

УДК 004.8: 81'1

DOI: 10.26456/vtfilol/2024.1.217

К ВОПРОСУ ОБ ИСПОЛЬЗОВАНИИ СРЕДСТВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЛИНГВИСТИКЕ

И. В. Дорофеева, Л. В. Никифорова, Е. Ю. Мягкова

Тверской государственной университет, г. Тверь

В статье обсуждаются проблемы, связанные с применением искусственного интеллекта в науке о языке. Данная тема является сейчас очень актуальной, так как, идя в ногу со временем, лингвистика всё больше становится объектом внимания ученых-программистов. За последние десятилетия было создано столько языковых платформ искусственного интеллекта для машинного перевода, систем распознавания речи, вопросно-ответных систем и других, сколько не было разработано за все годы прошлого века, начиная с 1940-х годов.

Ключевые слова: *искусственный интеллект (ИИ), машинный перевод, голосовые помощники, системы распознавания речи, вопросно-ответные системы.*

Я представляю себе время, когда мы будем для роботов тем же, чем собаки для людей, и я болею за машины. Клод Шеннон.

Искусственный интеллект (ИИ) прочно вошёл в нашу жизнь, и сейчас трудно представить себе, как можно обходиться без стиральных машин, компьютеров, смартфонов и других электронных приборов, сделавших нашу жизнь комфортнее.

В научной литературе существует много определений ИИ. Но, поскольку в данной статье не ставится задача анализа разных подходов к определению ИИ, то здесь мы будем считать, что ИИ – это способность компьютера выполнять действия, которые обычно свойственны человеку. Это целая совокупность компьютерных наук, занимающихся созданием и разработкой систем и программ, наделённых интеллектом, сопоставимым с человеческим интеллектом.

Сейчас люди ежедневно сталкиваются с ИИ, средства которого играют всё более значительную роль в нашей жизни. Мы пользуемся, например, поисковыми системами *Google* и *Yandex*, системами распознавания лиц и речи, голосовыми помощниками, типа Алисы и Маруси и тому подобное. Всё это свидетельствует о том, насколько тесно мы связаны ИИ. Технологии искусственного интеллекта применяются везде – от бытовой техники до смартфонов и компьютеров.

© Дорофеева И. В., Никифорова Л. В., Мягкова Е. Ю., 2024

Но следует отметить, что искусственный интеллект – далеко не новая сфера исследований. Считается, что даже в древности люди задумывались о том, как автоматизировать свою деятельность, чтобы облегчить жизнь. И всё-таки только в 40–50-е годы прошлого века началась история ИИ. В своём развитии искусственный интеллект прошёл этапы становления, забвения и бурного развития.

Началась история ИИ в середине прошлого столетия, когда еще в 1950 году Алан Тьюринг первым предложил идею мыслящей машины. После исследований Алана Тьюринга, Джона фон Неймана и Норберта Винера появились первые вычислительные машины, и исследователи стали задумываться о том, сможет ли машина имитировать человеческую деятельность, в том числе и мыслительную.

В 1956 году в Дартмутском колледже состоялась конференция о «механизации интеллекта», на которой Джон Маккарти, когнитивист и специалист по информатике, предложил термин «искусственный интеллект». Этот момент считается началом истории ИИ.

В 1960–1970-е годы бурно развивается компьютерная индустрия. Компьютеры становятся мощнее, дешевле, быстрее и начинают хранить гораздо больше информации.

Сначала системы искусственного интеллекта появились в медицине, финансах, робототехнике, игровой индустрии, промышленности.

В отличие от этих областей лингвистика всегда стояла особняком по причине крайней сложности автоматизации языковых процессов. В языке нет строгих правил и жестких ограничений, существующих в естественных науках. Язык всегда был и остаётся неким «черным ящиком». Мы знаем только то, что видим и слышим на входе. Все процессы, связанные с переработкой информации, остаются внутри «черного ящика». А тем более сложно структурировать и понять, что происходит в головном мозге человека, когда он читает, например, текст.

Поэтому проникновение искусственного интеллекта в лингвистику происходило очень сложно – через взлёты и падения, от эйфории до забвения. В этой связи можно вспомнить, какими неуклюжими и топорными были первые системы машинного перевода. Английское предложение *I am fine*, например, компьютер переводил на русский язык предложением *Я – штраф*. Конечно, ведь слово *fine* означает и *штраф*, и *прекрасно*. Это говорит о том, что компьютерная система не имеет экстралингвистических знаний.

Но в последнее десятилетие развитие ИИ в лингвистике движется семимильными шагами. Достигнуты значительные успехи и в машинном переводе. Компьютеры сейчас даже пишут тексты и стихи.

Причины стремительного развития ИИ в XXI веке объясняются следующими причинами:

1. Появилась огромное количество данных из социальных сетей и других медиа, с помощью которых системы ИИ могут полноценно обучаться.

2. Созданы мощные компьютеры, позволяющие обрабатывать и анализировать колоссальные объёмы информации с большей скоростью и эффективностью.

3. Разработаны новые технологии и подходы, которые поддерживают развитие искусственного интеллекта. Машинное обучение, нейронные сети, глубокое обучение стали доступными, и дали новые возможности для создания более умных и адаптивных систем [2].

В последние десять лет было создано больше различных систем ИИ, чем за всю его историю. Например:

– В 2011 году появилась *Watson (IBM)* – система вопросов и ответов на естественном языке. Она выиграла викторину Jeopardy, победив двух бывших чемпионов. В том же году Юджин Густман – говорящий компьютерный чат-бот – обманул судей во время теста Тьюринга, заставив их принять его за человека.

– В 2011 году *Apple* выпустила *Siri*, виртуального помощника, который с помощью технологии *NLP* (обработки естественного языка) делает выводы, изучает, отвечает и предлагает что-либо своему пользователю-человеку.

– В 2016 году появилась *София* – первый робот, который может менять выражение лица, видеть (с помощью распознавания изображений) и разговаривать с помощью искусственного интеллекта.

– В 2017 году *Facebook* разработал двух чат-ботов для переговоров друг с другом. В процессе коммуникации они обучались и совершенствовались тактики общения. В итоге эти чат-боты изобрели свой собственный язык, на котором разговаривали.

– 2023 – год прогресса для генеративных сетей (*GAN*), которые создают реалистичные изображения и видео, и больших языковых моделей (*LLM*), например *ChatGPT* [Там же].

Искусственный интеллект, наделённый способностью обрабатывать и анализировать большие объёмы данных, не только развил свою способность в смысле общения с человеком, он также внёс свой вклад в улучшение нашего понимания языка и коммуникации. Используя алгоритмы машинного обучения, ИИ может извлекать уроки из моделей использования человеческого языка и имитировать их, что приводит к разработке более сложных и более похожих на естественное взаимодействие между людьми и машинами.

Профессор компьютерных наук Стэнфордского университета Себастьян Траун считает, что «Никто не формулирует это таким образом, но я думаю, что искусственный интеллект – это почти гуманитарная дис-

циплина. На самом деле это попытка понять человеческий интеллект и человеческое познание» [6].

Однако, несмотря на значительный прогресс, современные системы ИИ всё еще не совершенны и главная проблема, с которой сталкиваются исследователи – это языковое обучение компьютера. Специалист по искусственному интеллекту Ноа Гудман объясняет: «Особенность языка заключается в том, что с одной стороны он зависит от большого количества знаний о нем, но с другой стороны, он также определяется и огромным количеством общеизвестной информации об окружающем мире, и оба эти знания очень тонко переплетаются друг с другом» [Там же].

Взаимодействие между искусственным интеллектом и лингвистикой является на самом деле более заметным и практичным, чем думают некоторые исследователи, поскольку лингвистика занимает центральное место в повседневных инструментах нашей жизни, таких как распознавание речи и поисковые системы. Здесь ключевую роль играет компьютерная лингвистика. Эта область приобретает все большее значение в языковых процессах, связанных с медициной, где механизмы компьютерной лингвистики могут использоваться для извлечения информации из файлов пациентов с помощью ключевых слов. Более того, методы компьютерной лингвистики были адаптированы для использования маркетинговыми компаниями для получения обратной связи от потребителей и лучшего понимания их потребностей. ИИ в лингвистике является областью, которая объединяет исследования в области компьютерной лингвистики и машинного обучения [3].

Основная цель ИИ в лингвистике – создание компьютерных систем, способных обрабатывать и понимать естественный язык так же, как это делают люди. Но ИИ используется во многих других разделах лингвистики. Одной из ключевых областей, где ИИ оказывает значительную помощь, является изучение древних языков. Алгоритмы ИИ используются для расшифровки древних письменностей, которые долго являлись загадкой для ученых. Например, ИИ применили при расшифровке загадочного средневекового текста, рукописного труда Войнич, (Рукопись Войнич или Манускрипт Войнич, также известный как *Beinecke MS 408* и *VMS*), – иллюстрированный кодекс, написанный неизвестным автором на неизвестном языке и неизвестном алфавите). Анализируя структуры и модели в тексте, ИИ смог предоставить ценное понимание возможных значений манускрипта, что тысячелетиями не удавалось ученым.

Кроме того ИИ также используется для анализа литературных произведений, открывая новые перспективы для изучения авторства, стиля и тематических моделей. Обработывая большие объемы текста, ИИ может выявлять тонкости и тенденции стиля автора, которые могут быть

упущены при человеческом анализе. Это, очевидно, приведёт к новым интерпретациям и пониманию литературных произведений, что будет обогащать наше восприятие литературы.

Инструменты искусственного интеллекта для *NLP* (обработка естественного языка) играют важную роль в языковом переводе, анализе и синтезе текста, обнаружении мошенничества, поиске нужной и достоверной информации, а также в процессе ответов на определенные вопросы.

Среди областей применения ИИ лингвистике можно также назвать разбор и анализ текста; автоматическое распознавание речи, транскрибирование записей; упомянутые выше машинный перевод и создание автоматических ответов в чат-ботах; анализ тональности и sentiment-анализ: это задача определения эмоциональной окраски текста [4].

Следует назвать основные методы применения ИИ в лингвистике:

1. Статистические модели: методы, основанные на анализе больших объемов текстовых данных и статистическом моделировании языка. Статистические модели могут использоваться для различных задач, таких как разбор предложений, машинный перевод и анализ тональности.

2. Нейронные сети: методы, основанные на моделировании работы мозга и использовании искусственных нейронных сетей. Нейронные сети могут применяться для решения задач обработки естественного языка, распознавания речи и генерации текста.

3. Глубокое обучение: подход к машинному обучению, основанный на использовании глубоких нейронных сетей с большим количеством слоев.

4. Комбинированные методы: методы, которые комбинируют различные подходы и модели для решения задач в области ИИ в лингвистике. Комбинированные методы могут быть эффективными для достижения более точных результатов и решения сложных задач [1].

Однако внедрение ИИ в филологию не обходится без проблем. Одной из основных опасностей является риск чрезмерного доверия ИИ, так как это может свести к минимуму человеческий фактор, являющийся главным в гуманитарных науках. В то время как ИИ может предоставлять ценную помощь, он не может заменить человеческую способность к критическому мышлению, креативности и эмоциональному пониманию. Поэтому нужно и важно найти баланс между использованием ИИ в качестве инструмента и сохранением сущности филологии.

Кроме того, существуют и этические соображения, особенно в отношении конфиденциальности данных и прав интеллектуальной собственности. Поскольку алгоритмы ИИ анализируют и обрабатывают вну-

шительные объемы текстовых данных, важно гарантировать ответственное и этичное использование этих данных.

Многие исследователи искусственного интеллекта даже опасаются чрезмерного внедрения ИИ в нашу жизнь.

«Я все больше склоняюсь к мысли, что должен быть какой-то регуляторный надзор, возможно, на национальном и международном уровнях, просто чтобы убедиться, что мы не совершаем каких-то очень глупых поступков. Я имею в виду, что с искусственным интеллектом мы вызываем демона», – предупредил Илон Маск на симпозиуме, посвященном столетию *АэроАстро* в Массачусетском технологическом институте [6].

Но существуют исследователи, которые тем не менее прогнозируют всё большее развитие ИИ. Известный футуролог и технический директор компании Google Рэй Курцвейл подсчитал, что «Искусственный интеллект достигнет человеческого уровня примерно к 2029 году. Далее, скажем, к 2045 году, мы увеличим интеллект, человеческий биологический машинный интеллект нашей цивилизации в миллиард раз» [4].

Несмотря на все сложности, потенциальные выгоды от внедрения ИИ в лингвистику огромны. Объединяя аналитическую мощь ИИ с интерпретационным и креативным пониманием ученых-людей, можно углубить и расширить наше понимание языков и литературы способами, которые раньше трудно было себе представить. Это слияние технологии и гуманитарных наук не только обогащает филологию, но также открывает новые возможности для междисциплинарного исследования и сотрудничества.

В заключение следует отметить, что, внедрение ИИ в лингвистику является свидетельством развития и мощи компьютерных технологий. Наступает новая эра изучения языков, литературы, понимания глубинных языковых процессов. За инструментами ИИ будущее, поскольку они предлагают точные и непредвзятые решения, открывают огромные перспективы для развития лингвистики. Илон Маск поддерживает эту идею. Он заявляет, что отношения «человек-машина» следует вывести на более высокий уровень: установить прямую связь «мозг-компьютер». Он основал компанию *Neuralink*, которая планирует подключать людей к системе, с помощью которой они смогут обмениваться мыслями, не тратя своё время и энергию на словесность [5].

Список литературы

1. Воронкова М. Д. Обзор зарубежных и отечественных компаний, использующих искусственный интеллект // Молодой ученый. 2023. № 30 (477). С. 42–49.

2. Денега А. О. История развития искусственного интеллекта [Электронный ресурс] // Справочник от Автор24. URL https://spravochnik.ru/informacionnye_tehnologii/istoriya_razvitiya_iskusstvennogo_intellekta/ (дата обращения: 15.03.2024).
3. Кузьмина А. А., Лифшиц М. А., Костенко В. Ю. Методы компьютерной лингвистики и обработки естественного языка: возможности ограничения для задач психологии личности // Современная зарубежная психология 2022. Т. 11. № 1. С. 104–115.
4. Ладов В. А. Язык в системе искусственного интеллекта: синтаксис и семантика // Гуманитарная информатика. 2004 № 1. С. 72–80.
5. Люшнина Д. Г. Искусственный интеллект: ответ или вызов? // Бюллетень медицинских интернет-конференций. 2014. Т. 4. Вып. 11. С. 1142.
6. Bernar Marr 28 best quotes fbout artificial intelligence [Электронный ресурс] // Huffpost. URL: https://www.huffpost.com/entry/28-best-quotes-about-artificial-intelligence_b_59afd546e4b0d0c16bb528d9 (Дата обращения: 13.03. 2024).

ON THE USE OF ARTIFICIAL INTELLIGENCE IN LINGUISTICS

I. V. Dorofeeva, L. V. Nikiforova, E. Yu. Myagkova

Tver State University, Tver

The article discusses the problems associated with the use of artificial intelligence in the science of language. This topic is very relevant now, because, keeping up with the times, linguistics is increasingly becoming the object of attention of computer scientists. In recent decades, there have been created as many artificial intelligence language platforms for machine translation, speech recognition systems, and question-and-answer systems as there have not been in all the years of the last century, since the 1940s.

Keywords: *artificial intelligence (AI), machine translation, voice assistants, speech recognition systems, question-and-answer systems.*

Об авторах:

ДОРОФЕЕВА Ирина Валентиновна – кандидат филологических наук, доцент кафедры иностранных языков гуманитарных факультетов Тверского государственного университета (170100, г. Тверь, ул. Желябова, 33), e-mail: Dorofeeva.IV@tversu.ru.

НИКИФОРОВА Любовь Васильевна – доцент кафедры фундаментальной и прикладной лингвистики Тверского государственного университета (170100, г. Тверь, ул. Желябова, 33), e-mail: Nikiforova.LV@tversu.ru.

МЯГКОВА Елена Юрьевна – доктор филологических наук, профессор кафедры теории языка, перевода и французской филологии, Тверской государственной университет (170100, Тверь, ул. Желябова, 33), e-mail: Myagkova.EY@tversu.ru.

About the authors:

DOROFEEVA Irina Valentinovna – Candidate of Philology, Associate Professor at the Foreign Languages Department of the Humanity Faculties, Tver State University, (170100, Tver, Zhelyabov str., 33), e-mail: iDorofeeva.IV@tversu.ru.

NIKIFOROVA Lyubov Vasilevna – Candidate of Philology, Associate Professor at the Department of Fundamental and Applied Linguistics, Tver State University (170100, Tver, Zhelyabov str., 33), e-mail: INikiforova.LV@tversu.ru.

MYAGKOVA Elena Yuryevna – Doctor of Philology, Professor at the Department of Theory of Language, Translation and French Philology, Tver State university, (170100, Tver, Zhelyabova str., 33), e-mail: Myagkova.EY@tversu.ru.

Дата поступления рукописи в редакцию: 17.02.2024 г.

Дата подписания в печать: 01.03.2024 г.