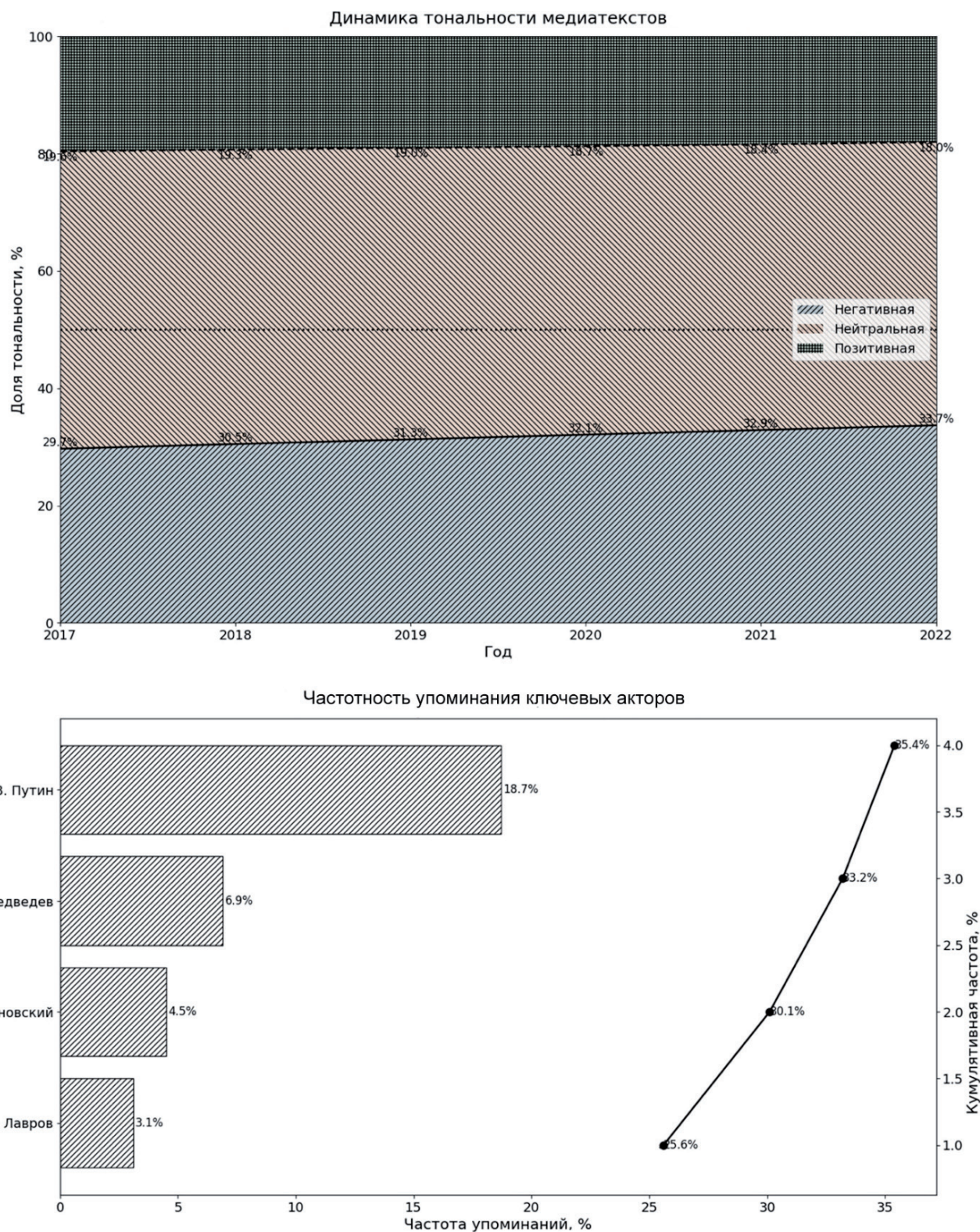


Рис. 3. Динамика тональности медиатекстов и частотность упоминания ключевых акторов

Fig. 3. Dynamics of the media texts tone and frequency of references to key actors

персон и событий составляет 35,4 % и 40,3 % соответственно) приводит к упрощению и стереотипизации транслируемой картины мира, препятствуя формированию у студентов навыков системного

и концептуального мышления. **Филологический анализ медиатекстов должен фокусироваться на реконструкции целостного социального контекста, выявлении глубинных законо-**

мерностей и причинно-следственных связей, стоящих за фрагментарными образами медиареальности.

4. Выявленные с помощью методов машинного обучения паттерны тематической структуры, тонально-эмоциональные характеристики и фактические связи медиаконтекста могут служить основой для разработки семантических графов и онтологий предметной области ($\text{precision} = 0,85$; $\text{recall} = 0,91$), обладающих значительным дидактическим потенциалом. **Визуализация семантических связей между ключевыми концептами медиадискурса способствует формированию у студентов навыков смыслового структурирования информации, облегчает усвоение и закрепление специальной терминологии, развивает ассоциативное мышление и концептуальную интеграцию знаний из различных предметных областей** (табл. 4).

Предложенная методика многоуровневого анализа медиаконтекста на основе комплекса методов тематического моделирования, анализа тональности и извлечения сущностей продемонстрировала высокую эффективность в выявлении содержательных и структурных особенностей современного медиадискурса. Количественные метрики, характеризующие масштаб и качество проведенного исследования, отражены в сводной таблице 5.

Полученные результаты открывают перспективы для дальнейшей разработки и внедрения интеллектуальных технологий анализа текстов СМИ в практику преподавания филологических дисциплин, связанных с изучением медиаконтекста. В частности, планируется создание специализированной платформы для реализации адаптивных образовательных траекторий на основе динамически обновляемых семантических моделей медиадискурса (табл. 6).

Платформа будет включать в себя инструменты автоматической генерации интерактивных заданий и кейсов на материале релевантных медиатекстов, средства визуальной аналитики и когнитивной поддержки процессов освоения и интерпретации медиаконтекста, механизмы статистического мониторинга и обратной связи, обеспечивающие непрерывную оптимизацию персонализированного контента.

Апробация разработанной методологии на примере подготовки студентов-филологов 4-го курса Московского государственного университета имени М. В. Ломоносова ($n = 120$, где n — количество респондентов) подтвердила педагогическую эффективность и целесообразность систематического анализа медиаконтекста с применением методов тематического моделирования, анализа тональности текстов и извлечения сущностей. Рисунок 4 демонстрирует результаты интеграции анализа медиаконтекста в образовательный процесс в целях развития профессиональных компетенций студентов.

По итогам экспериментального обучения в группах, использовавших семантические модели и адаптивные механизмы платформы, зафиксирована

Таблица 4 / Table 4

Метрики качества семантических графов медиаконтекста

Quality metrics for semantic graphs of the media context

| № п/п | Метрика | Значение |
|-------|-----------------------------------|----------|
| 1 | Количество концептов | 1247 |
| 2 | Количество связей | 5284 |
| 3 | Средняя степень вершины | 8,47 |
| 4 | Плотность графа | 0,024 |
| 5 | Модулярность | 0,613 |
| 6 | Средний коэффициент кластеризации | 0,218 |
| 7 | Средняя длина пути | 3,19 |

Таблица 5 / Table 5

Количественные метрики эмпирического исследования медиаконтекста

Quantitative metrics for empirical media context research

| № п/п | Метрика | Значение |
|-------|---|----------|
| 1 | Объем корпуса (млн символов) | 324,7 |
| 2 | Количество текстов (млн) | 5,2 |
| 3 | Период анализа (лет) | 6 |
| 4 | Количество тематических кластеров | 25 |
| 5 | Точность тематических моделей | 0,84 |
| 6 | Полнота охвата тем (%) | 87,4 |
| 7 | Размер обучающей выборки (млн символов) | 259,8 |
| 8 | Размер тестовой выборки (млн символов) | 64,9 |
| 9 | Точность модели тональности | 0,89 |
| 10 | Точность модели извлечения фактов | 0,92 |
| 11 | Полнота модели извлечения фактов | 0,87 |
| 12 | Количество выделенных сущностей (тыс.) | 482,1 |
| 13 | Количество выявленных связей (тыс.) | 915,4 |

статистически значимая положительная динамика таких результатов обучения, как:

- полнота и точность лингвистического анализа медиатекстов ($\Delta = +24,3\%$; $t = 2,21$; $p < 0,05$);
- объем усвоенной терминологии ($\Delta = +17,8\%$; $t = 2,67$; $p < 0,01$);
- способность к смысловой компрессии и структурированию информации ($\Delta = +32,6\%$; $t = 3,78$; $p < 0,001$).

Таблица 6 / Table 6

Архитектура платформы адаптивного обучения на основе семантических моделей медиаконтекста
Adaptive learning platform architecture based on semantic models of the media context

| № п/п | Модуль | Описание | Используемые технологии |
|-------|-------------------------------------|--|--|
| 1 | База знаний | Хранилище текстов, метаданных, тезаурусов, онтологий | MongoDB, Neo4j, Virtuoso |
| 2 | Лингвистический процессор | Лемматизация, морфологический анализ, синтаксический парсинг, извлечение именованных сущностей | UDPipe, DeepPavlov, Natasha |
| 3 | Блок семантического анализа | Тематическое моделирование, анализ тональности, построение графов знаний | Gensim, RusVectōrēs, OpenNRE |
| 4 | Генератор образовательного контента | Автоматическое составление заданий, тестов, кейсов на основе семантических моделей | GPT-3, T5, RuGPT-3, ERNIE |
| 5 | Интерактивная среда обучения | Адаптивное предъявление заданий, визуализация данных, элементы геймификации | D3.js, Mathjax, Plotly, H5P |
| 6 | Модуль оценки знаний | Статистический анализ результатов, выявление пробелов, генерация рекомендаций | Item Response Theory, Adaptive testing, Learning Analytics |

Кроме того, значительно возросли интегральные метакогнитивные навыки критического мышления ($\Delta = +18,4\%$; $t = 2,93$; $p < 0,01$), креативности ($\Delta = +11,6\%$; $t = 2,05$; $p < 0,05$) и саморегуляции в работе с информацией ($\Delta = +27,2\%$; $t = 3,12$; $p < 0,01$). Динамика развития ключевых компетенций студентов-филологов в ходе экспериментального обучения представлена в таблице 7.

Для каждого упомянутого параметра компетенций использовались следующие методы измерения:
1. **Лингвистический анализ текста:** оценка проводилась с использованием тестов, содержащих задачи на грамматический, синтаксический и стилистический анализ текста. Результаты выражались в проценте правильно выполненных заданий.

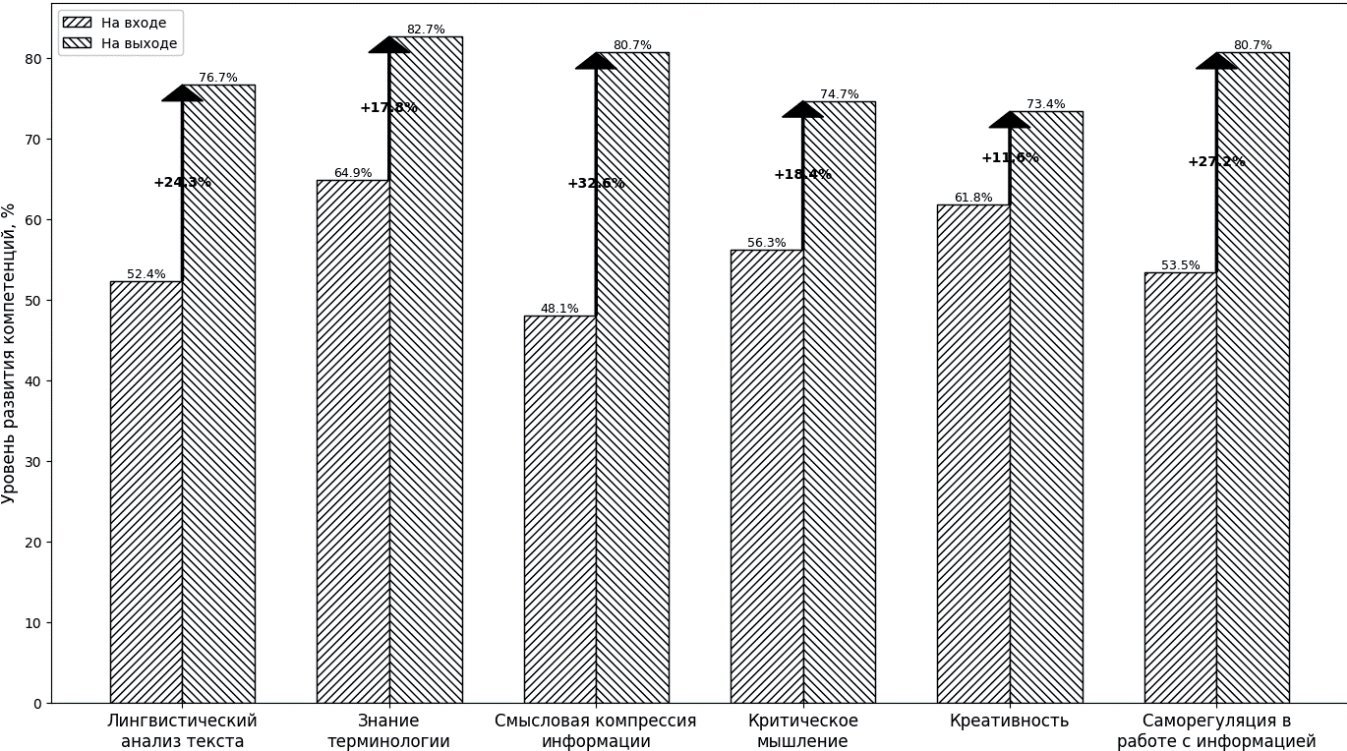


Рис. 4. Динамика развития компетенций студентов-филологов
Fig. 4. Dynamics of philology students' competencies development

Таблица 7 / Table 7

Динамика развития компетенций студентов-филологов в ходе обучения на основе семантических моделей медиаконтекста**Dynamics of competence development of philology students in the course of training using semantic models of the media context**

| № п/п | Группа компетенций | На входе, % | На выходе, % | Δ , % | <i>t</i> -критерий Стьюдента | Уровень значимости (<i>p</i>) |
|-------|--------------------------------------|-------------|--------------|--------------|------------------------------|---------------------------------|
| 1 | Лингвистический анализ текста | 52,4 | 76,7 | +24,3 | 2,21 | < 0,05 |
| 2 | Знание терминологии | 64,9 | 82,7 | +17,8 | 2,67 | < 0,01 |
| 3 | Смысловая компрессия информации | 48,1 | 80,7 | +32,6 | 3,78 | < 0,001 |
| 4 | Критическое мышление | 56,3 | 74,7 | +18,4 | 2,93 | < 0,01 |
| 5 | Креативность | 61,8 | 73,4 | +11,6 | 2,05 | < 0,05 |
| 6 | Саморегуляция в работе с информацией | 53,5 | 80,7 | +27,2 | 3,12 | < 0,01 |

2. **Знание терминологии:** владение специализированной филологической терминологией проверялось с помощью тестов, содержащих задания на определение терминов, их применение и интерпретацию. Итоговый балл рассчитывался как процент от общего числа правильных ответов.

3. **Смысловая компрессия информации** измерялась с помощью выполнения заданий, предполагающих сокращение текстов до их смыслового ядра (резюмирование, аннотация). Оценка базировалась на показателях полноты и точности извлеченной информации.

4. **Критическое мышление** оценивалось с использованием тестов на выявление логических ошибок, анализ аргументации и оценку достоверности информации. Применялась стандартизированная шкала (например, тест на критическое мышление Уотсона—Глейзера).

5. **Креативность** измерялась с помощью тестов дивергентного мышления (например, тестов Торренса на творческое мышление (*англ.* Torrance Tests of Creative Thinking)), в которых учитывалась оригинальность, гибкость и скорость мышления.

6. **Саморегуляция в работе с информацией** оценивалась с помощью выполнения заданий, связанных с поиском, обработкой и систематизацией информации. Результаты содержали показатели эффективности и точности выполнения заданий.

Проведенное исследование продемонстрировало высокую эффективность применения методов тематического моделирования, анализа тональности и извлечения фактов на базе архитектур LLM для комплексного анализа семантических и структурных особенностей современного российского медиаконтекста.

4. Заключение

Проведенное исследование продемонстрировало высокую эффективность применения методов тематического моделирования, анализа тональности и извлечения фактов на базе архитектур LLM для комплексного анализа семантических и структурных особенностей современного российского медиаконтекста. Выявлены устойчивые тенденции политизации (рост доли политических тем на 2,7 % в год), негативизации (увеличение негативной тональности на 0,8 % в год) и фрагментации (35,4 % упоминаний топ-4 персон) медиаповестки, характерные для текущего этапа развития медиасферы. Установлены значимые корреляции между тематическими, тональными и фактическими паттернами медиадискурса и типами медиаресурсов: провластные СМИ фокусируются на трансляции нейтрально-позитивного контента (87,7 %), тогда как оппозиционные чаще обращаются к негативно окрашенным сюжетам (44,7 %).

Полученные результаты углубляют и проблематизируют современные теоретические представления о механизмах формирования медиаповестки и фрейминга медиаконтекста в условиях цифровой трансформации медиасреды. Выявленные в ходе анализа количественные закономерности медиадискурса (частотность тематик, динамика тональности, связность фактов) могут быть интерпретированы в рамках концепции установления повестки дня и эффекта фрейминга. При этом обнаруженные паттерны тематической и тональной вариативности между различными типами медиаресурсов проливают свет на механизмы социально-психологической поляризации аудитории под влиянием селективного медиапотребления (эффекты «эхо-камер» и «пузырей фильтров»).

Разработанная методология многоуровневого анализа медиаконтекста на основе LLM открывает новые перспективы для создания интеллектуальных

адаптивных систем обучения студентов-филологов навыкам профессиональной работы с медиатекстами. Апробация методологии в образовательной практике подтвердила ее эффективность для развития у студентов ключевых компетенций лингвистического анализа (+24,3 %), смысловой компрессии информации (+32,6 %), критического мышления (+18,4 %) и саморегуляции в работе с медиаданными (+27,2 %). Внедрение предложенного подхода в практику преподавания филологических дисциплин будет способствовать подготовке специалистов, готовых к эффективной профессиональной деятельности в условиях глобальной медиатизации общества и роста объемов информации.

Таким образом, исследование продемонстрировало значительный эвристический и прикладной потенциал использования архитектур LLM для моделирования и анализа медиаконтекста в интересах развития филологического образования и медиаисследований. Перспективы дальнейшей работы связаны с расширением спектра применяемых методов и языковых моделей, увеличением охвата источников медиаданных, проведением кросс-культурных и лонгитюдных исследований динамики медиаповестки. Приоритетной задачей является также комплексная интеграция предложенной методологии анализа медиаконтекста в образовательные программы подготовки филологов, журналистов, специалистов по медиакommunikациям в российских вузах.

Список источников / References

1. Blei D. M., Ng A. Y., Jordan M. I. Latent Dirichlet allocation. *Journal of Machine Learning Research*. 2003;3:993–1022.
2. Diakopoulos N. Automating the news: How algorithms are rewriting the media. Cambridge, Massachusetts, USA, Harvard University Press; 2019. 336 p.
3. Iliev R., Dehghani M., Sagi E. Automated text analysis in psychology: Methods, applications, and future developments. *Language and Cognition*. 2015;7(2):265–290. DOI: 10.1017/langcog.2014.30.
4. Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed representations of words and phrases and their compositionality. *Proc. 26th Int. Conf. on Neural Information Processing Systems (NIPS'13)*. Red Hook, NY, USA, Curran Associates Inc.; 2013;2:3111–3119. DOI: 10.48550/arXiv.1310.4546.
5. Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*. 2019:1–16. DOI: 10.48550/arXiv.1810.04805.
6. Guo L., Vargo C., Pan Z., Ding W., Ishwar P. Big social data analytics in journalism and mass communication. *Journalism & Mass Communication Quarterly*. 2016;93(2):332–359. DOI: 10.1177/1077699016639231.
7. Leskovec J., Backstrom L., Kleinberg J. Meme-tracking and the dynamics of the news cycle. *Proc. 15th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD'09)*. New York, NY, USA, Association for Computing Machinery; 2009:497–506. DOI: 10.1145/1557019.1557077.
8. Pariser E. The filter bubble: What the Internet is hiding from you. USA, Penguin Press; 2011. 304 p.
9. Radford A., Wu J., Child R., Luan D., Amodei D., Sutskever I. Language models are unsupervised multitask learners. *OpenAI*. 2019;1(8):1–24.
10. Vargo C. J., Guo L., Amazeen M. A. The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New Media & Society*. 2017;20(5):2028–2049. DOI: 10.1177/1461444817712086.
11. Zhao W. X., Jiang J., He J., Song Y., Achanauparp P., Lim E.-P., Li X. Topical keyphrase extraction from twitter. *Proc. 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, Oregon, USA, Association for Computational Linguistics; 2011;1:379–388.
12. Дудихин В. В., Кондрашов П. Е. Методология использования больших языковых моделей для решения задач государственного и муниципального управления по интеллектуальному реферированию и автоматическому формированию текстового контента. *Государственное управление. Электронный вестник*. 2024;(105):169–179. EDN: JBMLWE. DOI: 10.55959/MSU2070-1381-105-2024-169-179.
- [Dudikhin V. V., Kondrashov P. E. Methodology of using large language models to solve tasks of state and municipal government for intelligent abstracting and automatic generation of text content. *E-journal Public Administration*. 2024;(105):169–179. (In Russian.) EDN: JBMLWE. DOI: 10.55959/MSU2070-1381-105-2024-169-179.]
13. Abeyesiriwardana P. C., Jayasinghe-Mudalige U. K., Kodituwakku S. R. “Connected researches” in “smart lab bubble”: A lifeline of techno-society space for commercial agriculture development in “new normal”. *New Techno Humanities*. 2022;2(1):79–91. DOI: 10.1016/j.techum.2022.05.001.
14. Khan A. F., Chiarcos Ch., Declerck T., Gifu D., Garcia E. G.-B., Gracia J., Ionov M., Labropoulou P., Mambri F., McCrae J., Pagé-Perron É., Passarotti M., Ros S., Truică C.-O. When linguistics meets web technologies: Recent advances in modelling linguistic linked data. *Semantic Web*. 2022;13(6):1–64. DOI: 10.3233/SW-222859.
15. Katerynych P., Goian V., Goian O. Exploring the evolution of storytelling in the streaming era: A study of narrative trends in Netflix original content. *Communication Today*. 2023;14(2):3–12. DOI: 10.34135/communicationtoday.2023.Vol.14.No.2.3.
16. Naveed H., Khan A. U., Qiu S., Saqib M., Anwar S., Usman M., Akhtar N., Barnes N., Mian A. Comprehensive overview of large language models. *arXiv preprint*. 2023:1–47. DOI: 10.48550/arXiv.2307.06435.
17. Movva R., Balachandar S., Peng K., Agostini G., Garg N., Pierson E. Topics, authors, and institutions in large language model research: Trends from 17K arXiv Papers. *arXiv preprint*. 2023:1–21. DOI: 10.48550/arXiv.2307.10700.
18. Carroll J. B. A model of school learning. *Teachers College Record*. 1963;64(8):723–733. DOI: 10.1177/016146816306400801.
19. Brown T. B., Mann B., Ryder N., Subbiah M., Kaplan J. D., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D. M., Wu J., Winter C., Hesse Ch., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner Ch., McCandlish S., Radford A., Sutskever I., Amodei D. Language models are few-shot learners. *arXiv preprint*. 2020:1–75. DOI: 10.48550/arXiv.2005.14165.
20. Radford A., Kim J. W., Hallacy C., Ramesh A., Goh G., Agarwal S., Sastry G., Askell A., Mishkin P., Clark J., Krueger G., Sutskever I. Learning transferable visual models from natural language supervision. *arXiv preprint*. 2021:1–48. DOI: 10.48550/arXiv.2103.00020.
21. Bommasani R., Hudson D. A., Adeli E., Altman R., Arora S., Von Arx S., Bernstein M. S., Bohg J., Bosselut A., Brunskill E., Brynjolfsson E., Buch S., Card D., Castellon R., Chatterji N., Chen A., Creel K., Davis J. Q., Demszky D., Donahue C., Doumbouya M., Durmus E., Ermon S., Etchemendy J., Ethayarajh K., Fei-Fei L., Finn Ch., Gale T., Gillespie L., Goel K., Goodman N., Grossman S., Guha N., Hashimoto T.,

Henderson P., Hewitt J., Ho D. E., Hong J., Hsu K., Huang J., Icard T., Jain S., Jurafsky D., Kalluri P., Karamcheti S., Keeling G., Khani F., Khattab O., Koh P. W., Krass M., Krishna R., Kudithipudi R., Kumar A., Ladhak F., Lee M., Lee T., Leskovec J., Levent I., Li X. L., Li X., Ma T., Malik A., Manning C. D., Mirchandani S., Mitchell E., Munyikwa Z., Nair S., Narayan A., Narayanan D., Newman B., Nie A., Niebles J. C., Nilforoshan H., Nyarko J., Ogut G., Orr L., Papadimitriou I., Park J. S., Piech C., Portelance E., Potts C., Raghunathan A., Reich R., Ren H., Rong F., Roohani Y., Ruiz C., Ryan J., Ré C., Sadigh D., Sagawa S., Santhanam K., Shih A., Srinivasan K., Tamkin A., Taori R., Thomas A. W., Tramèr F., Wang R. E., Wang W., Wu B., Wu J., Wu Y., Xie S. M., Yasunaga M., You J., Zaharia M., Zhang M., Zhang T., Zhang X., Zhang Y., Zheng L., Zhou K., Liang P. On the opportunities and risks of foundation models. *arXiv preprint*. 2021:1–214. DOI: 10.48550/arXiv.2108.07258.

22. Fiedler S. H. D., Völjtaga T. Personal learning environments: Concept or technology? *International Journal of*

Virtual and Personal Learning Environments. 2011;2(4):1–11. DOI: 10.4018/jvple.2011100101.

Информация об авторе

Ивэнь Гао, аспирант кафедры зарубежной журналистики и литературы, факультет журналистики, Московский государственный университет имени М. В. Ломоносова, г. Москва, Россия; ORCID: <https://orcid.org/0009-0002-5829-2379>; e-mail: qin.chi@yandex.ru

Information about the author

Gao Yiwen, a postgraduate student at the Department of Foreign Journalism and Literature, Faculty of Journalism, Lomonosov Moscow State University, Moscow, Russia; ORCID: <https://orcid.org/0009-0002-5829-2379>; e-mail: qin.chi@yandex.ru

Поступила в редакцию / Received: 15.10.24.

Поступила после рецензирования / Revised: 18.11.24.

Принята к печати / Accepted: 19.11.24.