



---

**ТЕОРЕТИЧЕСКАЯ, ПРИКЛАДНАЯ И СРАВНИТЕЛЬНО-СОПОСТАВИТЕЛЬНАЯ  
ЛИНГВИСТИКА/THEORETICAL, APPLIED AND COMPARATIVE LINGUISTICS**

---

DOI: <https://doi.org/10.60797/IRJ.2026.167.100> EDN: URIRQS**ЯЗЫКОВАЯ ИЕРАРХИЯ В ЭПОХУ ИИ**

Научная статья

**Пищерская Е.Н.<sup>1,\*</sup>, Бурцева А.Е.<sup>2</sup>, Лобанкова Я.А.<sup>3</sup>**<sup>1</sup>ORCID : 0000-0002-4487-3845;<sup>1</sup> Байкальский государственный университет, Чита, Российская Федерация<sup>2,3</sup> Читинский институт Байкальского государственного университета, Чита, Российская Федерация

\* Корреспондирующий автор (kishana[at]inbox.ru)

Предложена: 20.03.2026; Принята: 22.04.2026; Опубликовано: 18.05.2026

**Аннотация**

В данном исследовании рассматривается цифровая иерархия языков в эпоху ИИ, разделяющая их на высокоресурсные и низкоресурсные в зависимости от объема размеченных и неразмеченных данных в сети. Анализируются механизмы обработки естественного языка нейросетями на примере архитектуры трансформеров, а также факторы, влияющие на качество понимания языков: от морфологических особенностей до культурной предвзятости обучающих выборок. В работе поднимается проблема цифрового неравенства и необходимости сохранения языкового разнообразия в условиях технологического прогресса. Особое внимание уделяется классификации языков по уровню их цифровой представленности, что позволяет понять, почему лишь малая доля языков мира (около 5%) может активно использоваться и поддерживать свою жизнеспособность в цифровом пространстве. Целью работы является выявление факторов, определяющих цифровое неравенство языков в эпоху ИИ, и анализ причин, по которым одни языки обрабатываются нейросетями эффективнее других. В статье делается вывод о том, что технологический прогресс, при отсутствии целенаправленных усилий по инклюзивности, рискует закрепить доминирование узкой группы языков.

**Ключевые слова:** искусственный интеллект, нейросеть, промпт, язык, высокоресурсный язык, низкоресурсный язык.

**THE LINGUISTIC HIERARCHY IN THE AGE OF AI**

Research article

**Pishcherskaya E.N.<sup>1,\*</sup>, Burtseva A.Y.<sup>2</sup>, Lobankova Y.A.<sup>3</sup>**<sup>1</sup>ORCID : 0000-0002-4487-3845;<sup>1</sup> Baikal State University, Chita, Russian Federation<sup>2,3</sup> Chita Institute of the Baikal State University, Chita, Russian Federation

\* Corresponding author (kishana[at]inbox.ru)

Suggested: 20.03.2026; Accepted: 22.04.2026; Published: 18.05.2026

**Abstract**

This study examines the digital hierarchy of languages in the age of AI, categorising them as high-resource or low-resource ones depending on the volume of labelled and unlabelled data available online. The mechanisms of natural language processing by neural networks are analysed on the example of transformer architecture, as well as the factors influencing the quality of language understanding: from morphological features to cultural bias in training datasets. The paper raises the issue of digital inequality and the need to preserve linguistic diversity in the context of technological progress. Particular attention is paid to the classification of languages according to their level of digital representation, which helps to understand why only a small proportion of the world's languages (around 5%) can be actively used and maintain their viability in the digital space. The aim of the work is to identify the factors determining digital inequality among languages in the age of AI, and to analyse the reasons why some languages are processed more effectively by neural networks than others. The paper concludes that, in the absence of targeted efforts towards inclusivity, technological progress risks reinforcing the dominance of a narrow group of languages.

**Keywords:** artificial intelligence, neural network, prompt, language, high-resource language, low-resource language.

**Введение**

Исследование посвящено анализу цифровой языковой иерархии, формирующейся под влиянием стремительного развития технологий искусственного интеллекта (ИИ), в частности, больших языковых моделей (LLM) и методов обработки естественного языка (NLP). В центре внимания находится феномен цифрового неравенства языков, разделение их на высокоресурсные и низкоресурсные, а также факторы, определяющие качество и эффективность их обработки нейросетями.

Актуальность работы обусловлена всепроникающим характером ИИ в современной цифровой среде. Нейросети используются для генерации контента, перевода, поиска информации и решения широкого круга прикладных задач. В связи с этим критически важным становится вопрос о доступности и качестве языковых моделей для разных языков.

Распространено заблуждение, особенно среди российских пользователей, что использование английского языка в промптах гарантирует более точный и качественный ответ. Данное исследование направлено на выявление глубинных причин этого явления, связанных не только с техническими характеристиками моделей, но и с глобальной языковой политикой, объемом цифрового наследия и методологией сбора данных.

### Методы и принципы исследования

Теоретической и методологической базой исследования послужили работы зарубежных и отечественных специалистов в области компьютерной лингвистики, обработки естественного языка и машинного обучения.

В ходе работы были использованы такие методы исследования, как теоретический анализ и синтез (изучение научной литературы и технической документации для определения ключевых понятий); классификационный анализ (структурирование языков мира по критерию наличия размеченных и неразмеченных данных); сравнительный анализ (выявление качественных различий в обработке контекста, сравнение морфологических особенностей разных языков и их влияния на качество распознавания); метод моделирования (описание поэтапной обработки текста нейросетью на конкретном примере).

### Основные результаты

В эпоху развития технологий всё большее значение приобретают языковые инструменты для продвижения многоязыкового разнообразия во всём мире. Однако лишь малая доля языков мира (около 5% от общего количества) может активно использоваться в цифровом пространстве и поддерживать свою жизнеспособность.

Таким образом, в сетевом пространстве языки делятся на две группы: высокоресурсные и низкоресурсные. Высокоресурсными считаются те языки, которые преодолели порог в 250 тыс. статей на Википедии [1]. Это, согласно данным на февраль 2026 года: английский (7 137 220 статей), немецкий (3 097 141), русский (2 085 893), итальянский (1 956 619) и польский (1 684 874).

К низкоресурсным языкам относятся, например, южноазербайджанский (244 574 статьи), казахский (242 485), литовский (225 022), азербайджанский (211 911), грузинский (190 912).

Однако, в киберпространстве языки, с большим количеством ресурсов, делятся на размеченные и неразмеченные. К языкам с неразмеченными данными (unlabeled data) относятся языки с набором текстовой информации, не имеющей в себе меток, тэгов, раскрывающих значение. К ним относятся языки, которые представлены в виде постов в социальных сетях, аудио, изображений, статей в Википедии. В свою очередь, к языкам с размеченными данными (labeled data) относятся языки, которые имеют метку, описывающую его категорию или значение. Данное подразделение языков позволяет нейросетям быстрее обрабатывать информацию. Покажем это на следующем примере.

Представим, что у нас есть изображение кога. Само по себе оно будет относиться к неразмеченным данным, но если будет добавлен к изображению тег «кот», то изображение перейдет в категорию размеченных данных. Благодаря этому нейросети дадут правильный «ответ», избавляя ее от необходимости самостоятельно разбираться в огромных структурах данных в поисках закономерности.

В 2020 году индийские специалисты в своем исследовании «The State and Fate of Linguistic Diversity and Inclusion in the NLP World» [2] разделили языки мира на 6 классов на основе того, сколько размеченных и неразмеченных ресурсов у них есть в сетевом пространстве (показано на рисунках 1, 2).

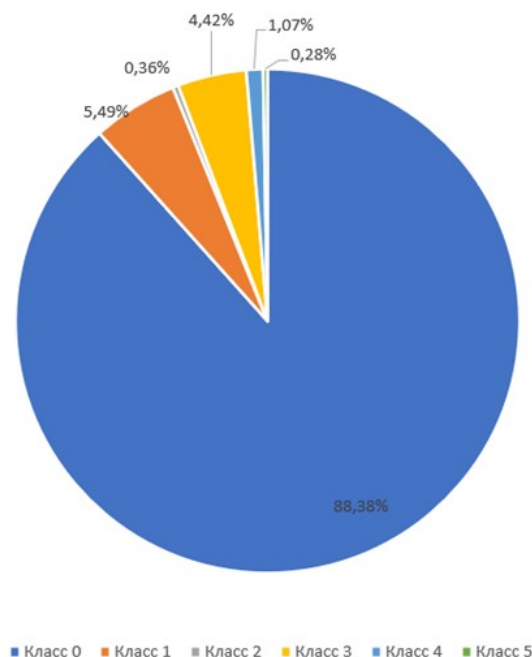


Рисунок 1 - % от общего числа языков  
DOI: <https://doi.org/10.60797/IRJ.2026.167.100.1>

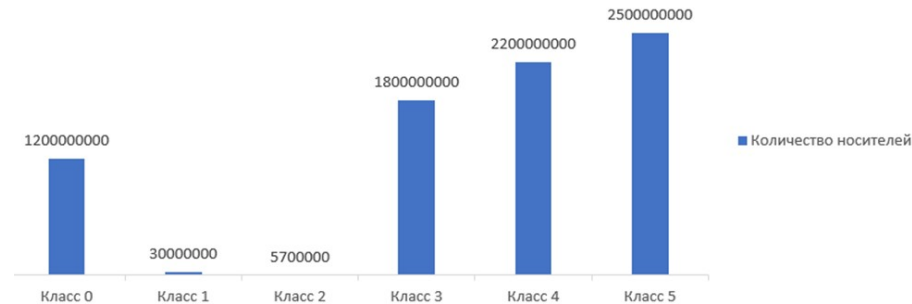


Рисунок 2 - Количество носителей языка  
DOI: <https://doi.org/10.60797/IRJ.2026.167.100.2>

Поясним данные классы.

Класс 0 — The Left-Behinds: эти языки всегда игнорировались в языковых технологиях. У них практически нет размеченных данных для использования (дахало, варлпири, пополока, валлиссский, бора и т.д. — 2191 языков).

Класс 1 — The Scraping-Bys: имеют некоторое количество размеченных данных, также есть вероятность, что в будущем повысится осведомленность об этих языках в связи с усилением сбора размеченных данных (чероки, фиджийский, гренландский, бходжпури, навахо и т.д. — 222 языков).

Класс 2 — The Hopefuls: имеют небольшой набор размеченных данных, что означает существование исследователей, стремящихся сохранить их в цифровом пространстве (зулу, конкани, лаосский, мальтийский, ирландский и т.д. — 19 языков).

Класс 3 — The Rising Stars: предварительное обучение происходило без учителя, но имеют сильную информационную базу в сети Интернет. Однако не имеют достаточное количество размеченных (индонезийский, украинский, себуано, африкаанс, иврит и т.д. — 28 языков).

Класс 4 — The Underdogs: перспективные языки, обладающие большим количеством размеченных данных, но ограниченные меньшим количеством размеченных данных. У них есть все шансы стать лидерами в цифровой среде (русский, венгерский, вьетнамский, нидерландский, корейский и т.д. — 18 языков).

Класс 5 — The Winners: языки, благодаря своим богатым ресурсам занявшие доминирующие позиции довольно давно и продолжающие стремительно развиваться (английский, испанский, немецкий, японский, французский и т.д. — 7 языков).

Обобщая вышесказанное, можно найти связь между числом размеченных и размеченных ресурсов языка, влияют на иерархическое положение в информационном пространстве.

Для понимания качества обработки информации нейросетью, также важно понимание механизма восприятия ею естественных языков.

Естественные языки — это языки, появившиеся в процессе эволюции человеческого общения, характеризующиеся гибкостью и многозначностью, которую люди могут истолковывать по-разному, а искусственный интеллект — нет. Например, фраза «Он встретил девушку с собакой» имеет двойную интерпретацию: у него была собака или девушка была с собакой. И чтобы научить нейросеть видеть все тонкости, совершенствуются методы обработки естественного языка (NLP), а также модели машинного обучения такие как LLM [3].

NLP (Natural Language Processing, обработка естественного языка) — это направление в машинном обучении, посвященное распознаванию, генерации и обработке устной и письменной человеческой речи.

LLM (Large Language Model, Большая языковая модель) — это модель машинного обучения на базе нейронной сети, называемой моделью-трансформером, на больших объемах текстовых данных.

LLM — это основа современных технологий обработки естественного языка. Их главная задача — научиться эффективно понимать смысл текстов и воспроизводить человеческий язык (YandexGPT от «Яндекса», GigaChat от «Сбера» и т.д.).

Для этого модель должна освоить как грамматику, так и скрытую семантику языка, что без NLP невозможно.

Этапы обработки естественного языка в NLP представлено на рисунке 3.



Рисунок 3 - Этапы обработки естественного языка в NLP  
DOI: <https://doi.org/10.60797/IRJ.2026.167.100.3>

Остановимся на данных этапах при рассмотрении следующего примера.

Например, в окно нейросети был введен следующий запрос: «Бабушка обещала испечь любимый яблочный пирог с корицей в воскресенье».

На этапе предобработки текста происходит:

- нормализация (приводит всё к нижнему регистру): «бабушка обещала испечь любимый яблочный пирог с корицей в воскресенье».
- токенизация (разбивает на слова): [бабушка, обещала, испечь, любимый, яблочный, пирог, с, корицей, в, воскресенье].
- удаление стоп-слов (убирает предлоги/союзы): [бабушка, обещала, испечь, любимый, яблочный, пирог, корицей, воскресенье].
- лемматизация (приводит слова к начальной форме): [бабушка, обещать, испечь, любимый, яблочный, пирог, корица, воскресенье].

После чего понимание грамматических основ происходит на этапе синтаксического анализа [4]:

1) Синтаксический анализ:

а) POS-тэггинг (части речи):

- бабушка (существительное, NOUN),
- обещала (глагол, VERB),
- испечь (глагол, VERB в инфинитиве),
- яблочный (прилагательное, ADJ),
- пирог (существительное, NOUN).

б) построение дерева зависимостей (определяет связи в предложении):

- сказуемое «обещала» связано с подлежащим «Бабушка».
- инфинитив «испечь» зависит от глагола «обещала» (что обещала?).
- дополнение «пирог» — объект действия «испечь».
- определения «любимый» и «яблочный» относятся к слову «пирог».
- обстоятельство времени «в воскресенье» связано с глаголом «обещала».

2) Семантический анализ:

а) векторные представления (эмбеддинги): векторы слов «бабушка», «испечь», «пирог» окажутся семантически близки к понятиям «семья», «готовка», «выпечка».

б) распознавание именованных сущностей (Named Entity Recognition):

- Бабушка → PER
- воскресенье → DATE

в) разрешение многозначности: система понимает, что «яблочный» в данном контексте — это вид пирога (десерт), а не сок или дерево, благодаря связи со словами «испечь» и «пирог».

3) Прагматический анализ:

а) анализ намерения: классифицирует высказывание как обещание / планы (а не как констатацию факта или инструкцию).

б) анализ тональности: определяет позитивный, семейный эмоциональный окрас (использование слов «бабушка», «любимый»).

Так как NLP обрабатывает данные на уровне предложений и фраз, то LLM использует механизмы внимания, чтобы отслеживать контекст в целых документах.

Современные LLM используют архитектуру трансформера, впервые предложенную исследователями Google Brain в 2017 году и подробно описанную в статье *Attention is All You Need* [5].

Для последующей обработки трансформером текст разбивается на токены (слова), которые, в свою очередь, преобразуются в эмбеддинги — наборы чисел, представляющих координаты вектора в многомерном пространстве. Однако при векторизации не учитывается контекст, поэтому одно и то же слово, принимающее разное значение в независимых предложениях будет иметь одинаковый эмбеддинг.

Например, в окно поиска нейросети ввели: «Сегодня мне удалили *корень* зуба» и «Нужно решить уравнение и найти его *корень*».

В этих предложениях слово «корень» имеет разный смысл, но эмбединг для него будет идентичным в обоих случаях. Решает эту проблему механизм внимания (attention).

Внимание делится на два типа: само-внимание (self-attention) и перекрестное внимание (cross-attention). Само-внимание работает с одной входной последовательностью и направлено на выявление влияния одних токенов на другие. Перекрестное внимание принимает две последовательности, устанавливая связи между их токенами (примером может служить перевод на иностранный язык). Обратимся к предложению из примера: «Нужно решить уравнение и найти его корень».

Механизм внимания преобразует эмбединг слова «корень» и переместит его вектор в многомерном пространстве ближе к вектору слова «уравнение», создав контекстное представление.

Для правильной интерпретации также необходимо учитывать позицию слова в предложении, поэтому в эмбединг добавляется кодировка порядкового номера токена (Position encoding).

В самом упрощенном понимании трансформеры работают по такому же принципу, как и Т9 на наших смартфонах: предсказывают следующее слово. Однако, Т9 опирается на частоту использования, только предшествующее слово и полностью игнорирует семантику и грамматику. Трансформеры же «смотрят» на текст целиком и выбирают самый оптимальный токен, оценивая общий контекст.

На качество понимания естественных языков нейросетями влияет не только понимание строения языковых моделей, но и некоторые факторы [6], [7].

1. *Качество токенизации.* Чем эффективнее механизм деления текста на части, тем меньше потеря смысла. Если токены более крупные и осмысленные, то они позволяют модели точнее воспринимать слова целиком, а не дробить их на отдельные символы.

2. *Объем и релевантность обучающих данных.* Нейронная сеть строит статистические закономерности на основе прочитанных текстов. Качество ответов напрямую зависит от того, насколько часто подобные образцы встречались в книгах, статьях и форумах, на которых училась модель [8].

3. *Ресурсность языка и объем обучающих данных.* Фундаментальный фактор, определяющий качество понимания модели. Если длина контекста будет увеличиваться пропорционально сложности задачи, разрыв между высокоресурсными и низкоресурсными языками может лишь возрасти. При этом наблюдается прямая зависимость с Википедией: чем больше материалов представлено на языке, тем эффективнее модель способна извлекать и обрабатывать информацию на нём.

4. *Технические особенности токенизации и архитектуры.* На способность модели удерживать смысл контекста влияет то, как именно один и тот же текст разные нейросетевые модели могут разделить на токены.

5. *Специфика токенизации и Диакритические знаки.* Для взаимодействия с языковой моделью наиболее эффективным будет выбор языка, использующего латиницу и диакритические знаки. Алфавит уже знаком модели за счет большого количества данных на английском, а включение в токенизацию диакритики позволяет нейросети лучше понять язык.

6. *Культурная предвзятость и языковое прайминг (Cultural Priming).* Нейросеть может перенимать культурные особенности в зависимости от языка промпта. Например, при запросах на китайском языке модель в ответах склонна опираться на коллективизм, а на английском — на индивидуализм. В таком случае, ответы на прикладные задачи будут различаться в зависимости от выбранного языка коммуникации.

7. *Структура инструкций и промптов.* Искусственный интеллект не понимает речь так как люди, хотя модели, обученные на больших объемах данных, способны распознавать непрямую коммуникацию, не во всех случаях они выдают желаемый результат. Для лучшего понимания задачи нейросетью промпт должен: иметь четкие и ясные формулировки; быть структурирован и разделен на логические части; использовать глаголы для конкретизации необходимых действий [9], [10].

8. *Морфологические особенности языка.* Языки с богатой морфологией позволяют моделям быстрее и эффективнее отслеживать связи между элементами текста и исключать двусмысленность за счет изобилия падежей, окончаний, изменений формы по родам, временам, лицам и т.д. Высокая информационная плотность дает возможность сделать инструкции более компактными, не теряя ясности.

## Заключение

Развитие искусственного интеллекта и технологий обработки естественного языка (NLP) не просто отражает существующую языковую реальность, но и активно формирует новую цифровую иерархию языков. Анализ языковой ситуации в киберпространстве показывает фундаментальный разрыв между высокоресурсными и низкоресурсными языками. Критерием систематизации выступает не столько количество носителей, сколько объем размеченных и неразмеченных данных, доступных для машинного обучения. Кроме того, качество понимания зависит от морфологического изобилия языка, культурной предвзятости обучающих данных и технических решений, таких как поддержка диакритических знаков.

Таким образом, будущее языкового разнообразия напрямую зависит от усилий исследователей и разработчиков. Преодоление цифрового разрыва требует совершенствования архитектур моделей с учетом морфологических особенностей разных языковых групп и борьбы с культурной предвзятостью. В противном случае технологический прогресс, вместо создания единого цифрового пространства, рискует закрепить доминирование узкой группы языков, превратив остальные в игнорируемую «цифровую тишину».

**Конфликт интересов**

Не указан.

**Рецензия**

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

**Conflict of Interest**

None declared.

**Review**

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

**Список литературы / References**

1. List of Wikipedias // Wikimedia. — 2026. — URL: [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias). (дата обращения: 21.03.26)
2. Joshi P. The State and Fate of Linguistic Diversity and Inclusion in the NLP World / P. Joshi, S. Santy, A. Budhiraja et al. // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL). — 2020. — 58. — URL: <https://arxiv.org/abs/2004.09095> (accessed: 21.03.26). — DOI: 10.48550/arXiv.2004.09095
3. Дударова Х.Х.М. Применение искусственного интеллекта в решении задач обработки естественного языка / Х.Х.М. Дударова, З.С. Фаргиева // Интернаука. — 2023. — 43-1(313). — С. 19–21.
4. Бедняк С.Г. Методы предварительной обработки для интеллектуального анализа текста / С.Г. Бедняк, Е.П. Пономарева // Наука и бизнес: пути развития. — 2023. — 12(150). — С. 48–52.
5. Vaswani A. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar et al. // CoRR. — 2017. — abs/1706.03762. — URL: <https://arxiv.org/abs/1706.03762> Это 4 (accessed: 21.03.26). — DOI: 10.48550/arXiv.1706.03762
6. Kim Y. One ruler to measure them all: Benchmarking multilingual long-context language models / Y. Kim, J. Russell, M. Karpinska et al. // COLM. — 2025. — 3. — URL: <https://arxiv.org/pdf/2503.01996> (accessed: 21.03.26).
7. Lu J. Cultural tendencies in generative AI / J. Lu, L.L. Song, L. Zhang // Nature Human Behaviour. — 2025. — 9. — P. 2360–2369. — URL: [https://www.researchgate.net/publication/391856548\\_Cultural\\_tendencies\\_in\\_generative\\_AI](https://www.researchgate.net/publication/391856548_Cultural_tendencies_in_generative_AI) (accessed: 21.03.26). — DOI: 10.1038/s41562-025-02242-1
8. Уралова Е.А. ИИАСУ'24 — Искусственный интеллект в автоматизированных системах управления и обработки данных / Е.А. Уралова, П.В. Мартынова, М.Ф. Симонов. // Сборник статей III Всероссийской научной конференции; — 1. — Москва: КДУ, 2025. — С. 104–107.
9. Язык и культура : сборник статей XXXIV Международной научной конференция (15, 16, 17 октября 2024 г.) / отв. ред. С.К. Гураль. — Томск : Издательство Томского государственного университета, 2024. — 372 с.
10. Комарова Е.В. Выбор промта для больших языковых моделей: деловые коммуникации / Е.В. Комарова // Филология и культура. — 2025. — 1(79). — С. 66–74. — DOI: 10.26907/2782-4756-2025-79-1-66-74

**Список литературы на английском языке / References in English**

1. List of Wikipedias // Wikimedia. — 2026. — URL: [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias). (accessed: 21.03.26)
2. Joshi P. The State and Fate of Linguistic Diversity and Inclusion in the NLP World / P. Joshi, S. Santy, A. Budhiraja et al. // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL). — 2020. — 58. — URL: <https://arxiv.org/abs/2004.09095> (accessed: 21.03.26). — DOI: 10.48550/arXiv.2004.09095
3. Dударова Х.Х.М. Primenenie iskusstvennogo intellekta v reshenii zadach obrabotki estestvennogo yazyka [The use of artificial intelligence in solving natural language processing problems] / X.X.M. Dudarova, Z.S. Fargieva // Interscience. — 2023. — 43-1(313). — P. 19–21. [in Russian]
4. Bednyak S.G. Metody' predvaritel'noj obrabotki dlya intellektual'nogo analiza teksta [Preprocessing methods for text mining] / S.G. Bednyak, E.P. Ponomareva // Science and business: ways of development. — 2023. — 12(150). — P. 48–52. [in Russian]
5. Vaswani A. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar et al. // CoRR. — 2017. — abs/1706.03762. — URL: <https://arxiv.org/abs/1706.03762> Это 4 (accessed: 21.03.26). — DOI: 10.48550/arXiv.1706.03762
6. Kim Y. One ruler to measure them all: Benchmarking multilingual long-context language models / Y. Kim, J. Russell, M. Karpinska et al. // COLM. — 2025. — 3. — URL: <https://arxiv.org/pdf/2503.01996> (accessed: 21.03.26).
7. Lu J. Cultural tendencies in generative AI / J. Lu, L.L. Song, L. Zhang // Nature Human Behaviour. — 2025. — 9. — P. 2360–2369. — URL: [https://www.researchgate.net/publication/391856548\\_Cultural\\_tendencies\\_in\\_generative\\_AI](https://www.researchgate.net/publication/391856548_Cultural_tendencies_in_generative_AI) (accessed: 21.03.26). — DOI: 10.1038/s41562-025-02242-1
8. Uralova E.A. IIASU'24 — Iskusstvenny'j intellekt v avtomatizirovanny'x sistemax upravleniya i obrabotki danny'x [IIASU'24 — Artificial Intelligence in automated control and data processing systems] / E.A. Uralova, P.V. Marty'nova, M.F. Simonov. // Proceedings of the III All-Russian Scientific Conference; — 1. — Moscow: KDU, 2025. — P. 104–107. [in Russian]
9. Yazyk i kul'tura : sbornik statej XXXIV Mezhdunarodnoj nauchnoj konferenciya (15, 16, 17 oktyabrya 2024 g.) [Language and culture : collection of articles XXXIV International Scientific Conference (October 15, 16, 17, 2024)] / ed. by S.K. Gural. — Tomsk : Tomsk State University Press, 2024. — 372 p. [in Russian]
10. Komarova E.V. Vy'bor promta dlya bol'shix yazykovy'x modelej: delovy'e kommunikacii [Choosing prompt for large language models: business communications] / E.V. Komarova // Philology and Culture. — 2025. — 1(79). — P. 66–74. — DOI: 10.26907/2782-4756-2025-79-1-66-74 [in Russian]