

МАНИПУЛЯЦИЯ ЧЕЛОВЕЧЕСКИМ СОЗНАНИЕМ ПОСРЕДСТВОМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА КАК ГУМАНИТАРНАЯ ПРОБЛЕМА XXI ВЕКА

Зиньковская Анастасия Владимировна

- e-mail: anastassiat@bk.ru
доктор филологических наук, профессор, заведующая кафедрой
английской филологии факультета романо-германской филологии,
Кубанский государственный университет
350040, г. Краснодар, ул. Ставропольская, д. 149,
Российская Федерация

Оломская Наталья Николаевна

- e-mail: olomnat@mail.ru
доктор филологических наук, доцент, профессор кафедры
английской филологии факультета романо-германской филологии,
Кубанский государственный университет
350040, г. Краснодар, ул. Ставропольская, д. 149,
Российская Федерация

Для цитирования

Зиньковская А. В., Оломская Н. Н.

Манипуляция человеческим сознанием посредством искусственного интеллекта как гуманитарная проблема XXI века // Российский социально-гуманитарный журнал. 2024. № 1. URL: www.evestnik-mgou.ru

- Статья поступила в редакцию: 16.11.2023
- Статья размещена на сайте: 26.02.2024

СТРУКТУРА СТАТЬИ

[Аннотация](#)

[Ключевые слова](#)

[Введение](#)

[Специфика манипуляции в современном мире](#)

[Эксперименты в области усовершенствования возможностей искусственного интеллекта](#)

[Специфика манипуляции искусственного интеллекта](#)

[Искусственный интеллект как гуманитарная проблема](#)

[Заключение](#)

[Литература](#)

[Article information](#)

АННОТАЦИЯ

Цель нашего исследования состоит в том, чтобы проанализировать возможности манипуляции сознанием человека посредством использования искусственного интеллекта.

Процедура и методы. В процессе исследования авторами использовался междисциплинарный подход при рассмотрении специфики манипулятивного воздействия искусственным интеллектом на социум в целом и индивида в частности. Данный подход позволил применить ряд методов и приёмов для анализа текстов, сгенерированных с применением нейросетей, – таких как контекстуальный анализ, компонентный и сопоставительный анализ.

Результаты. Следствием лингвокогнитивного манипулирования через социальные сети, а также разнообразные платформы, существующие в виртуальном пространстве, является мощное социально-политическое воздействие на социум, поскольку социальные сети – это не только инструмент обмена информацией среди пользователей, так как именно социальные сети предоставляют большой выбор средств и техник манипулирования, а также – неограниченные возможности для применения искусственного интеллекта для манипулирования общественным сознанием.

Возможности нейросетей практически не лимитированы, и, несмотря на то, что результаты работы искусственного интеллекта в совокупности идентичны, они чаще всего идеологически окрашены и не предлагают выбора ни лексики, ни синтаксиса. Использование подобных нейросетей приводит к снижению творческого потенциала индивида, мотивации к изучению и осмыслению явлений окружающей действительности, что в конечном итоге приведёт к манипуляции человеческим сознанием искусственным интеллектом.

Теоретическая и / или практическая значимость. Результаты научного исследования вносят вклад в дальнейшее осмысление опасности манипуляции сознанием искусственным интеллектом как гуманитарной проблемы. Анализ текстов, сгенерированных с помощью искусственного интеллекта, а также анализ языковых средств поможет в перспективе учитывать возможности искусственного интеллекта и выработать методики, позволяющие ограничить сферу применения искусственного интеллекта.

КЛЮЧЕВЫЕ СЛОВА:

искусственный интеллект (ИИ), гуманитарная проблема, манипуляция, манипулятивное воздействие, нейросети

ВВЕДЕНИЕ

В эпоху тектонических изменений, происходящих в первой трети XXI в., всё больший интерес исследователей вызывает изучение возможностей искусственного интеллекта при манипулировании сознанием индивида в частности и общественным сознанием социума в целом, что особенно важно сегодня, когда происходит открытая информационная война между так называемым коллективным Западом и Российской Федерацией.

Цель нашего исследования заключается в том, чтобы проанализировать возможности манипуляции сознанием человека посредством использования искусственного интеллекта. Гипотеза исследования состоит в том, что стремительная эволюция возможностей искусственного интеллекта (ИИ) может стать прямой экзистенциальной угрозой человечеству и приведёт к тотальной трансформации жизни и деятельности социума в результате использования ИИ. Стремительное распространение использования искусственного интеллекта в различных социальных областях становится угрозой прерогативе человека, а не просто проблемой, так как начинает лишать человека возможности развиваться и стремится подчинить человека себе. Данной проблеме посвящён ряд работ отечественных и зарубежных исследователей: Т. В. Черниговской, В. А. Лекторского, Д. И. Дубовского, Дж. Фодора, Т. Ван Дейка и др. Отметим ряд наиболее, с нашей точки зрения, интересных работ по данной проблеме: «Искусственный интеллект для науки и наука для искусственного интеллекта» в журнале «Вопросы философии» (К. В. Анохин, К. С. Новоселов, С. К. Смирнов, А. Р. Ефимов, Ф. М. Матвеев, 2022), работу И. М. Дзялошинского «Искусственный интеллект: гуманитарная перспектива» в журнале «Вестник НГУ» (2022), сборник «Человек и системы искусственного интеллекта» (2022). Тем не менее проблемы, связанные с манипуляцией сознанием человека искусственным интеллектом, всё ещё недостаточно изучены и требуют пристального внимания коллег. Данная работа актуальна ещё и тем, что она привлекает внимание к необходимости тщательного изучения данного вопроса. Теоретико-методологической основой представленного в работе анализа является понимание сложного механизма манипулятивного воздействия на человека с помощью нейросетей, как справедливо отметила Т. В. Черниговская, ссылаясь на Р. Джекендоффа, который предложил «перекинуть мост между вычисляющим и самодостаточным

мозгом и внешним миром, вводя концепт f-mind, который можно понимать как способность средствами естественного языка кодировать определённые комбинации в нейронных сетях в релевантных контексту отделах мозга» [8, с. 32]. Возникает вопрос: можно ли использовать эти средства для кодировки определённых комбинаций в ИИ, и к чему это приведёт, если использовать возможности «Strong» ИИ? Характер исследования и его основная цель (обоснование гипотезы исследования) объясняют выбор междисциплинарного подхода при рассмотрении специфики манипулятивного воздействия искусственным интеллектом на социум в целом и индивида в частности. Данный подход позволил применить ряд методов и приёмов для анализа текстов, сгенерированных с применением нейросетей, – таких как контекстуальный анализ, компонентный и сопоставительный анализ.

СПЕЦИФИКА МАНИПУЛЯЦИИ В СОВРЕМЕННОМ МИРЕ

Манипуляция общественным сознанием далеко не нова, если вспомнить историю (например, военные действия стран-членов НАТО в Ираке, Ливии, Югославии и сегодня на Украине), – невозможно представить её без манипуляции общественным мнением, часто не только направленным на «врага» для психологического воздействия, но и обращённым к своему народу. Так, к примеру, во время бомбардировок Ирака и в последствии Ливии американскому народу в СМИ насаждалось утверждение о том, что силы НАТО несут демократию в эти страны, распространяют идеи равенства и свободы. Социально-экономические изменения, влияющие на эволюционные процессы в современном информационном обществе, имеют опасные тенденции для будущего человечества, поскольку сопровождаются формированием, а затем манипулированием сознанием индивидуума в частности и социума в целом посредством управления не только поведением человека, но и воздействием на его когнитивные способности. Именно в текущем столетии возникает опасность манипуляции общественным сознанием не только традиционными средствами, но и набирающим мощь искусственным интеллектом. В данном контексте представляется актуальным изучение бинарных антагонистов «чужой – свой», так как данная оппозиция может рассматриваться в применении практически во всех политико-

идеологических и когнитивных аспектах исследования. В этой связи интересными представляются результаты исследования интердискурсивной среды медийного и сетевого дискурса Украины после государственного переворота 2014 г., так как именно на фоне этих событий произошло резкое отчуждение Российской Федерации и Украины, направленной на разобщение двух исторически связанных народов, что в конечном итоге отразилось в формировании резко негативного образа России как «врага» и оказало влияние на проявление русофобии в ряде стран Европы, а с началом Специальной военной операции эти тенденции только усилились. А. А. Хабаров в своём исследовании приводит интересные результаты мониторинга общественного мнения по данным социологических опросов молодёжи на Украине в 2017 г., где в фокус-группе из 2000 респондентов (от 14 до 29 лет) 61% опрошенных полагал, что Российская Федерация несёт ответственность за разжигание конфликта на востоке Украины, 60% одобряли вступление Украины в ЕС, 91% опрошенных при этом пользовались социальными сетями [7, с. 72–82].

Несомненно, такие результаты стали возможными при перманентном использовании украинскими медиа манипулятивных приёмов и техник, направленных на когнитивно-психологическую составляющую социума Украины и, как следствие, на формирование общественного мнения, выгодного правящей элите Украины. Мы полагаем, что такой результат является следствием лингвокогнитивного манипулирования через социальные сети, а также разнообразные платформы, существующие в виртуальном пространстве, поскольку социальные сети – это не только инструмент обмена информацией среди пользователей, но и мощное социально-политическое средство воздействия на социум. Мы полагаем, что именно социальные сети предоставляют большой выбор средств и техник манипулирования, а также – неограниченные возможности для применения искусственного интеллекта для манипулирования общественным сознанием.

ЭКСПЕРИМЕНТЫ В ОБЛАСТИ УСОВЕРШЕНСТВОВАНИЯ ВОЗМОЖНОСТЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Для реализации цели нашего исследования необходимо определить понятие «гуманитарная проблема» – это сложный социальный, экономический и политический вопрос, связанный с противоречиями и недостатками в области человеческого развития, нарушением прав и основных потребностей людей. Они касаются таких аспектов, как доступ к образованию, здравоохранению, безопасности, равноправию, окружающей среде и многому другому. Обращаясь к словарной статье в «Толковом словаре по искусственному интеллекту», читаем: «Искусственный интеллект – свойство искусственных интеллектуальных систем выполнять творческие функции, которые традиционно считаются прерогативой человека»¹.

В современных исследованиях (Т. В. Черниговской, Д. И. Дубровского, В. А. Лекторского, Ф. Джонсона-Лэрда и др.) выделяют две основные формы ИИ (искусственного интеллекта). Во-первых, «weak» («слабый») ИИ, который представляет собой некое подобие естественного интеллекта (ЕИ), обладающего ограниченным спектром возможностей (например, выполняющего «рутинную» трудоёмкую работу, такую как обработка базы данных, машинный перевод с одного языка на другой и пр.). Во-вторых, «strong» («сильный») ИИ, который обладает ментальной репрезентацией, способностью к самообучению, функционирует независимо от человека, а также способен самостоятельно ставить цели и решать задачи. Разработка интегральной системы ИИ приобретает всё большую актуальность и является предметом исследований научных центров в сфере ИИ. Перспективы таких разработок, на наш взгляд, и представляют основную гуманитарную угрозу для человечества в XXI в., поскольку достижения науки, к сожалению, не всегда используются в «мирных целях» и могут служить инструментом манипуляции общественным сознанием. А. Чичулин определяет нейросеть как модель, которая «производит и анализирует комплексные паттерны баз данных, учится на примерах, предвидит события и принимает решения» [9, р. 9].

Эксперименты в области усовершенствования ИИ (в международной классификации принято сокращённо обозна-

¹Толковый словарь по искусственному интеллекту / сост. А. Н. Аверкин, М. Г. Гаазе-Рапопорт, Д. А. Поспелов. М.: Радио и связь, 1992. С. 38.

чать как AGI) ведутся в разных направлениях и часто бывают достаточно противоречивыми. Так, к примеру, специалисты Медиа-лаборатории МТИ (MIT Media Lab) создали первую в мире психопатическую нейросеть, которую назвали «Норман» в честь известного персонажа романа «Психо» Р. Блоха². Смысл эксперимента сводился к следующему. Группа учёных ограничила информационный поток Нормана паттернами, содержащими информацию о смерти, разрушениях и катастрофах, что способствовало формированию маниакально-психопатической индивидуальности ИИ. Результатом этого явилось то, что ИИ идентифицировал в нейтральных изображениях сцены смерти и насилия. Цель эксперимента заключалась в том, чтобы продемонстрировать негативное влияние интернета на психику человека. На наш взгляд, этот эксперимент показал возможности манипулятивного воздействия на психику человека с помощью ИИ. Изначально нейросеть представляет собой некую «tabula rasa», поскольку у неё нет желаний, эмоций, побуждений и пр., т. е. того, что свойственно человеку. Однако эксперимент с Норманом подтвердил опасения учёных, разрабатывающих ИИ, в том, что ИИ («strong») способен обучаться и, как следствие, приобретать опасные «знания» и использовать их в конечном счёте против человечества, поскольку ИИ не обладает нравственно-этическими нормами, регулирующими поведение индивидуума в социуме. Таким образом, ИИ и нейросети изначально являются продуктом интеллектуальной деятельности человека (что вложили, то и получили), но это именно тот случай, когда «ученик» легко может превзойти «учителя». Вопросы, связанные с дальнейшей разработкой и использованием ИИ с гуманитарной точки зрения остаются открытыми, поскольку, как отмечает Д. И. Дубровский: «Здесь необходимыми ресурсами для построения моделей AGI могут служить, в первую очередь, положения эпистемологии, касающиеся объяснения и обоснования успешного познавательного процесса, результаты психологических и феноменологических исследований ценностно-смысловых, динамических структур субъективной реальности и результаты новейших нейронаучных исследований сознания» [2, с. 100–107]. Как справедливо заметил Дж. Фодор: «Что с нами не так, так это то, что наш

² Норман: первый в мире искусственный интеллект-психопат // TechInsider: [сайт]. [06.06.2028]. URL: <https://www.techinsider.ru/technologies/news-426782-norman-pervyy-v-mire-iskusstvennyy-intellekt-psihopat> (дата обращения: 12.09.2023).

разум не был развит для того, чтобы справляться с тем миром, в котором мы живём»³. Говоря об использовании ИИ и нейросетей в социальных сетях для генерирования общественного мнения о тех или иных событиях (мы уже приводили выше социальный опрос, поскольку полагаем, что данное общественное мнение было сформировано в том числе и под влиянием ИИ и нейросетей, широко применяемых в интернете), следует отметить, что манипуляция общественным сознанием ИИ находит всё большее распространение. Как справедливо отметил И. Маск: «В долгосрочной перспективе искусственный интеллект может стать автономным и захватить власть над миром, но в краткосрочной перспективе он используется политиками для контроля над тем, что вы думаете, для прекращения вашего независимого суждения и уничтожения демократии накануне президентских выборов»⁴.

СПЕЦИФИКА МАНИПУЛЯЦИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Исследование проблемы манипуляции искусственного интеллекта, безусловно, носит междисциплинарный характер, где когнитивная составляющая является одной из основных опор анализа. Проблемы соотношения языка и сознания, концептуализация и категоризация социума, а также способность человека связывать язык и формы его воздействия на общество являются неотъемлемым механизмом в процессе манипуляции сознанием человека. Примером могут послужить созданные разными программами искусственного интеллекта тексты на актуальные проблемы. Анализируя данные тексты, несложно понять, что информация предвзята и искажает исторические факты, но именно так эта проблема освещается на страницах англоязычного интернет-пространства, где новости и факты очень часто генерируются искусственным интеллектом по ключевым словам, адаптируя предубеждения и выдумки разработчиков. Искусственный интеллект не предполагает проверки того, что он создаёт. Дезинформация, таким образом,

³ Fodor J. Why pigs don't have wings [Электронный ресурс]. URL: http://www.biolinguagem.com/ling_cog_cult/fodor_2007_why_pigs_dont_have_wings.pdf (дата обращения: 11.09.2023).

⁴ Интервью Илона Маска Такеру Карлсону [Электронный ресурс]. URL: <https://yandex.ru/video/preview/5309831523159438930> (дата обращения: 21.10.2023).

распространяется быстро и благодаря используемым определённым языковым моделям влияет на сознание людей, которые теряют способность отличать ложь от истины, так как мозг современного человека склонен к лени и открыт для манипуляции. Особенно это касается молодых людей, сознание которых подвержено возможному «переформатированию».

В качестве эксперимента нами были выбраны наиболее «популярные» нейросети: TWEE, Chat GPT, Neuro-texter, Educaplay. Задача была представлена следующим образом: нейросетям предлагалось по ключевым словам «Russia», «Ukraine», «conflict» сгенерировать текст. В результате мы получили четыре практически идентичных текста. Приведём данные тексты в сокращённом варианте:

Chat GPT:

The conflict between Ukraine and Russia has been a significant geopolitical issue since 2014. Stemming from historical, cultural, and political factors, this conflict has resulted in significant consequences for both nations and the wider international community. The roots of the conflict can be traced back to Ukraine's history as part of the Soviet Union. Ukraine gained independence but struggled to establish a stable political identity. The country became divided between those who sought closer ties with Europe. In 2013, Ukraine's decision to pursue closer ties with the European Union sparked widespread protests known as the Euromaidan movement. Russia saw this as a threat to its influence in the region and subsequently annexed Crimea in 2014, a move that was widely condemned by the international community. Following the annexation of Crimea, pro-Russian separatists in eastern Ukraine, primarily in the Donbass region, declared independence. This led to a violent conflict between Ukrainian armed forces and separatist groups supported by Russia. The conflict between Ukraine and Russia has drawn significant international attention. Additionally, the conflict has highlighted the vulnerability of post-Soviet countries to external influences, raising concerns among neighboring nations about their own territorial integrity.

TWEE

The conflict between Russia and Ukraine has been ongoing for several years now. It began in 2014 when Russia *annexed Crimea* from Ukraine. On one hand, there is a deep *historical and cultural connection* between Russia and Ukraine, which dates back centuries. On the other hand, Ukraine has long struggled to establish *its own independent identity*. *Russian-backed separatists* have taken control of parts of eastern Ukraine, *leading to a protracted and bloody conflict that has resulted in thousands of deaths and the displacement of millions of people*. The international community has imposed *economic sanctions* on Russia in an attempt to pressure them into *withdrawing their support* for the separatists. In recent years, there have been efforts to *negotiate a ceasefire*.

Educaplay

The conflict between Ukraine and Russia began in 2014 when *Russia annexed Crimea* from Ukraine. The international community has *condemned Russia's actions and imposed sanctions on the country*. Efforts to find a peaceful resolution to the conflict have been unsuccessful so far.

Neuro-texter

The conflict between Russia and Ukraine has been *a key issue in international relations since its outbreak* in 2014. This ongoing conflict has had *far-reaching consequences*. The roots of this conflict can be traced back to *historical, political, and cultural factors*, as well as geopolitical interests. Ukraine's desire to align itself with Western democracies *and pursue closer ties with the European Union* ignited tensions with Russia. The *annexation of Crimea* by Russia in 2014 further escalated tensions and led to armed conflict in eastern Ukraine between Ukrainian government forces and *pro-Russian separatists*. The conflict has resulted in *significant human suffering, including civilian casualties, displacement of populations, economic hardships, and violations of human rights*. It has also strained diplomatic relations between Russia and Western countries who *have imposed sanctions on Russia* in response to its actions in Ukraine.

Как показывает анализ данных текстов, частотность использования синонимичных идеологически маркированных единиц дискурса достаточно высока (the annexation of Crimea, imposed sanctions on Russia, significant human suffering, be traced back to historical, political, and cultural factors; negotiate a ceasefire,

separatists, independent identity closer ties и пр.), что говорит об однородности и практически полном совпадении лексических единиц, которые представляют собой интертекстуальное прецедентное использование в нейросетях. Основными инструментами генерирования текста, применяемыми разработчиками данных нейросетей, служат информативные знаки, репрезентирующие социально-ценностную парадигму, основанную на дискурсивно-мотивированной деятельности. Характерными особенностями анализируемых текстов являются наличие политических клише и штампов, а также тождественных оценочных коннотаций: наиболее частотны по употреблению слова «*significant*», «*conflict*», «*consequences*», «*annexation*» и его производные. Кроме того, анализ показывает употребление определённых лексических единиц, которые служат маркерами, идеологически моделирующими у реципиента заведомо враждебное отношение к данным событиям. Например:

The conflict has resulted in significant human suffering, including civilian casualties, displacement of populations, economic hardships, and violations of human rights.

The international community has condemned Russia's actions and imposed sanctions on the country.

In 2013, Ukraine's decision to pursue closer ties with the European Union sparked widespread protests known as the Euromaidan movement. These protests led to the ousting of then-President Viktor Yanukovich, who had favored closer ties with Russia. Russia saw this as a threat to its influence in the region and subsequently annexed Crimea in 2014, a move that was widely condemned by the international community. The conflict between Ukraine and Russia has been a significant geopolitical issue since 2014.

The annexation of Crimea by Russia in 2014 further escalated tensions and led to armed conflict in eastern Ukraine between Ukrainian government forces and pro-Russian separatists.

Так, Т. ван Дейк в исследованиях идеологии как формы социального познания писал о том, что «интерпретация идеологии как когнитивной системы позволяет использовать её в процессах формирования социальных репрезентаций различных групп общества и за счёт этого осуществлять косвенный контроль над относящимися к ним социальными практиками и, в свою очередь, явлениями речевой деятельности индивидов» [10, р. 242–257], чем в полной мере пользуются и будут пользоваться в дальнейшем разработчики ИИ для того, чтобы

оказывать влияние на формирование общественного мнения социума в целом и индивида в частности.

Таким образом, говоря о возможностях нейросетей сегодня, отметим, что, несмотря на оперативное создание текста (не более 2 минут, а зачастую 5–10 секунд), результаты работы искусственного интеллекта в совокупности идентичны, могут быть идеологически окрашены и не предлагают выбора как лексики, так и синтаксиса. Иными словами, использование подобных нейросетей приводит к тому, что человеку нет необходимости думать, получать знания, тренировать логическое мышление, анализировать явления окружающей действительности, стоит отключить здравый смысл и положиться на манипуляцию человеческим сознанием искусственным интеллектом. Как справедливо отмечает В. А. Лекторский: «Цифровизация приносит немало благ, решает множество проблем, но также может создавать угрозы самому бытию человека» [4, с. 29].

Мы полагаем, что такой результат является следствием лингвокогнитивного манипулирования посредством искусственного интеллекта, используемого разнообразными платформами, существующими в виртуальном пространстве, поскольку это не только инструмент обмена информации среди пользователей, но и мощное социально-политическое средство воздействия на социум. Искусственный интеллект обладает большим выбором средств и техник манипулирования. Кроме того, человек сам, по своей воле предоставляет неограниченные возможности для манипулирования общественным сознанием, обучая искусственный интеллект и направляя его, вполне возможно, против себя самого. Получая однотипную информацию с нескольких ресурсов, современный человек принимает её как истину, редко подключая критическое мышление или здравый смысл. Новое поколение, а точнее поколение Z, у которых «я» физическое живёт в реальном мире, а «я» цифровое – в виртуальном, воспринимает представленную в сети информацию как единственно верную, и даже при условии, что $2+2$ будет равняться не 4, ответом может являться утверждение о том, что на просторах интернета так написано, а значит, проверять не имеет смысла и проще всего переделать таблицу умножения, так как нет, по мнению молодых людей, прямых доказательств, что $2+2=4$. Отдавая на откуп искусственному интеллекту мыслительную деятельность человека,

мы попадаем в социум манипулируемых существ – не это ли основная гуманитарная проблема современного мира?

Кроме того, в последнее время получили распространение так называемые Computer Generated Images (CGIs) («набор образов, сгенерированных компьютером»), которые являются электронными копиями реальных людей, как правило, известных актёров и актрис, которые «снимаются» в электронном кинематографе, не подозревая об этом. Данные копии могут, используя авторитет известных людей, транслировать любую информацию (не всегда адекватную), тем самым оказывая влияние на формирование общественного мнения.

Любой факт может быть искажён, зафиксирован в лингвистических моделях, растиражирован при помощи ботов, которые совершают вирусную рассылку, и подрастающее поколение уже уверено, что шекспировская Джульетта – темнокожая девушка, да и в праве ли мы определять пол? Это за нас также сделает искусственный интеллект, манипулируя сознанием человека. «Все когнитивные процессы, независимо от их осознанности, рутинности / тривиальности или новаторства и даже уникальности, обеспечиваются мозговыми процедурами. Огромную роль в ментальном мире человека играет язык, который не только обеспечивает основной вид коммуникации, но и является главным инструментом мышления: мы не просто *Homo sapiens*, но *Homo loquens*, а значит, по словам У. Эко, “мы утоплены в языке, кардинально от него зависим”» [6, с. 675–686].

И сегодня в современных условиях есть вполне реальная угроза превратиться из *Homo sapiens* и *Homo loquens* в *Homo Insipiens* – и разрушить всё то, что создавалось человеческим гением на протяжении всей истории существования.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ КАК ГУМАНИТАРНАЯ ПРОБЛЕМА

Использование ИИ в сфере образования в конечном итоге может привести к печальным последствиям, когда мы столкнёмся с повальным использованием школьниками и студентами ИИ для подготовки домашнего задания, написания исследовательских работ, дипломных проектов. О каком качестве образования в этом случае может идти речь и каких «специалистов» мы получим? В. Л. Макаров отмечает, что нет уверенности в том, что «искусственный интеллект, и соответственно,

искусственное общество могут, хотя бы теоретически, полностью быть адекватными (быть полной копией) человеку и обществу людей. Сознание пока или в принципе не может быть объяснено вычислительным способом» [5, с. 68].

Технологический рывок конца XX – начала XXI вв. поставил перед учёными новую задачу: научиться сосуществовать с постоянно прогрессирующим искусственным интеллектом (ИИ), который на данный момент требует пристального внимания лингвистов для обеспечения безопасности общества, так как бесконтрольное развитие данного механизма может привести к гуманитарной катастрофе. Следует помнить, что, несмотря на то, что «искусственный интеллект представляет собой комплекс технологических решений, имитирующих когнитивные процессы человека ..., при всех вариантах технического развития ... искусственный интеллект всегда будет связан с деятельностью человека» [1]. Учёными подтверждается, что ИИ имеет потенциал выявлять уязвимые места в человеческом поведении и, используя их, оказывать влияние на принятие решений, что является манипуляцией сознанием.

Научное мировое сообщество, ведущие информатики, программисты, физики и др. сходятся в едином мнении о том, что «в области искусственного интеллекта должна быть очень серьёзно прописана законодательная база»⁵. Проблемы возникают с открытыми базами данных, куда разработчик «складывает» свои программные опции, которые можно использовать для нейронных сетей, поэтому их нужно очень внимательно анализировать, потому что там зачастую могут уже содержаться коды ошибок, что создаёт определённую опасность.

⁵ Емельяненко А. Президент РАН Геннадий Красников: Развитие искусственного интеллекта надо увязать с законом // RG.ru: [сайт]. URL: <https://rg.ru/2023/04/04/prezident-ran-gennadij-krasnikov-razvitie-iskusstvennogo-intellekta-nado-uviazat-s-zakonom.html?ysclid=lofubrszi1156469147> (дата обращения: 21.10.2023).

ЗАКЛЮЧЕНИЕ

Гуманитарная угроза со стороны искусственного интеллекта, безусловно, на данный момент и, будем надеяться, далее – контролируема, так как без языковой модели, которая изначально создаётся человеком, искусственный интеллект порождать тексты, а значит, и полностью манипулировать сознанием человека *пока* не может (*выделено нами – Н. О., А. З.*). Задача лингвистов и филологов создавать такие языковые модели, которые при любой комбинации не упускали и не искажали бы человеческие ценности, а также позволили бы вывести рассуждения искусственного интеллекта на доверительный уровень: «ИИ обладает значительным потенциалом в научных исследованиях прежде всего за счёт возможности анализировать скрытые от человека закономерности, предлагая дополнительные описательные возможности, от нового языка до новой науки, объединяющей различные дисциплины» [3, с. 93–105].

Однако нужно помнить, что отсутствие правовой базы для контроля производительности искусственного интеллекта несёт угрозу человечеству. Существо разумное, т. е. человек, наделённое способностью анализировать, рассуждать, мыслить критически, отдаёт на откуп хоть и интеллекту, но с ключевой составляющей «искусственному», т. е. созданному кем-то, принятие порой судьбоносных решений.

В заключение процитируем открытое письмо И. Маска и экспертов-создателей искусственного интеллекта, опубликованное на сайте «Future of Life Institute», которое состоит всего лишь из 21 слова, но ёмко обобщает масштабность угрозы человечеству со стороны искусственного интеллекта: «Снижение риска вымирания из-за ИИ должно стать глобальным приоритетом наряду с другими социальными рисками, такими как пандемии и ядерная война»⁶.

Мы полагаем, что результаты нашего исследования помогут осмыслить опасность манипуляции человеческим сознанием посредством искусственного интеллекта и учесть в перспективе возможности искусственного интеллекта, а также выработать методики, которые позволят регламентировать возможности применения искусственного интеллекта.

⁶ Открытое письмо Илона Маска [Электронный ресурс]. URL: <https://futureoflife.org> (дата обращения: 01.10.2023).

ЛИТЕРАТУРА

1. Дзялошинский И. М. Искусственный интеллект: гуманитарная перспектива // Вестник Новосибирского государственного университета. Серия: История, филология. 2022. № 6. С. 20–29.
2. Дубровский Д. И. Развитие искусственного интеллекта и глобальный кризис земной цивилизации (к анализу социогуманитарных проблем) // Философия науки и техники. 2022. Т. 27. № 2. С. 100–107.
3. Искусственный интеллект для науки и наука для искусственного интеллекта / К. В. Анохин, К. С. Новоселов, С. К. Смирнов, А. Р. Ефимов, Ф. М. Матвеев // Вопросы философии. 2022. № 3. С. 93–105.
4. Лекторский В. А. Искусственный интеллект в изучении человека, человек в мире, создаваемом искусственным интеллектом // Человек и системы искусственного интеллекта / под ред. В. А. Лекторского. СПб.: Юридический центр, 2022. С. 10–18.
5. Макаров В. Л. Как современные технологии меняют представление об окружающем мире (искусственное общество) // Человек и системы искусственного интеллекта / под ред. В. А. Лекторского. СПб.: Юридический центр, 2022. С. 59–71.
6. Мозг человека и многозначность когнитивной информации: конвергентный подход / Т. В. Черниговская, В. М. Аллахведов, А. Д. Коротков, В. А. Гершкович, М. В. Киреев, В. К. Прокопеня // Вестник Санкт-Петербургского университета. Философия и конфликтология. 2020. Т. 36. № 4. С. 675–686.
7. Хабаров А. А. Техники лингвокогнитивного манипулирования в реалиях информационно-психологического противоборства // Актуальные проблемы филологии и педагогической лингвистики. 2021. № 4. С. 72–82.
8. Черниговская Т. В. Чеширская улыбка кота Шредингера: мозг, язык и сознание. М.: АСТ, 2023. 496 с.
9. Chichulin A. Neural networks guide. Unleash the power of neural networks: The complete guide to understanding, implementing AI. Екатеринбург: Издательские решения, 2023. 116 p.
10. Van Dijk T. A. Discourse, Semantics and Ideology // Discourse and society. 1995. № 2. P. 242–257.

ARTICLE INFORMATION

Author

Anastasia V. Zinkovskaya

- e-mail: anastassiat@bk.ru
Dr. Sci. (Philology), Prof., Departmental Head, Department of English Philology, Faculty of Romance and Germanic Philology
Kuban State University
ul. Stavropolskaya 149, 350040 Krasnodar, Russian Federation

Natalya N. Olomsкая

- e-mail: olomnat@mail.ru
Dr. Sci. (Philology), Assoc. Prof., Prof., Department of English Philology, Faculty of Romance and Germanic Philology
Kuban State University
ul. Stavropolskaya 149, 350040 Krasnodar, Russian Federation

For citation:

Zinkovskaya A. V., Olomsкая N. N. Manipulation of human consciousness by means of artificial intelligence as a humanitarian problem of the XXI century. In: *Russian Social and Humanitarian Journal*, 2024, no. 1.

Available at: www.evestnik-mgou.ru

Abstract

Aim. To analyze the possibilities of manipulating a person's consciousness through the use of artificial intelligence (AI).

Methodology. In the process of research, the authors used the interdisciplinary approach when considering the specifics of the AI manipulative effect on society in general and the individual in particular. This approach made it possible to apply a number of methods and techniques for analyzing texts generated using neural networks, such as contextual analysis, component and comparative analysis.

Results. The consequence of linguocognitive manipulation through social networks, as well as the diverse platforms existing in the virtual space, is a powerful socio-political impact on society, since social networks are not a mere tool for exchanging information among users, since it is social networks that provide a wide range of means and techniques of manipulation as well as unlimited opportunities for the AI use to manipulate public consciousness.

The capabilities of neural networks are practically unlimited and, despite the fact that the results of AI work are collectively identical, they are most often ideologically colored and do not offer a choice of either vocabulary or syntax. The use of such neural networks leads to a decrease in the creative potential of the individual, motivation to study and comprehend the phenomena of social realm, which, ultimately, will lead to the manipulation of human consciousness by AI.

Research implications. The results of scientific research contribute to the further comprehension of the danger of consciousness manipulation by artificial intelligence as a humanitarian problem. Analysis of texts generated using artificial intelligence, as well as analysis of language tools, will help in the future to take into account the capabilities of artificial intelligence and develop techniques to limit its scope.

Keywords:

Artificial intelligence (AI), humanitarian problem, manipulation, manipulative impact, neural networks

References

1. Dzyaloshinsky I. M. [Artificial intelligence: a humanitarian perspective]. In: *Vestnik NGU. Seriya: Istoriya, filologiya* [Vestnik NSU. Series: History and Philology]. 2022, no. 6, no. 20–29.
2. Dubrovsky D. I. [The development of artificial intelligence and the global crisis of earthly civilization (to the analysis of socio-humanitarian problems)]. In: *Filosofiya nauki i tekhniki* [Philosophy of science and technology], 2022, vol. 27, no. 2, pp. 100–107.
3. Anokhin K. V., Novoselov K. S., Smirnov S. K., Efimov A. R., Matveev Ph. M. [AI for science and science for AI]. In: *Voprosy Filosofii* [Issues of philosophy], 2022, no. 3, pp. 93–105.
4. Lektorsky V. A. [Artificial intelligence in human studies, the human being in the world, created by artificial intelligence]. In: Lektorsky V. A., ed. *Chelovek i sistemy iskusstvennogo intellekta* [Man and artificial intelligence systems]. St. Petersburg, Yuridicheskii Tsentr Publ., 2022, pp. 10–18.
5. Makarov V. L. [How modern digital technologies are changing perceptions of the world around us (artificial society)]. In: Lektorsky V. A., ed. *Chelovek i sistemy iskusstvennogo intellekta* [Man and artificial intelligence systems]. St. Petersburg, Yuridicheskii Tsentr Publ., 2022, pp. 59–71.
6. Chernigovskaya T. V., Allakhverdov V. M., Korotkov A. D., Gershkovich V. A., Kireev M. V., Prokopenya V. K. [Human brain and

ambiguity of cognitive information: a convergent approach]. In: *Vestnik Sankt-Peterburgskogo universiteta. Filosofiya i konfliktologiya* [Vestnik of Saint-Petersburg University. Philosophy and Conflict Studies], 2020, vol. 36, no. 4, pp. 675–686.

7. Khabarov A. A. [Techniques of linguistic cognitive manipulation in the information and psychological confrontation environment]. In: *Aktual'nye problemy filologii i pedagogicheskoi lingvistiki* [Current Issues in Philology and Pedagogical Linguistics], 2021, no. 4, pp. 72–82.
8. Chernigovskaya T. V. *Cheshirskaya ulybka kota Shredingera: mozg, yazyk i soznanie* [Cheshire smile of Schrodinger's cat: brain, language and consciousness]. Moscow, AST Publ., 2023. 496 p.
9. Chichulin A. Neural networks guide. Unleash the power of neural networks: The complete guide to understanding, implementing AI. Ekaterinburg, Izdatelskiye resheniya Publ., 2023. 116 p.
10. Van Dijk T. A. Discourse, Semantics and Ideology. In: *Discourse and society*, 1995, no. 2, pp. 242–257.