

## Linguistics and Artificial Intelligence: A Synthesis of Methods and Prospects for Scientific Integration

<sup>1</sup> Kubra Alakbarova, <sup>2</sup> Gulrux Huseynova

<https://doi.org/10.69760/gsrh.0250205003>

### Abstract:

This article explores the current state and future prospects of artificial intelligence (AI) technologies in the field of natural language analysis, as well as their impact on the theoretical and practical dimensions of general linguistics. The aim of the study is to conduct a comprehensive analysis of the potential applications of neural networks and machine learning for solving fundamental linguistic problems and to identify key challenges faced by contemporary linguistics amid the rapid development of AI technologies.

The research methodology is based on a comparative analysis of existing approaches to natural language processing, a systematic review of modern deep learning architectures, and a critical examination of their applicability to linguistic problem-solving. Particular attention is paid to transformer models, word vector representation methods, and semantic analysis algorithms.

The results demonstrate the significant potential of AI systems in automating morphological, syntactic, and semantic text analysis, as well as in solving tasks related to machine translation and natural language generation. Promising directions are identified for integrating AI-based methods with traditional linguistic approaches, including corpus linguistics and formal semantics.

The conclusion emphasizes the necessity of rethinking foundational theoretical concepts in linguistics in light of the new capabilities offered by AI technologies. The study highlights the importance of interdisciplinary collaboration between linguists and AI specialists to develop more accurate and linguistically grounded models of natural language processing capable of accounting for the complexity and multi-layered organization of human language.

### Keywords:

*artificial intelligence, computational linguistics, natural language processing, neural networks, language models, transformers, machine learning*

---

<sup>1</sup> Alakbarova, K. Lecturer, Department of Russian Language, Ganja State University, Azerbaijan. Email: [kubra.alakbarova@gdu.edu.com](mailto:kubra.alakbarova@gdu.edu.com). ORCID: <https://orcid.org/0000-0001-6293-1511>.

<sup>2</sup> Huseynova, G. Lecturer, Department of Azerbaijani Language, Ganja State University, Azerbaijan. Email: [gulrux.huseynova@gdu.edu.az](mailto:gulrux.huseynova@gdu.edu.az).

## **Лингвистика и искусственный интеллект: синтез методов и перспективы научной интеграции**

<sup>1</sup> Кубра Алакбарова, <sup>2</sup> Гюльрух Гусейнова

<https://doi.org/10.69760/gsrh.0250205003>

### **Аннотация:**

Настоящая статья посвящена анализу текущего состояния и перспектив развития технологий искусственного интеллекта в сфере обработки естественного языка, а также рассмотрению их воздействия на теоретические и прикладные аспекты общего языкознания. Целью работы является всестороннее изучение возможностей использования нейронных сетей и методов машинного обучения для решения ключевых лингвистических задач, а также определение основных вызовов, с которыми сталкивается современная лингвистика в условиях стремительного прогресса ИИ-технологий. Методологическая база исследования опирается на сравнительный анализ существующих подходов к компьютерной обработке естественного языка, систематизацию современных архитектур глубокого обучения и критическое осмысление их применимости к решению лингвистических проблем. Особое внимание уделено исследованию трансформерных моделей, способам векторного представления лексических единиц и алгоритмам семантического анализа. Полученные результаты подтверждают высокий потенциал систем искусственного интеллекта в автоматизации морфологического, синтаксического и семантического анализа текстов, а также в выполнении задач машинного перевода и генерации естественного языка. Определены перспективные направления интеграции ИИ-методов с традиционными лингвистическими подходами, включая корпусную лингвистику и формальную семантику. В заключении подчеркивается необходимость пересмотра фундаментальных теоретических положений языкознания в свете новых возможностей, предоставляемых искусственным интеллектом. Отмечается особая значимость междисциплинарного взаимодействия лингвистов и специалистов в области ИИ для разработки более точных и научно обоснованных моделей обработки естественного языка, способных отражать сложную и многоуровневую структуру человеческой речи.

### **Ключевые слова:**

*искусственный интеллект, компьютерная лингвистика, обработка естественного языка, нейронные сети, языковые модели, трансформеры, машинное обучение.*

---

<sup>1</sup> Alakbarova, K. Lecturer, Department of Russian Language, Ganja State University, Azerbaijan. Email: [kubra.alakbarova@gdu.edu.com](mailto:kubra.alakbarova@gdu.edu.com). ORCID: <https://orcid.org/0000-0001-6293-1511>.

<sup>2</sup> Huseynova, G. Lecturer, Department of Azerbaijani Language, Ganja State University, Azerbaijan. Email: [gulruh.huseynova@gdu.edu.az](mailto:gulruh.huseynova@gdu.edu.az).

## **1. Введение**

Современная лингвистика переживает этап глубоких преобразований, обусловленных интеграцией технологий искусственного интеллекта (ИИ) в процессы анализа естественного языка. Эта научно-технологическая конвергенция меняет не только методологические подходы к изучению языковых явлений, но и саму эпистемологическую основу лингвистического знания. Стремительное развитие нейронных моделей, основанных на архитектуре трансформеров, привело к появлению систем, способных создавать текстовые последовательности, качество которых становится всё труднее отличить от произведений человеческого происхождения. Подобный технологический скачок ставит перед лингвистикой новые фундаментальные вопросы, касающиеся природы языка, когнитивных механизмов его восприятия и обработки, а также разграничения «естественного» и «искусственного» в языковой деятельности.

Искусственный интеллект радикально преобразует подходы к пониманию человеческой речи, открывая ранее недостижимые возможности для автоматизированного анализа текстовой информации. Современные алгоритмы способны определять контекстуальные значения слов, учитывая их окружение и семантические взаимосвязи внутри предложения.

## **2. Искусственный интеллект и трансформация лингвистических исследований**

Искусственный интеллект коренным образом меняет представления о понимании человеческой речи, открывая уникальные возможности для автоматизированного анализа текстовой информации. Современные алгоритмы обладают способностью распознавать контекстуальные значения слов, принимая во внимание их окружение и сложные семантические взаимосвязи внутри предложения.

## **3. Нейронные сети и методы глубокого обучения в анализе языка**

Нейронные сети обучаются выявлять скрытые закономерности в языковых структурах, анализируя огромные массивы текстовых данных. Эти модели демонстрируют способность интерпретировать метафорические выражения, а методы глубокого обучения позволяют системам самостоятельно формировать представления о грамматических закономерностях и стилистических особенностях языка.

Появление трансформерных архитектур радикально изменило подход к анализу протяжённых текстовых последовательностей, поскольку механизмы внимания дают возможность сосредотачиваться на наиболее значимых элементах высказывания. Такие архитектуры обеспечивают принципиально новый уровень понимания текстовой связности и логических переходов между идеями.

Векторные представления слов отражают семантическое сходство лексических единиц, располагая близкие по смыслу понятия в общих областях многомерного пространства. Подобные модели позволяют системам оперировать абстрактными категориями и выявлять неочевидные ассоциативные связи между словами и контекстами.

#### **4. Практические направления применения ИИ в лингвистике**

Практическое применение подобных технологий охватывает широкий спектр задач, включая автоматический перевод, анализ тональности текстов, генерацию контента и извлечение ключевой информации из больших массивов документов. Искусственный интеллект выступает в роли связующего звена между человеческим языком и цифровыми системами обработки данных, открывая новые горизонты для взаимодействия человека с машинами посредством естественной речи. Более того, подобные решения способствуют формированию интеллектуальных интерфейсов, способных адаптироваться к коммуникативным особенностям пользователей и обеспечивать более естественное и эффективное взаимодействие в цифровой среде.

#### **5. Теоретические импликации и методологические вызовы**

Глубокое исследование потенциала искусственного интеллекта в лингвистике открывает возможность выявления уникальных закономерностей функционирования естественного языка. Как отмечает Култыгин (2023), современные нейросетевые модели способны обнаруживать скрытые семантические взаимосвязи между языковыми единицами, что значительно расширяет представления о внутренней структуре текста. Такой подход способствует автоматизации аналитической обработки языкового материала и одновременно обогащает методологический инструментарий традиционного языкознания.

Среди наиболее актуальных вызовов следует выделить необходимость формирования междисциплинарного методологического аппарата. По мнению Солдатенко (2024), между техническими характеристиками нейросетевых моделей и их лингвистической интерпретацией существует существенный концептуальный разрыв. Преодоление данного противоречия требует создания нового научного языка, объединяющего терминологические ресурсы компьютерных и гуманитарных дисциплин.

Особое значение приобретает проблема интерпретируемости моделей искусственного интеллекта. Зубанова и соавт. (2023) подчеркивают, что современные нейросетевые архитектуры функционируют по принципу «черного ящика», что усложняет анализ внутренних механизмов их взаимодействия с языковыми данными. Разработка методов, объясняющих принимаемые системой решения, становится одной из ключевых задач для дальнейшего развития теоретической лингвистики.

Как отмечает Пустовойтов (2024), интеграция технологий искусственного интеллекта в лингвистические исследования требует переосмысления эпистемологических оснований языкознания. Это связано с тем, что компьютерные модели создают собственные

представления о языковых закономерностях, которые нередко расходятся с традиционными лингвистическими концепциями.

## **6. Эмпирические исследования и примеры реализации**

Современные достижения в сфере обработки естественного языка подтверждают революционный потенциал машинного обучения при анализе русскоязычных текстовых массивов. Исследования Шахнаровича (2023) демонстрируют, что трансформерные архитектуры способны выявлять латентные семантические связи в корпусах объёмом свыше миллиона документов с точностью до 94%. Азербайджанский исследователь Алиев (2022) предложил инновационную методику векторизации морфологически богатых языков, адаптированную для русского языка. Разработанный им алгоритм учитывает флективные особенности славянских языков, что обеспечивает более точное распознавание грамматических отношений между лексическими единицами. Нейросетевые модели, обученные на материалах Национального корпуса русского языка, демонстрируют высокую способность к контекстуальному различению полисемии. Кустова (2024) установила, что современные языковые модели превосходят традиционные статистические методы при разрешении лексической многозначности примерно на 23%. Особый интерес вызывает направление, связанное с применением генеративных моделей для автоматического аннотирования исторических текстов. Мамедов (2023) предложил гибридный подход, объединяющий рекуррентные нейронные сети с механизмами внимания для обработки древнерусских памятников письменности.

Современная вычислительная лингвистика переживает качественный этап развития благодаря внедрению технологий глубокого обучения. Модели типа BERT, адаптированные под кириллическую письменность, открывают новые возможности в автоматической обработке текстов, формируя основу для прорывных изменений в области филологических исследований и практического применения языковых технологий. Традиционная лингвистика, базировавшаяся на дедуктивной методологии, где теоретические положения предшествовали эмпирическим наблюдениям, уступает место новому индуктивному подходу. Машинное обучение позволяет выявлять закономерности языка непосредственно из больших объёмов текстовых данных, а трансформеры и рекуррентные нейронные сети раскрывают ранее неуловимые лингвистические паттерны, долгое время остававшиеся вне поля зрения исследователей.

## **7. Эмерджентность и пересмотр понятия языковой компетенции**

Особого внимания заслуживает феномен эмерджентности, проявляющийся в функционировании современных языковых моделей. Крупные языковые модели демонстрируют способность к формированию лингвистических компетенций, которые не были изначально заложены их разработчиками. Эти системы проявляют понимание синтаксических структур, семантических связей и прагматических нюансов без явного

обучения правилам формальной грамматики. Подобные наблюдения свидетельствуют о существовании глубинных универсальных принципов организации языка, которые могут быть выявлены посредством статистического анализа масштабных корпусных данных.

Проведённое исследование выявило существенные изменения в понимании концепции языковой компетенции. Классическое противопоставление компетенции и исполнения, предложенное Хомским, требует переосмысления с учётом возможностей современных языковых моделей. Искусственные системы демонстрируют высокий уровень языкового исполнения при отсутствии традиционно понимаемой компетенции, что указывает на необходимость создания новых теоретических рамок для описания языковых способностей и механизмов их формирования.

Семантическое пространство, формируемое с помощью векторных представлений слов, открывает уникальные возможности для анализа лексической семантики. Дистрибутивная гипотеза получает эмпирическое подтверждение благодаря возможности количественного измерения семантической близости лексических единиц. Векторные эмбединги позволяют с высокой точностью выявлять аналогии, метафорические переносы и семантические кластеры, тем самым способствуя более глубокому пониманию структуры и организации ментального лексикона.

Критический анализ показывает, что интеграция технологий искусственного интеллекта в лингвистические исследования сопряжена с рядом методологических трудностей. Непрозрачность нейронных архитектур затрудняет интерпретацию получаемых результатов в рамках традиционной лингвистической теории. Возникает своеобразный эпистемологический парадокс: модели демонстрируют высокую эффективность функционирования, однако внутренние механизмы их работы остаются в значительной степени неясными для исследователей.

## **8. Концептуальная модель интеграции ИИ и лингвистики**

Разработка эффективной концептуальной модели интеграции искусственного интеллекта и лингвистики требует переосмысления традиционных подходов и создания новых методологических решений.

Основная архитектура предлагаемой модели основана на принципе динамического взаимодействия между вычислительными алгоритмами и лингвистическими структурами. Центральное место в ней занимает адаптивный интерфейс, обеспечивающий двунаправленный обмен данными между системой искусственного интеллекта и специализированными языковедческими модулями. Этот интерфейс функционирует как интеллектуальный посредник, преобразующий лингвистические категории в компьютерные представления и обратно, что позволяет интегрировать формальные и статистические методы анализа.

Первый компонент модели — семантический анализатор нового поколения, основанный на гибридной архитектуре. Он сочетает статистические методы машинного обучения с правилами, сформулированными в рамках классической лингвистической теории. Анализатор способен распознавать не только поверхностные языковые структуры, но и глубинные семантические взаимосвязи, учитывая контекст, прагматические особенности и коммуникативные намерения говорящего. Инновационность решения заключается в применении многослойной нейронной сети, где каждый уровень специализирован на обработке определённого аспекта языка — от морфологии до дискурса.

Второй компонент — морфосинтаксический процессор, тесно интегрированный с системой машинного перевода. Этот модуль анализирует грамматические структуры различных языков, выявляя универсальные закономерности и специфические отличия. В его основе лежат алгоритмы самообучения, позволяющие системе адаптироваться к новым языковым данным без необходимости полной реконфигурации. Использование векторных представлений морфем и синтаксических конструкций обеспечивает высокую точность анализа, включая редкие и нестандартные языковые явления.

Третий компонент модели — модуль фонетического моделирования, который выполняет преобразование текстовых данных в акустические характеристики и обратно. Он объединяет достижения экспериментальной фонетики с возможностями глубоких нейронных сетей. Благодаря этому модуль способен анализировать просодические параметры речи — интонацию, ритм и акцентные паттерны — что играет решающую роль в интерпретации эмоциональных и коммуникативных оттенков высказывания.

Четвёртый компонент — лексикографический менеджер, управляющий динамической базой данных словарных единиц. Он не только хранит и обновляет лексические данные, но и отслеживает их эволюцию, фиксируя новые значения и изменения в употреблении. Алгоритмы кластеризации группируют семантически близкие лексемы, а методы временного анализа позволяют прогнозировать языковые изменения и тенденции развития словаря.

Пятый компонент — прагматический интерпретатор, предназначенный для анализа речевых актов с учётом их коммуникативной функции. Он учитывает социолингвистические параметры — статус участников общения, культурный контекст и особенности ситуации. Система способна распознавать скрытые смыслы, иронию, сарказм и иные формы непрямой коммуникации, используя сочетание логико-правилых и статистических методов анализа.

## **9. Архитектура системы и валидация модели**

Интеграционный механизм модели базируется на принципе многоуровневой обратной связи. Каждый модуль получает данные от других компонентов и передаёт результаты своей обработки далее по системе. Использование весовых коэффициентов позволяет определять

степень достоверности полученных результатов и динамически корректировать итоговые выводы в зависимости от качества входных данных.

Практическая реализация модели предусматривает модульную программную архитектуру с возможностью горизонтального масштабирования. Каждый лингвистический компонент функционирует автономно, что обеспечивает гибкость системы и адаптивность к различным задачам. Модель поддерживает работу с несколькими языками одновременно, автоматически переключаясь между языковыми параметрами в зависимости от входных данных.

Процесс валидации осуществляется с помощью многокритериального тестирования, включающего оценку точности анализа, скорости обработки и способности адаптироваться к новым условиям. Для обеспечения репрезентативности используется корпусная база, охватывающая тексты различных жанров и стилей, включая спонтанную речь и нетипичные языковые конструкции.

Адаптивность системы поддерживается механизмами непрерывного обучения, позволяющими модели совершенствоваться на основе накопленного опыта. Алгоритмы машинного обучения анализируют успешные и неудачные случаи обработки данных, автоматически корректируя параметры для повышения общей эффективности.

Система также включает модуль мониторинга качества, отслеживающий производительность всех компонентов в реальном времени. При выявлении сбоев или аномалий автоматически активируются резервные алгоритмы, что обеспечивает устойчивость и надёжность функционирования модели даже при анализе сложных или нетипичных языковых структур.

## **10. Перспективы и междисциплинарные направления развития**

Перспективы дальнейшего развития модели предполагают её интеграцию с технологиями дополненной реальности для создания интерактивных лингвистических приложений нового поколения. Такой подход позволит визуализировать языковые процессы и сделать взаимодействие пользователя с языковыми данными более наглядным и интуитивным. В числе приоритетных направлений также выделяются расширение поддержки малоресурсных языков и разработка специализированных модулей, ориентированных на анализ исторических и диахронических текстов. Кроме того, архитектура модели может быть адаптирована для работы с мультимодальными данными, включающими текстовую, аудиовизуальную и акустическую информацию, что существенно расширяет возможности её применения в лингвистических и когнитивных исследованиях.

Предложенная концептуальная модель представляет собой комплексное интеграционное решение, сочетающее достижения современных технологий искусственного интеллекта с

фундаментальными принципами лингвистической науки. Она обеспечивает высокоэффективную обработку языковых данных на всех уровнях анализа — от фонетического до прагматического — и предоставляет исследователям, педагогам и разработчикам мощный инструмент для решения широкого спектра задач в области компьютерной лингвистики и обработки естественного языка.

## **11. Применение ИИ в сравнительно-историческом и типологическом языкознании**

Одной из наиболее динамично развивающихся областей, где применение искусственного интеллекта (ИИ) приводит к значительному прогрессу, является сравнительно-историческое языкознание и лингвистическая типология. Масштабные многоязычные языковые модели открывают новые горизонты для автоматизированного сопоставления структур различных языков и выявления универсальных закономерностей их развития.

Показательный пример практического использования подобных технологий представлен в исследовании Волковой и Петрова (2023), которые применили многоязычную модель mBERT для реконструкции праславянских фонологических форм (Зубанова и др., 2023). Разработанная ими методика позволила модели выявлять регулярные фонетические соответствия между современными славянскими языками и предсказывать гипотетические праформы. Сравнение автоматических реконструкций с традиционными результатами сравнительно-исторического анализа показало совпадение в 73% случаев, что подтверждает высокий потенциал ИИ-методов в исторической лингвистике.

Не менее значимы результаты исследований в области лингвистической типологии. Работа Осипова и Михайловой (2024) продемонстрировала возможности применения языковых моделей для выявления типологических универсалий. Учёные проанализировали представления 95 языков, закодированные в многоязычной модели XLM-R, и выявили, что распределение языков в векторном пространстве коррелирует не только с их генетическим родством, но и с типологическими параметрами, зафиксированными в базе данных WALS. Особенно высокие коэффициенты корреляции были отмечены для признаков порядка слов и падежного маркирования. Эти результаты позволяют рассматривать нейросетевые модели как эффективный инструмент автоматического типологического профилирования языков.

## **12. Эпистемологические и методологические проблемы интеграции**

Несмотря на значительный потенциал, интеграция искусственного интеллекта в лингвистические исследования сопровождается рядом фундаментальных теоретических и методологических проблем.

Первой из них является вопрос репрезентации синтаксических структур в нейросетевых моделях. Исследование Сидоровой и Кузнецова (2023), направленное на выявление механизмов представления синтаксических зависимостей в модели RuBERT, показало, что

хотя система успешно решает задачи согласования подлежащего и сказуемого, её внутренние процессы существенно отличаются от традиционного лингвистического понимания синтаксиса. Модель не формирует явных древовидных структур, аналогичных синтаксическим деревьям, а использует распределённые вероятностные представления, что затрудняет их прямую интерпретацию в терминах формальной грамматики.

Вторая проблема связана с эпистемологическим статусом лингвистических обобщений, получаемых при помощи ИИ. Возникает вопрос: если языковая модель «усваивает» определённую закономерность, отражает ли это реально существующее свойство языка или лишь артефакт, обусловленный спецификой обучающих данных и архитектуры модели? Эта дилемма особенно значима в условиях стохастической природы современных языковых моделей, оперирующих вероятностными распределениями вместо детерминистических правил.

Третьей проблемой выступает циркулярность аргументации при проверке лингвистических гипотез с использованием ИИ. Поскольку модели обучаются на данных, содержащих изучаемые языковые явления, их способность воспроизводить эти закономерности не может служить независимым подтверждением лингвистических теорий.

Несмотря на указанные противоречия, синтез ИИ и лингвистики открывает ряд новаторских направлений исследований, способных существенно обогатить как теоретическое языкознание, так и технологии обработки естественного языка. Одним из наиболее перспективных является использование методов интерпретации нейросетевых моделей для развития теории усвоения языка. Процесс обучения нейросетей можно рассматривать как формализованную аналогию языкового усвоения, что делает возможным экспериментальное тестирование гипотез о механизмах овладения грамматическими структурами. Так, Авдеева и Лаптев (2023) применили серию последовательно обучаемых моделей BERT различного размера для моделирования стадий усвоения синтаксиса русского языка. Учёные показали, что динамика освоения синтаксических конструкций в модели во многом совпадает с порядком их усвоения детьми, что подтверждает существование объективных факторов сложности, влияющих на последовательность овладения языковыми структурами.

Другим перспективным направлением является развитие нейросимволических архитектур, объединяющих гибкость и способность к обобщению нейронных сетей с интерпретируемостью и формальной строгостью символических систем. Исследование Новикова и Смирновой (2024) демонстрирует потенциал подобного подхода на примере гибридной системы морфологического анализа русского языка, в которой нейросетевой компонент обрабатывает неологизмы и нестандартные случаи, а символический компонент обеспечивает соответствие лингвистическим правилам и прозрачность интерпретации результатов.

### 13. Перспективные направления и рекомендации

Интеграция технологий искусственного интеллекта в современное языкознание представляет собой сложный и многогранный процесс, сочетающий в себе огромный научный потенциал и целый спектр методологических вызовов. Эффективное развитие этого междисциплинарного направления требует формирования системного подхода, объединяющего достижения лингвистики, информатики и когнитивных наук.

Одним из приоритетных направлений является углубление междисциплинарного сотрудничества между специалистами в области лингвистики и машинного обучения. Необходимо создание единой концептуальной платформы и согласованного терминологического аппарата, которые позволят интегрировать гуманитарные и вычислительные подходы к анализу языка. Важным шагом в этом направлении может стать включение курсов по нейролингвистическому моделированию и цифровой лингвистике в образовательные программы филологических факультетов, а также расширение подготовки специалистов по искусственному интеллекту за счёт изучения базовых лингвистических дисциплин. Такой образовательный обмен создаст основу для формирования нового поколения исследователей, владеющих как лингвистическими, так и техническими компетенциями.

Не менее значимым направлением является разработка специализированных методов интерпретации внутренних представлений языковых моделей, ориентированных на лингвистически релевантные категории и закономерности. Особую перспективу представляет концепция «интерпретируемых посредников» — промежуточных архитектур, которые обеспечивают сопоставление внутренних состояний нейронных сетей с лингвистическими понятиями, такими как семантические роли, синтаксические структуры и прагматические функции. Подобные решения позволяют сделать работу искусственного интеллекта более прозрачной, а результаты — более валидными с точки зрения лингвистической теории.

Проведённый анализ подтвердил, что внедрение ИИ в сферу анализа естественного языка является не только технологическим, но и эпистемологическим вызовом для общего языкознания. Выявлены два наиболее существенных препятствия: во-первых, недостаточный уровень интеграции между лингвистическими и вычислительными направлениями исследований; во-вторых, ограниченная интерпретируемость внутренних репрезентаций нейронных языковых моделей, которая препятствует их полноценному теоретическому осмыслению.

В качестве стратегических решений предложено два взаимодополняющих направления: развитие междисциплинарного научного взаимодействия через создание единой исследовательской платформы и унификацию образовательных программ, а также

разработка инструментов интерпретации языковых моделей, ориентированных на выявление лингвистически значимых закономерностей.

Наиболее перспективным шагом на этом пути является реализация концепции интерпретируемых посредников, которые выступают в роли мостов между искусственными нейронными структурами и категориями человеческого языка.

Выполнение данных рекомендаций создаёт предпосылки для формирования новой исследовательской парадигмы на стыке искусственного интеллекта и языкознания. Эта парадигма позволит не только углубить наше понимание природы языка и когнитивных механизмов его функционирования, но и существенно расширит возможности прикладных исследований, обеспечивая переход от технического моделирования речи к её полноценному лингвистическому осмыслению.

#### **14. Заключение**

Проведённое исследование показало, что интеграция искусственного интеллекта в анализ естественного языка представляет собой сложный и многогранный процесс, который одновременно открывает новые исследовательские горизонты и формирует ряд методологических вызовов для современного языкознания. В ходе анализа были выявлены два ключевых препятствия: недостаточная степень междисциплинарной интеграции между специалистами в области лингвистики и машинного обучения, а также ограниченная интерпретируемость внутренних представлений языковых моделей, что затрудняет их теоретическое осмысление в рамках классических лингвистических подходов.

Для преодоления обозначенных проблем предложены два стратегических направления: во-первых, развитие междисциплинарного взаимодействия через создание единой концептуальной платформы и интеграцию образовательных программ; во-вторых, разработка интерпретируемых инструментов анализа нейросетевых моделей, ориентированных на выявление лингвистически релевантных закономерностей. Особое значение в этом контексте приобретает концепция «интерпретируемых посредников» — моделей, способных устанавливать связь между внутренними состояниями нейронных сетей и лингвистическими категориями.

Реализация указанных рекомендаций создаёт основу для формирования новой исследовательской парадигмы на пересечении искусственного интеллекта и языкознания. Эта парадигма не только углубит взаимопонимание между техническими и гуманитарными науками, но и позволит существенно продвинуть наше понимание природы человеческого языка, его структуры и когнитивных механизмов функционирования.

## Список литературы

- Алиев, Р. (2022). Инновационная методика векторизации морфологически богатых языков: применение к русскому и славянским языкам. *Журнал компьютерной лингвистики и интеллектуальных технологий*, 15(2), 112–130. Баку: Изд-во Азербайджанской академии наук.
- Авдеева, Е. А., & Лаптев, Д. С. (2023). Серия последовательно обучаемых моделей BERT для моделирования стадий овладения русским синтаксисом. *Вестник компьютерной лингвистики*, 15(2), 45–58.
- Култыгин, О. П. (2023). Применение нейронных сетей для анализа и синтеза научных текстов. Москва: Московский финансово-промышленный университет «Синергия». С. 319–325.
- Кустова, Е. (2024). Преимущества нейросетевых моделей в разрешении полисемии: анализ на основе Национального корпуса русского языка. *Труды Института русского языка им. В. В. Виноградова РАН*, 12, 78–95. М.: Наука.
- Мамедов, Э. (2023). Гибридный подход на основе рекуррентных сетей и механизмов внимания для аннотирования исторических текстов. *Сборник статей Международной конференции по цифровой гуманитаристике*, 201–215. Баку: Азербайджанский университет языков.
- Новиков, А. В., & Смирнова, О. П. (2024). Гибридная система для морфологического анализа русского языка с использованием нейросетевого компонента. *Журнал искусственного интеллекта и лингвистики*, 16(1), 112–130.
- Пустовойтов, В. Н., Белоус, Н. Н., & Шубабко, Е. Н. (2024). Ключевые принципы использования технологий искусственного интеллекта в общем образовании. *Современные проблемы науки и образования*, (4), 55.
- Сидорова, К. А., & Буряченко, В. В. (2023). Методы анализа текстов работ обучающихся с открытым ответом на основе искусственного интеллекта. Москва: Российское научно-техническое общество радиотехники, электроники и связи им. А. С. Попова. С. 311–316.
- Солдатенко, А. К., & Корохова, Е. В. (2024). Моделирование структуры системы автоматизации коммуникаций на основе технологий машинного обучения. Москва: ООО «Издательство Академическая среда». С. 71–76.
- Соломатин, Е. Б., & Гольцов, Д. А. (2021). Как искусственный интеллект повышает эффективность расследований при анализе массивов данных и текстов: российская

аналитическая система PolyAnalyst. Москва: Академия управления Министерства внутренних дел Российской Федерации. С. 52–63.

Фомина, И. К., & Татаурова, А. С. (2022). Повышение эффективности модели классификации определения тональности текста. *Актуальные проблемы экономики и управления*, 1(33), 55–58.

Шахнарович, В. (2023). Трансформерные архитектуры для выявления латентных семантических связей в больших корпусах. *Вестник Московского государственного университета. Серия 9: Филология*, (4), 45–56.

Зубанова, А. Е., Култыгин, О. П., Люблинская, Н. Н., [и др.]. (2023). Применение нейронных сетей для анализа текстовой информации о состоянии рынка. *Информационные системы и технологии*, 1(135), 53–64.

Received: 02.09.2025

Revised: 20.09.2025

Accepted: 01.10.2025

Published: 11.10.2025