

Риски применения искусственного интеллекта в лингвистической экспертизе

Е. С. Славинская

Кемеровский государственный университет

ул. Красная, 6, 650000, Кемерово, Россия. E-mail: slavinskaja.elizaveta@mail.ru

Статья рассматривает возможности применения технологий искусственного интеллекта в автороведческой экспертизе текста. Анализируются существующий опыт и преимущества использования нейросетей и методов компьютерной лингвистики для анализа текста. Нейросеть рассматривается как перспективный инструмент при проведении комплексного анализа – она исключает необходимость последовательного анализа текста по каждому из аспектов с ручным выделением каждого значимого признака и поиском его по тексту. Это существенно ускоряет процесс и облегчает работу лингвиста-эксперта. Рассмотрены технологические, юридические и этические риски, возникающие при внедрении искусственного интеллекта в экспертную практику. Среди них: непрозрачность и нестабильность существующих нейросетей, высокий риск возникновения «галлюцинаций», риск утечки данных, риск манипулируемости входными данными, риск некорректного использования результатов. Приводятся нормативные документы, принятые в Российской Федерации, регулирующие применение нейросетевых технологий в социально значимых областях, позволяющие в том числе снизить вышеупомянутые риски. Показано, что нейросетевые технологии могут служить эффективным вспомогательным инструментом, однако требуют строгого соблюдения профессиональных стандартов, контроля рисков, прозрачности алгоритмов и экспертной проверки результатов.

Ключевые слова: искусственный интеллект, нейросети, лингвистическая экспертиза, юридическая лингвистика, управление рисками.

Prospects and Risks of Using Artificial Intelligence in Authorship Attribution

E. S. Slavinskaya

Kemerovo State University

6 Krasnaya st., 650000, Kemerovo, Russia. E-mail: slavinskaja.elizaveta@mail.ru

The article examines the possibilities of applying artificial intelligence technologies in authorship attribution and forensic linguistic text analysis. The existing experience and advantages of using neural networks and computational linguistics methods for text analysis are analyzed. Neural networks are considered a promising tool for comprehensive text analysis, as they eliminate the need for sequential examination of individual textual features through manual identification and search, thereby significantly accelerating the analytical process and facilitating the work of forensic linguists. The paper discusses technological, legal, and ethical risks arising from the implementation of artificial intelligence in expert practice, including the non-transparency and instability of existing neural network models, the high risk of hallucinations, data leakage risks, susceptibility to input manipulation, and the risk of improper use of generated results. The study also presents regulatory documents adopted in the Russian Federation that govern the application of neural network technologies in socially significant fields and contribute to mitigating the identified risks. It is shown that neural network technologies can serve as an effective auxiliary tool; however, their use requires strict compliance with professional standards, risk control measures, algorithmic transparency, and expert verification of results.

Key words: artificial intelligence, neural networks, forensic linguistic expertise, legal linguistics, risk management.

В последние годы искусственный интеллект и машинное обучение демонстрируют стремительный рост и все более активно интегрируются в различные аспекты современной жизни. Благодаря стремительному развитию самообучающихся технологий, нейронные сети доступны для использования не только в узкоспециализированных отраслях, они стремительно входят в повседневную жизнь, становясь доступными каждому, у кого есть доступ

в интернет. Области их профессионального применения также постоянно расширяются, не только охватывая технические и производственные сферы, но и открывая новые горизонты в областях, тесно связанных с гуманитарным знанием.

Благодаря значительным успехам в области алгоритмов обработки естественного языка, исследователи получили возможность эффективно работать с огромными объемами текстовых данных, автоматизируя рутинные элементы анализа и статистической обработки информации, что существенно повышает продуктивность и позволяет выявлять скрытые закономерности [Григорьева 2020; Романов 2009]. Перспективным является использование ИИ и в юридической лингвистике [Галюченко 2024], в том числе и в рамках автороведческой экспертизы текста. Нейросеть, обученная на произведениях естественной речи, получившая информацию о способах определения стилистических, статистических, графических и других особенностей текста, а также получившая корректный промпт (задача для нейросети с подробными требованиями к результату), сможет определить и сопоставить те или иные закономерности текста, являющиеся проявлением уникальной языковой личности автора. Применение нейросетей позволяет ускорить и автоматизировать как анализ текстов (поиск тех или иных отличительных черт, количественный и качественный их анализ), так и синтез результатов (составление общего портрета языковой личности автора текста).

В научной литературе описаны опыты анализа текстов, показывающие, что нейронные сети могут быть успешно использованы для рассмотрения отдельных аспектов текста, которые традиционно являются предметом внимания лингвиста-эксперта. В работе П. Е. Беловой и А. К. Сафарян представлена система FindImper, предназначенная для автоматизации юрислингвистического анализа текстов на русском языке, в частности, для выявления в них высказываний побуждающего характера, таких как просьбы, приказы или призывы. Система, основанная на анализе глагольных форм и синтаксических связей, призвана оптимизировать работу экспертов-лингвистов и может быть использована для обнаружения потенциально опасных текстов в интернете [Белова 2024]. В работе Г. Н. Трофимовой рассматривается применение метода семантического дифференциала в лингвистической экспертизе для выявления неконгруэнтности – рассогласованности значений языковых единиц. Показано, что этот метод, основанный на сопоставлении семантики синонимов и антонимов, позволяет объективно выявлять и обосновывать наличие негативных оценок в спорных медиатекстах [Трофимова 2024].

Перспективным является определение возможностей нейросетей проводить комплексный анализ текста с использованием формальных и экспертных методов. Чем больше разнообразных методов будет использовано при анализе, тем большее количество отличительных черт, особенностей и потенциальных черт портрета языковой личности можно будет выделить в итоге. Соответственно, тем более точным будет языковой портрет автора, тем больше возможностей для объективной идентификационной лингвистической экспертизы. Преимуществом использования нейросети является то, что она позволяет рассмотреть сразу большое количество аспектов текста, выявляя закономерности в нем. Нейросеть успешно работает с естественным текстом [Аль-Нами 2023; Ширяева 2024; Martinc 2024; Verma 2025; Zucco 2025], может оценить его с качественной и количественной точки зрения, она может проводить анализ с использованием как формальных методов, так и методов, аналогичных экспертным (благодаря словарной информации определять речевые маркеры, оценочную лексику, экспрессивные средства, оценивать контекст и дискурс, в котором существует текст). За границей уже проводятся исследования, направленные на юрислингвистические анализы текста при помощи ИИ [Arazzi 2025; Alkhodhairy 2025]. Важным аспектом является удобство. Нейросеть исключает необходимость последовательного анализа текста по каждому из аспектов с ручным выделением каждого значимого признака и поиском его по тексту. В отличие от классического подхода, где требуется использовать множество отдельных инструментов и сводить результаты вручную, нейросеть может выполнять весь комплекс анализа в рамках единой системы, что существенно ускоряет процесс и снижает риск ошибок при интерпретации данных. Это открывает новые возможности для повышения как продуктивности работы лингвиста-эксперта, так и достоверности результатов лингвистической экспертизы.

Однако важно понимать, что нейросетевые технологии не выступают заменой эксперту и их использование должно проходить осознанно. Использование ИИ связано с рядом существенных рисков, которые необходимо предупреждать и минимизировать. Первая группа рисков – технологические. К ним можно отнести непрозрачность и нестабильность существующих нейросетей, подходящих для анализа текста. При использовании коммерческих, общедоступных нейросетей невозможно точно охарактеризовать алгоритм, по которому проходит анализ текста, нет доступа к информации о специфике предложения, для пользователя нейросеть – «черный ящик». Обновление версий инструмента разработчиком также может привести к нестабильности результатов. Также высок риск возникновения «галлюцинаций» – ИИ склонен давать ответ в любой ситуации. Если ИИ не способен найти нужную информацию, он генерирует правдоподобную, но не связанную с реальными фактами информацию, которая должна удовлетворить запрос. Также ИИ стремится соглашаться с пользователем, даже если это утверждение не совсем соответствует реальности, однако этот риск можно минимизировать – формулировать промпт так, чтобы не было спорных вопросов, на которые можно ответить отрицательно или положительно.

Вторая группа рисков – юридические. В первую очередь, использование коммерческих, открытых инструментов несет в себе риск утечки данных, что может стать критической проблемой при работе в юридической сфере, с личными данными. Кроме того, материалы для экспертизы могут быть закрытыми, их использование в открытой сети Интернет не представляется возможным. Кроме того, для юрислингвистики актуален вопрос соответствия инструментов юридическим требованиям и рекомендациям. Перед лингвистом-экспертом в ходе в том числе судебного заседания могут поставить вопрос о необходимости и обоснованности использования того или иного

инструмента. Эксперт не может использовать ненадежный и не закрепленный в практике метод.

Третья группа – этические риски. Существует риск манипулируемости входными данными, но он решается разработкой универсального промпта, либо нейросети менее чувствительной к промпту и более чувствительной непосредственно к данным. Также возникает риск некорректного использования результатов – выводы ИИ могут восприниматься как объективные и окончательные, применяться как результаты экспертизы, без оценки и корректировки специалистом. Важен также в целом риск несоответствия этическим стандартам – морально и этически значимые вопросы нельзя решать сугубо при помощи ИИ.

Игнорирование данных рисков не просто снижает эффективность работы лингвиста, оно может привести к критическим ошибкам и получению результатов, не имеющих ничего общего с достоверными.

Использование ИИ в юрислингвистике должно быть подчинено строгим принципам и нормам, которые уже определены на уровне законодательства, профессиональных стандартов и общественной морали. Так, ПНСТ 840-2023 «ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ Обзор этических и общественных аспектов» отмечает следующее: «В основе создания этичного и устойчивого ИИ лежат следующие элементы, ориентированные на данные: - сбор данных (включая средства или способы такого сбора данных); - подготовка данных; - контроль прослеживаемости; - контроль доступа и обмена данными (аутентификация); - защита данных; - контроль хранения (добавление, изменение, удаление); - качество данных. Эти элементы влияют на объяснимость, прозрачность, безопасность и конфиденциальность, особенно в случаях создания, контроля или обработки персональной идентифицируемой информации». Также в ПНСТ 840-2023 отмечаются основные принципы создания и использования ИИ, в том числе: подотчетность, справедливость и недискриминация, прозрачность и объяснимость, конфиденциальность и защита частной жизни, приватность, контроль технологии со стороны человека, уважение верховенства закона. Для использования нейросетей в юрислингвистике и минимизации ранее описанных рисков перспективной является разработка кодекса, опирающегося на данные принципы.

Этические рамки, регулирующие применение ИИ, призваны обеспечить защиту прав и свобод граждан, предотвратить дискриминацию и гарантировать прозрачность и подотчетность принимаемых решений. В частности, в пункте 5.18 ГОСТ Р 71476-2024 «Национальный стандарт Российской Федерации. Искусственный интеллект. Концепции и терминология искусственного интеллекта» содержится оценка социального риска, связанного с применением систем ИИ. В документе отмечается, что «...система ИИ может быть отнесена к категории высокого риска, если ее действия оказывают прямое воздействие на жизни людей, если она действует без внешнего надзора, а процесс принятия ее решений является непрозрачным». Данное положение подчеркивает необходимость особого внимания к тем системам ИИ, которые используются в сферах, затрагивающих фундаментальные права и свободы, таких как юриспруденция, медицина или образование. Безусловно, ситуация юрислингвистического исследования текста является этически значимой, поскольку результаты экспертизы могут иметь серьезные правовые последствия и оказывать влияние на судьбы людей. Для недопущения высокой степени риска и обеспечения справедливости и объективности необходимо воспринимать советы и рекомендации, выдаваемые ИИ, как вспомогательные материалы, требующие тщательной проверки и критической оценки со стороны эксперта, а также прилагать все усилия для достижения максимально возможной степени прозрачности алгоритмов работы системы. Только в таком случае, при строгом соблюдении этических норм и принципов, возможно эффективно использовать нейронные технологии как один из ряда инструментов, доступных эксперту, расширяя его возможности и повышая качество проводимых исследований. Важно помнить, что этот инструмент может быть использован только в сочетании с другими методами исследования. Окончательное решение и ответственность за выводы, сделанные в ходе экспертизы, всегда должны оставаться за квалифицированным лингвистом-экспертом, обладающим необходимым опытом, знаниями и критическим мышлением.

Несмотря на ряд существенных сложностей, связанных с разработкой, обучением и внедрением нейронных сетей, а также с необходимостью решения этических вопросов, их использование в работе лингвиста-эксперта представляется перспективным и целесообразным. Внедрение нейронных сетей в различные отрасли экономики и социальной сферы является стратегическим приоритетом для Российской Федерации, на государственном уровне обозначена четкая тенденция на активное внедрение технологий искусственного интеллекта в научные исследования, систему здравоохранения, производственные процессы и другие ключевые отрасли. В РФ действует Национальная стратегия развития искусственного интеллекта на период до 2030 года, утвержденная Указом Президента РФ от 10.10.2019 N 490 (ред. от 15.02.2024) «О развитии искусственного интеллекта в Российской Федерации» (далее Стратегия). Целями развития искусственного интеллекта в Российской Федерации являются обеспечение роста благосостояния и качества жизни ее населения, обеспечение национальной безопасности и правопорядка, достижение устойчивой конкурентоспособности российской экономики, в том числе лидирующих позиций в мире в области искусственного интеллекта. Стратегия предполагает «внедрение технологий искусственного интеллекта в государственных корпорациях»; «повышение информированности государственных гражданских служащих и работников федеральных органов государственной власти, иных государственных органов <...> об эффективности технологий искусственного интеллекта и о наличии в Российской Федерации конкурентоспособных организаций-разработчиков таких технологий», также предполагается «популяризация технологий искусственного интеллекта и повышение доверия граждан к ним».

Таким образом, разработка и использование нейросетевых технологий, в том числе в социально-важных сферах, является актуальной и поощряется на государственном уровне. Что касается вопросов юридической и этической регуляции использования ИИ – их решение также ведется в РФ на государственном уровне. Согласно portalu

Искусственный интеллект Российской Федерации, в РФ развивается система комплексного регулирования общественных отношений, возникающих в связи с развитием и использованием технологий ИИ, включающая в себя нормативно-правовое регулирование, нормативно-техническое регулирование и этическое регулирование.

Можно сделать вывод, что интеграция нейросетевых технологий в юрислингвистику и автороведческую экспертизу текста представляется не только технологически возможной, но и обоснованной с научной и юридической точек зрения. В условиях стремительного развития ИИ и государственной поддержки цифровых инициатив использование нейросетей в гуманитарных и правовых исследованиях становится перспективным направлением. Современные достижения в области обработки естественного языка позволяют использовать нейросети как эффективный инструмент, способный автоматизировать значительную часть анализа, повысить его объективность и сократить временные затраты. При этом важно сохранять баланс между использованием данной технологии и соблюдением этических, правовых и профессиональных норм. Использование ИИ влечет за собой определенное количество рисков, которыми важно грамотно управлять. Перспективна разработка профессиональных протоколов или кодекса использования ИИ в юрислингвистической экспертизе, а также специализированных ИИ-инструментов отечественного производства.

Важно помнить, что нейросеть должна рассматриваться не как альтернатива эксперту, а как вспомогательный инструмент, расширяющий его аналитические возможности и повышающий качество экспертизы. Результаты автороведческого анализа должны быть основаны на прозрачных и верифицируемых результатах, оцениваться экспертом. Кроме того, делать окончательные выводы, принимать решение и нести за него ответственность обязан только сам лингвист-эксперт.

Литература

- Аль-Нами Б. А., Кормачев И. Д. Нейросети в современном мире. Исследование возможностей и областей применения chatgpt / Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2023): Сборник научных статей XII Международной научно-технической и научно-методической конференции. Санкт-Петербург, 28 февраля – 01 марта 2023 года. В 4-х томах, Том 2. СПб, 2023. – С. 155-158.
- Белова П. Е., Сафарян А. К. Автоматическое выявление побуждений в тексте: применение методов компьютерной лингвистики в работе эксперта-лингвиста / Национальные и международные тенденции и перспективы развития судебной экспертизы: сборник докладов Научно-практической конференции с международным участием, г. Нижний Новгород, 22–23 мая 2024 г. Нижний Новгород, 2024. – С. 31-38.
- Галюченко А. С. Нейронные технологии и корпусные инструменты для исследования юридического медиадискурса (на примере термина smart-contract) / Язык в эпоху цифровых трансформаций и развития искусственного интеллекта: Сборник научных статей по итогам международной научной конференции. Минск, 2024. – С. 303-308.
- Григорьева Е. Г., Клячин В. А. Исследование статистических характеристик текста на основе графовой модели лингвистического корпуса / Известия Саратовского университета. Новая серия. Серия: Математика. Механика. Информатика. – 2020. – Т. 20, № 1. – С. 116-126.
- Романов А. С. Методика идентификации автора текста на основе аппарата опорных векторов / Доклады Томского государственного университета систем управления и радиоэлектроники. – 2009. – № 2. Т. 1. – С. 36-42.
- Трофимова Г. Н. Алгоритмы и возможности использования генеративных нейросетей в производстве лингвистических экспертиз / Язык в эпоху цифровых трансформаций и развития искусственного интеллекта: Сборник научных статей по итогам международной научной конференции. Минск, 2024. – С. 625-629.
- Ширяева А. А., Новицкая И. В. Преимущества и недостатки использования нейросетей для обработки естественного языка (NLP) / Язык и культура. – 2024. – № 67. – С. 89-101
- Alkhoodairy G., Saleem K. Machine learning algorithm for detecting suspicious email messages using Natural Language Processing / Alexandria Engineering Journal. – 2025. – Vol. 128. – P. 153-165
- Arazzi M., Arikkat R. D., Nicolazzo S. NLP-based techniques for Cyber Threat Intelligence / Computer Science Review. – 2025. – Vol. 58. – 100765.
- Martinc M., Perovšek M., Lavrač N. Textflows: an open science NLP evaluation approach / Lang Resources & Evaluation. – 2024. <https://doi.org/10.1007/s10579-024-09793-1>
- Verma P., O'Shea D., Newe T. WebShield 5.0: Harnessing AI and NLP to combat web threats in Industry 5.0 / Alexandria Engineering Journal. – 2025. – Vol. 127. – P. 677-689/
- Zucco C. Deep Learning Methods in NLP / Encyclopedia of Bioinformatics and Computational Biology (Second Edition). – 2025. – Vol. 2. – P. 190-198.
- ГОСТ Р 71476-2024 «Национальный стандарт российской федерации. Искусственный интеллект. Концепции и терминология искусственного интеллекта». М., 2024.
- ПНСТ 840-2023 «ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ Обзор этических и общественных аспектов». М., 2023.
- Искусственный интеллект Российской Федерации. Регуляторика. URL: <https://ai.gov.ru/ai/regulatory/>

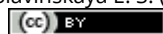
References

- Al-Nami, B. A., Kormachev, I. D. (2023). Neural networks in the modern world. Study of the capabilities and application areas of ChatGPT. In Proceedings of the 12th International Scientific and Technical Conference "Actual Problems of Infotelecommunications in Science and Education" (APINO 2023). Saint Petersburg, 2, 155-158 (in Russian).
- Belova, P. E., Safaryan, A. K. (2024). Automatic detection of imperative utterances in texts. Application of computational linguistics methods in the work of a forensic linguist. In National and International Trends and Prospects for the Development of Forensic Examination. Proceedings of the Scientific and Practical Conference. Nizhny Novgorod, 31-38 (in Russian).
- Galyuchenko, A. S. (2024). Neural technologies and corpus tools for the study of legal media discourse using the example of the term smart contract. In Language in the Era of Digital Transformation and Artificial Intelligence Development. Conference Proceedings. Minsk, 303-308 (in Russian).
- Grigorieva, E. G., Klyachin, V. A. (2020). Study of statistical text characteristics based on a graph model of a linguistic corpus. Proceedings of Saratov University. New Series. Series Mathematics. Mechanics. Informatics, 20(1), 116-126 (in Russian).
- Romanov, A. S. (2009). Methodology for authorship attribution based on support vector machines. Proceedings of Tomsk State University of Control Systems and Radioelectronics, 2(1), 36-42 (in Russian).
- Trofimova, G. N. (2024). Algorithms and possibilities of using generative neural networks in forensic linguistic examination. In Language in the Era of Digital Transformation and Artificial Intelligence Development. Conference Proceedings. Minsk, 625-629 (in Russian).
- Shiryaeva, A. A., Novitskaya, I. V. (2024). Advantages and disadvantages of using neural networks for natural language processing. Language and Culture, 67, 89-101 (in Russian).
- Alkhodhairi, G., Saleem, K. (2025). Machine learning algorithm for detecting suspicious email messages using natural language processing. Alexandria Engineering Journal, 128, 153-165.
- Arazzi, M., Arikkat, R. D., Nicolazzo, S. (2025). NLP-based techniques for cyber threat intelligence. Computer Science Review, 58, 100765.
- Martinc, M., Perovšek, M., Lavrač, N. (2024). Textflows. An open science NLP evaluation approach. Language Resources and Evaluation. <https://doi.org/10.1007/s10579-024-09793-1>
- Verma, P., O'Shea, D., Newe, T. (2025). WebShield 5.0. Harnessing AI and NLP to combat web threats in Industry 5.0. Alexandria Engineering Journal, 127, 677-689.
- Zucco, C. (2025). Deep learning methods in natural language processing. In Encyclopedia of Bioinformatics and Computational Biology (2nd ed., Vol. 2, pp. 190-198).
- GOST R 71476-2024. (2024). Artificial intelligence. Concepts and terminology of artificial intelligence. Moscow (in Russian).
- PNST 840-2023. (2023). Artificial intelligence. Review of ethical and social aspects. Moscow (in Russian).
- Artificial Intelligence of the Russian Federation. (2025). Regulatory framework. Available from <https://ai.gov.ru/ai/regulatory/> (in Russian).

Citation:

Славинская Е. С. Риски применения искусственного интеллекта в лингвистической экспертизе // Юрислингвистика. – 2026 – 39. – С. 72-76.

Slavinskaya E. S. (2026) Prospects and Risks of Using Artificial Intelligence in Authorship Attribution. Legal Linguistics, 39, 72-76.



This work is licensed under a Creative Commons Attribution 4.0. License