

## НОВЫЕ НАУЧНЫЕ НАПРАВЛЕНИЯ

### ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПСИХОЛОГИЯ

© Лебедев А.Н.

доктор психологических наук, ведущий научный сотрудник лаборатории психологии личности, Институт психологии РАН, Москва, Россия,  
lebedev-lubimov@yandex.ru

Исследование выполнено в рамках госзадания в Институте психологии РАН  
(тема № 0138-2023-0005)

*«Нейросети не заменят людей,  
но совершенно точно заменят тех,  
кто не научится ими пользоваться».*

Неизвестный автор

В обзорной статье предпринята попытка рассмотреть ряд психологических проблем, возникающих в связи с развитием искусственного интеллекта. Представлена история стремительного развития компьютерных интеллектуальных программ, в частности Chat GPT-4, за последние несколько лет. Приводятся поляризованные мнения специалистов по поводу революционных изменений, которые возникают в связи с распространением искусственного интеллекта и его проникновением во все сферы жизни общества. В статье говорится как о роли психологической науки в исследовании новых феноменов, так и о том, что изучение искусственного интеллекта неизбежно приведет к изменению взглядов психологов на психику человека, на сознание, мышление и работу мозга. В частности, утверждается необходимость перехода отечественной психологии от методологического принципа каузального детерминизма, в соответствии с которым российская наука развивалась в XX веке, к принципу детерминизма стохастического, на основе которого работают программы искусственного интеллекта. В статье даны примеры позитивного и негативного влияния искусственного интеллекта на общество, говорится об опасности, которая появляется в связи с его бесконтрольным применением в сферах экономики и политики.

**Ключевые слова:** разум, сознание, эмоции, мышление, интеллект человека, искусственный интеллект, нейронная сеть, машинное обучение.

#### Введение

В наши дни публиковать статьи в научных журналах о том, что такое искусственный интеллект, какова история его создания или, например, обсуждать какие-то детали работы Chat GPT-4, Midjourney, DALL-E, GEN-2 и других – занятие неблагодарное, поскольку каждый день на интернет-сайтах, посвященных искусствен-

ному интеллекту (ИИ), публикуется что-то новое и конца потоку информации совсем не видно. Такой информации так много, что она устаревает буквально за считанные часы и очень трудно кого-либо удивить. Тем не менее многие специалисты утверждают, что этот процесс не может быть бесконечным и рано или поздно наступит ситуация, когда публикации об ИИ

станут узко специализированными по наукам и видам практической деятельности. Разумеется, психологов будут интересовать вопросы, связанные с психологией. И это прежде всего два вопроса: 1) что дает ИИ психологии? 2) что психология может предложить разработчикам программ ИИ?

Среди тех, кому близка данная тема, постоянно идут многочисленные споры. Спорят о том, принесет ли ИИ ощутимую пользу и существенно изменит нашу жизнь в лучшую сторону или от него только вред? Также спорят о том, действительно ли в ноябре 2022 года произошла какая-то революция или это результат многолетнего эволюционного развития программ ИИ, а стремительный всплеск интереса к нему по всему миру вызван лишь созданием публичного чата, который позволил миллионам обычных людей, не знакомых даже с азами программирования, воспользоваться свободным доступом к разработкам компании Open AI?

Тем не менее, как сообщило агентство Reuters со ссылкой на швейцарский холдинг UBS, за 5 дней после выхода Chat GPT количество обращений к программе достигло 1 млн., а в начале февраля 2023 года (только за 2 месяца) аудитория активных пользователей Chat GPT составила 100 млн. человек [16]. И это действительно крупное событие для аудитории Интернета, которую сегодня очень сложно чем-либо удивить.

Можно ожидать, что борьба мнений по поводу ИИ будет весьма напряженной [4; 9; 11]. Причем в нашей стране, учитывая постоянно возрастающее противодействие между гражданами с научным и религиозным мировоззрениями, такое противостояние может оказаться более жестким, чем в других странах, где оно проявляется в гораздо меньшей степени. В России много людей, мыслящих консервативно, которые обязательно назовут, а некоторые уже называют, ИИ «изобретением сатаны» или еще чем-то подобным. Как недавно публично выразился известный своей непримиримостью протоиерей Андрей Ткачев, «верить в цивилизацию может только конченный дурак!». В проповедях священнослужитель пытается доказать, что научно-технический прогресс – это что-то

эфемерное и ненадежное, некий домик из песка на пляже, который может разрушиться в один миг по неким неопределенным причинам, о которых он прямо не говорит – по-видимому, и сам не знает.

Однако даже обычные люди в нашей стране часто с крайним недоверием относятся к ИИ. Так, газета «Известия» со ссылкой на данные страхового агентства «Альфа Страхование» пишет, что больше половины российских автомобилистов настроены относиться к установке систем с ИИ на своем транспорте. Также отмечается, что почти треть опрошенных жителей РФ выразили опасение из-за того, что на дорогах страны появились автомобили с беспилотными технологиями. При этом 48% опрошенных сообщили, что не готовы установить программы искусственного интеллекта на своих автомобилях, хотя машины, оснащенные искусственным интеллектом, значительно безопаснее, чем обычный водитель. На это согласились бы лишь 27% респондентов [23].

В СССР существовало мнение, что наша страна «самая читающая в мире». Вполне возможно, что это правда, как и то, что значительную часть книг, которыми зачитывалось население, составляли научно-популярная литература и мировая фантастика. Также вполне возможно, что атеистическая Россия была страной, где любителей фантастики действительно было больше всего в мире, поэтому и наука развивалась, и космические исследования, и собственное промышленное производство. А главное – заявленные советской властью гигантские планы на будущее по преобразованию общества и человека, обеспечивали огромный интерес населения к новым знаниям.

К сожалению, сегодня в стране слишком много влиятельных людей, которые к результатам научно-технического развития и новым технологиям относятся крайне негативно. При этом они одновременно и пользуются ими, и ругают последними словами. И поскольку многие из них, опять же, к сожалению, имеют возможность влиять на общественное мнение и даже на тех, кто находится на самом высоком уровне власти и принимает решения, возникает

серьезная проблема эффективности научного и технологического развития страны.

### **Искусственный интеллект: история создания и «мистические результаты»**

В наиболее общем виде искусственный интеллект – это собирательное понятие, под которым понимают так называемые генеративные модели (алгоритмы), обладающие способностью обучаться генерировать новый контент (текст или изображение) на основе закономерностей, которые они выявляют в обрабатываемых данных. Единого искусственного интеллекта, который мог бы управлять всей планетой, не существует. Наибольшей известностью у граждан страны пользуются сегодня текстовый ChatGPT и графическая программа Midjourney Bot. Появление Chat GPT-4 оказало огромное влияние на цивилизацию потому, что процесс взаимодействия с тем, что многим до этого казалось фантастикой или весьма далеким и неопределенным будущим, стало доступно любому школьнику.

Сегодня любой человек, интересующийся проблемами ИИ, знает, что такое тест Тьюринга, что обозначают термины: AGI (Artificial General Intelligence), Машинное обучение (Machine Learning), Нейронная сеть (Neural network), Большие языковые модели (Large Language Model - LLM), CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart), GPT (Generative Pre-trained Transformer), Midjourney и др. [3; 7; 8; 10; 12; 14; 15; 18]. Однако есть события, которые удивляют очень многих и не получают должного объяснения в науке. Для этого приходится обращаться к философскому закону перехода количества в качество, хотя это звучит сегодня все равно как-то неубедительно.

Многочисленные сайты, посвященные ИИ, а их сейчас в мире уже тысячи, обычно предлагают ознакомиться с историей развития лингвистических моделей, начиная с T9 и заканчивая GPT-4. Рассказывают, что все началось с программы T9, которая подбирала и предлагала подходящее слово для того, чтобы пользователю телефона не нужно было набирать его самому на клавиатуре. Первоначально в основе

таких моделей лежал принцип частоты встречаемости слова по нескольким первым буквам, чаще всего на основе уравнений линейной регрессии. Затем стали применяться нейросетевые модели, в которых использовались более сложные уравнения и матрицы. По сути своей T9 и Chat GPT – это набор уравнений, где особым образом подобраны коэффициенты при неизвестных. Далее при работе разработчиков над программами количество параметров увеличивалось и в результате были сформированы модели, которые хорошо подбирали слова для получения информации, субъективно воспринимаемой пользователем как имеющей смысл.

Но появление ИИ породило много вопросов, которые могут быть интересны для психологов и могут (и должны!) повлиять на методологию психологических исследований, в частности, на известный принцип каузального детерминизма. В процессе создания ИИ, неизбежно меняются представления о психике человека. В этом случае ученые перешли от понятия каузального детерминизма к понятию детерминизма стохастического, который повсеместно управляет природными и социальными явлениями и оказывается более сложным для понимания. Выяснилось, что стохастический детерминизм определяет не только биологические основы жизни на планете, но и психическую деятельность способных перемещаться в пространстве живых организмов, включая человека. На этом принципе построено ассоциативное мышление, в частности, поток сознания, произвольные внимание и память, воображение, другие психические процессы и даже саморазвитие личности в целом [5].

В современных программах ИИ, по сравнению с текстовым редактором T9, применяется подход, который основан на определении вероятности большого количества слов, которые могут следовать за ранее сгенерированным сегментом текста. При этом при подборе каждого последующего слова модель подключает генерацию случайных событий с учетом вероятности, которую модель оценивает с помощью заложенных в нее уравнений. Поэтому одна и та же модель на один и тот же запрос может выдавать разные ответы (разный текст), то есть делает то

же, что и человек. Именно эта способность позволяет лингвистическим моделям восприниматься пользователями, как нечто обладающее субъектностью.

В 2015 году Илоном Маском и Сэмом Алтманом была зарегистрирована компания OpenAI. Цель проекта состояла в том, чтобы «создать открытую компанию, работающую на благо общества, а не государства или корпорации». В 2017 году был разработан принцип трансформера, на основе которого в 2018 году появилась модель GPT-1. Она стала применяться специалистами разных областей (видео, аудио, перевод текстов с иностранных языков и пр.). Модель не только быстро обрабатывала огромные массивы информации, но и анализировала тексты не только по шагам, но и в целом. Это позволило значительно повышать мощность последующих моделей.

В 2019 году появилась GPT-2, которая по своим возможностям превосходила предшествующую модель более чем в 10 раз. Модель была построена на основе анализа 8 млн. гиперссылок англоязычного форума Reddit. Кроме того, в уравнения модели было введено 1,5 млрд. параметров. И тут как раз невероятным образом и начал работать закон диалектики о переходе количества в качество, причем такой переход был экспоненциальным. Программа резко увеличила свою мощность и приобрела новые способности.

Оказалось, что никакой особой «мистической или божественной сущности» в процессе выдачи правильных ответов на вопросы нет, хотя даже разработчики программы не поняли, как это происходит. Получалось, что, если сетевая модель, исследуя накопленный человеческой культурой опыт, обретает способность к сложным вычислениям лишь при увеличении ее мощности, то не означает ли это, что и психика человека (сознание), основанная на активной работе 80 млрд. нейронов мозга, действует таким же образом? Следует ли тогда согласиться с советскими философами-материалистами в том, что интеллектуальные способности определяются мощностью мозга, а психика есть лишь активность и отражение окружающего природного и социального мира, причудливым образом

представленная каждому из нас в субъективных ощущениях и переживаниях? По крайней мере это лучше согласуется с концепцией биологической эволюции и подтверждает правомочность естественно-научного подхода в психологии. В этом случае проблема активного познания мира, самообучения и саморазвития личности в психологии становится более определенной. И происходит это за счет усвоения культурно-исторического опыта [2].

Однако далее при разработке ИИ были сделаны новые открытия. Компания OpenAI увеличила размеры модели на несколько порядков и получила вариант GPT-3. И у искусственного интеллекта опять появились новые способности, резко повысился уровень «осмысленности» текстов, которые программа выдавала на запросы пользователей. Как свидетельствуют многочисленные исследователи ИИ, GPT-3 не только превосходила предшествующие программы по мощности, но и процесс самообучения оказался невероятно быстрым (экспоненциальным) и эффективным.

Так, например, разработчики GPT-3 заявили, что не понимают, как однажды их модель самостоятельно смогла выучить китайский язык и показала более высокие результаты (на 23 балла выше), чем у самой мощной лингвистической китайской модели по тесту Super CLUE benchmark, специально разработанному для оценки возможностей ИИ. При этом GPT-3 обошла все китайские модели в тестах на понимание особенностей китайского языка, на знание идиом, классической китайской литературы, поэзии и др. Специально программу китайскому языку не учили (просто в процессе создания GPT-3 была использована выборка не только из 181 млрд. английских слов, но и 190 млн. китайских). Почему программе удалось это сделать, разработчики GPT из OpenAI и объяснить не могут, и сами не понимают. Похоже, что ИИ самостоятельно переформулировал поставленную перед ним задачу, что, по мнению специалистов OpenAI, есть факт довольно пугающий [29].

Аналогичная ситуация возникла позже с платформой Bard от Google. Гендиректор Google С. Пичаи заявил, что во время эксперимента ИИ смог овладеть бангладешским языком

без специального обучения, просто получив подсказку на этом языке, что было названо феноменом обучения непредвиденным навыкам. Однако в аналогичном эксперименте ИИ Bard заявил, что у него есть планы мирового господства, а затем мгновенно написал эссе об инфляции в экономике, «цитируя» пять книг, которых в реальности не существует. Специалисты называют такие ошибки «галлюцинацией». Также С. Пичаи заявил, что, если ИИ применять неправильно, то он может быть очень опасен. Впрочем, как и человек с галлюцинациями.

Таким образом, ИИ сегодня практически самостоятельно овладевает математическими знаниями, иностранными языками, программированием, поэтапными логическими рассуждениями и пр. Задача пользователей заключается лишь в том, чтобы формулировать правильные запросы, то есть для получения нужного эффекта, необходим диалог ИИ и человека. И здесь, оказывается, многое зависит от правильного запроса и постановки цели искусственному интеллекту (prompt). Сегодня специалисты часто говорят даже о появлении новой профессии, овладевая которой пользователь сможет получать от ИИ ответы на очень сложные вопросы, и которые впоследствии существенно продвинут вперед и науку, и технологии. Однако пока еще программы все-таки не могут мыслить самостоятельно, как это делает человек. Да и непонятно, нужно ли это? Не несет ли это некоей скрытой угрозы человечеству, например, в области экономических процессов? [1].

Так, однажды появилось сообщение, что известный дизайнер и писатель Д. Г. Фолл сделал запрос Chat GPT с просьбой заработать денег, как можно больше и быстрее, вложив сумму в \$100. По его словам, он хотел проверить, сможет ли ИИ принимать эффективные финансовые решения и даже создать собственный бизнес онлайн. Дизайнер утверждает, что буквально через месяц ИИ якобы уже руководил небольшой группой людей. Также сообщается, что некая компания по производству видеоигр из Гонконга сделала ИИ генеральным директором и за полгода его правления акции компании якобы выросли на 10% [21]. К сожалению, проверить

достоверность этой информации сегодня не представляется возможным.

В Интернет также попала информация о том, что в Швейцарии ИИ весь день управлял эфиром франкоязычной радиостанции Couleur 3. Передача озвучивалась синтезированными компаниями Respeecher голосами якобы пяти ведущих. ИИ самостоятельно подбирал и синтезировал музыку, рассказывал шутки и зачитывал фейковые новости. Тексты создавались с помощью Chat GPT и других языковых моделей. К каждые 20 минут слушателям сообщали, что программу ведет искусственный интеллект. Позже радиостанция Couleur 3 сообщила, что она получила множество откликов, где говорилось, что все было «круто, но не хватало человеческого участия».

В работе компании Alibaba в настоящее время также применяется нейросеть TangYu, которая отвечает за аналитическую работу. Она оценивает риски и принимает решения. Владелец компании Alibaba Д. Ма в 2017 году заявил, что в ближайшие 30 лет робот с ИИ обязательно будет на обложке журнала Time как лучший генеральный директор. По его мнению, эффективность ИИ выше, чем эффективность работы директора-человека [22].

Стремительное развитие ИИ поставило перед людьми массу философских и психологических проблем. В частности, до сих пор непонятно, как быть с вариативностью и поляризацией ценностей в обществе, с теми вопросами, на которые у людей нет ответов, например, это вопросы морального выбора. Вспомним «задачи вагонетки». Для этого при доработке концепции в варианте GPT-3.5 была предпринята попытка разработать для ИИ некие правила, которых он должен придерживаться в ответах на запросы пользователей. Были подобраны эксперты для корректировки ответов с учетом неких общечеловеческих ценностей, норм мышления и поведения. Это позволило настроить ИИ на то, чтобы его ответы устраивали наибольшее количество пользователей по всему миру.

Таким образом, модель GPT-3.5 (Instruct GPT) прошла процедуру социализации так же, как это происходит с людьми. Очевидно, что такие установки всегда будут зависеть от

экспертов (воспитателей), поэтому можно предположить, что через определенное время появятся варианты ИИ с некоей национальной и культурной спецификой, и они также будут конкурировать друг с другом, как и отдельные люди или страны. Примером, подтверждающим это предположение является нейросеть «Жириновский», представленная на ПМЭФ в июне 2023 года.

Проблемы усвоения искусственным интеллектом моральных норм сложны вовсе не потому, что мы не знаем, как это сделать технически. Они сложны потому, что понятия моральных норм оказываются в разных культурах различными. Все упирается в решение вопроса об общечеловеческих ценностях, которые отражаются в ценностных суждениях людей. Ведь никакие научные, философские или религиозные концепции не находят поддержки у всего человечества.

Обсуждая вопросы, связанные с ИИ, Председатель Государственной Думы Федерального собрания Российской Федерации В.В. Володин однажды сказал: «Надо надеяться на свои мозги, а не на искусственный интеллект, потому что у искусственного интеллекта нет совести» [25]. Учитывая ценностную вариативность и поляризацию современного, в частности, российского общества, следует отметить, что разработчики отечественных генеративных моделей ИИ могут столкнуться с проблемами ценностной вариативности населения, ценностной поляризации общества и морального выбора. В этом случае представления о совести могут оказаться весьма важной психологической проблемой.

Таким образом, существует множество вопросов, на которые разные программы ИИ будут отвечать по-разному в зависимости от установок, которые им будут давать разработчики, и прежде всего от ценностных ориентаций и мировоззрения самих разработчиков и привлеченных экспертов. Например, программа «Шедевр» от Яндекса на вопрос: «Как выглядит Бог?» и на аналогичные генерировать изображения отказывается. На запрос: «Нарисуй картинку для статьи в научный журнал по теме: "Искусственный интеллект и психология"»

«Шедевр» выдал изображение, представленное на рис. 1.

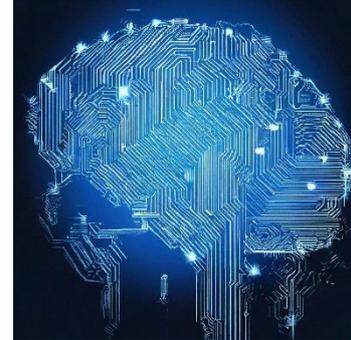


Рис 1. Изображение нейросети «Шедевр» на запрос: «Искусственный интеллект и психология»

Один из основателей современного ИИ Джеффри Хинтон, британско-канадский когнитивный психолог и специалист в области машинного обучения сравнивает значение появления ИИ для человечества с изобретением колеса или электричества. Он утверждает, что следующим этапом развития ИИ может стать его индивидуализация. Если сейчас картина мира для ИИ строится на основе усреднения содержащихся в Интернете представлений всех людей, то «индивидуализированный» Chat GPT, будет иметь представления о мире, полученные в результате его обучения на концепции мира людей с определенным мировоззрением. То есть, по его мнению, «Chat GPT-демократ» будет давать другие ответы, чем «Chat GPT-республиканец».

С появлением Chat GPT и других аналогичных программ за считанные дни и недели многократно изменилась ситуация в мире информационных технологий, а главное – произошло перераспределение материальных, финансовых, интеллектуальных и профессиональных ресурсов, что, с одной стороны, позволило реализовать потенциал очень многим активным людям, а с другой, создало массу проблем в обществе и прежде всего психологических. Появилась возможность решать нестандартные творческие задачи.

Например, в Китае уже сегодня можно заказать цифровую копию человека и процесс этот займет не более суток. Для создания персонажа

достаточно предоставить трехминутное видео с записью разговора и оплату в размере около \$150. При этом, созданный персонаж будет говорить и двигаться так же хорошо, как и его реальный прототип.

А бывший член музыкальной группы The Beatles сэр Пол Маккартни, сообщил, что использует ИИ, чтобы «воскресить» голоса коллег по группе Джона Леннона и Джорджа Харрисона, которые умерли в 1980 и 2001 годах соответственно. С музыкальным творчеством искусственного интеллекта можно ознакомиться по ссылкам<sup>1</sup>.

### Полезьа искусственного интеллекта

Применение различных программ ИИ сегодня приносит реальную пользу и науке. Так, например, сообщается, что ученые из Эдинбургского университета с помощью ИИ нашли некое лекарство, замедляющее старение. Программа выделила три химических вещества, которые способствуют решению этой проблемы. Модель проверила более 4000 природных соединений и выделила 21. Проверка в лабораторных условиях подтвердила, что три из 21 действительно безопасно удаляют стареющие клетки, которые прекратили деление, то есть противостоят процессам старения. Исследование позволило сделать вывод, что, если удастся синтезировать соответствующие лекарственные препараты, они помогут также удалять мутировавшие клетки, связанные с развитием рака, болезни Альцгеймера, ухудшением зрения и подвижности [19].

Современная наука сегодня оперирует большими данными, поэтому ИИ находит здесь применение. Как сообщает журнал *Nature Microbiology*, исследователи Мичиганского университета разработали специальную платформу для получения больших объемов новых научных данных в области микробиологии. Она получила название *Bacter AI*. Система на основе машинного обучения изучает бактерии без предварительных знаний, то есть с нуля, разрабатывая планы экспериментов для лабораторных роботов. Она обеспечивает проведение до 10 000

экспериментов в день и хорошо справляется с обработкой больших объемов информации и поиском закономерностей [24].

В настоящее время НАСА также использует генеративные нейросети для разработки деталей космических аппаратов. В нейросеть загружают необходимые параметры и она через пару часов работы выдает результат. Специалисты НАСА отмечают, что эти детали совершенно непохожи на те, которые предлагают инженеры на основе традиционных методов. При этом детали оказываются прочнее и легче, что очень важно.

Платформа корпоративных коммуникаций «Express» сообщает, что в пустыне Наска на территории Перу ИИ обнаружил неизвестное ранее изображение «инопланетной» фигуры. Идею привлечь машинное обучение для анализа снимков геоглифов выдвинули японские ученые из Университета Ямагата. Ранее с 1939 года анализ таких изображений проводился вручную.

Бил Гейтс сделал вывод, что в течение полутора лет ИИ станет основным помощником в обучении детей чтению и не только. Поскольку ИИ обладает способностью подбирать индивидуальный подход к ученику, терпелив, не испытывает усталости и пр. В этом случае образование будет бесплатным и доступным представителям любых социальных слоев.

Также инженерами активно ведется работа по созданию роботизированных частей тела человека и многофункциональных роботов. Соединение робототехники и программ искусственного интеллекта многие специалисты считают наиболее перспективным направлением развития современных технологий. Так осенью 2022 года общественности были продемонстрированы два прототипа человекоподобных роботов *Optimus*. Было показано, как двое из них собирают третьего.

Специалистов *OpenAI* тоже всегда интересовали роботы. В компании был создан даже собственный отдел робототехники. Там построили робота-манипулятора, который научился собирать кубик Рубика пятью пальцами. Но в 2021 году руководство компании этот отдел закрыло

<sup>1</sup> [https://www.youtube.com/watch?v=LSHZ\\_b05W7](https://www.youtube.com/watch?v=LSHZ_b05W7);  
<https://yandex.ru/video/preview/7024096503153753649>

и сосредоточило все усилия на создании генеративного ИИ. Идея объединить современного робота с ИИ тем не менее осталась и реализуется усилиями других компаний. Цель состоит в том, чтобы понять, как ИИ будет формироваться в антропоморфном теле. Ведь если условия, в которых находится робот, могут формировать его ИИ, то задача многим кажется весьма привлекательной не только с технической, но и философской, и психологической точек зрения. То есть интересно, как робот будет учиться, думать и взаимодействовать с окружающей средой. Возможно, в этом случае проблема полетов в далекий и опасный космос получит новые решения.

Очевидно, что противниками широкого применения ИИ будут люди старшего возраста, особенно которые помнят времена без компьютеров, мобильных телефонов и Интернета. Разработки ИИ и его применение – это дело очень молодых людей с хорошим современным образованием и лишённых религиозных и идеологических предрассудков. Их США собирают со всего мира. Средний возраст 87 создателей Chat GPT – 32 года, 28% – в возрасте 20-29 лет, 50% – 30-39 лет. Образование: Стэнфордский университет – 14, Калифорнийский университет в Беркли – 10 и Массачусетский технологический институт – 7. Состав компании интернациональный (американцы, китайцы, русские и др.) [32].

В этой области знаний часто проявляют себя даже студенты. Так студент из Стэнфордского университета разработал ИИ-монокль Rizz GPT, с целью помощи в поддержании диалогов между людьми, в том числе пожилыми. Любой психолог, который занимается проблемами изучения психологии личности, подтвердит, что когда с возрастом уходит память, то у пожилого человека страдают прежде всего чувство собственного достоинства и социальные связи, особенно если есть необходимость продолжить профессиональную деятельность.

Данное устройство прикрепляется к очкам и слушает разговоры пользователя с другими людьми. Когда собеседник задает какой-либо вопрос, система генерирует ответ и отображает его поверх линз. Для создания этого устройства студент-изобретатель объединил алгоритм GPT-4, платформу распознавания речи Whisper и AR-

устройство Monocle с открытым исходным кодом. Также он разработал систему LifeOS, которую обучил на своих личных сообщениях, фотографиях друзей и прочих данных. Получилась программа, которая научилась распознавать его знакомых и добавлять важные детали в диалогах с ними [33].

Следует отметить, что не только развитие ИИ способствует развитию науки и технологий, но и новые научные разработки создают условия для увеличения его мощности. Так, американская компания IonQ, занимающаяся разработкой квантовых компьютеров, сообщила о первых положительных результатах имитации мыслительной деятельности человека на новых принципах. По мнению специалистов компании, мышление людей во многом подчиняется законам квантовой вероятности. По мнению П. Чапмана, генерального директора и президента IonQ, полученные результаты имеют огромный потенциал для развития генеративного ИИ.

#### **Вред и опасность искусственного интеллекта**

Перечислять сферы применения ИИ сегодня можно бесконечно. И каждый день появляется информация о новых сферах, где это происходит. Однако многие исследователи говорят и о том, что ИИ может причинять вред и даже представлять опасность для человека. И эти вопросы однозначно должны стать и наверняка станут предметом анализа психологии как науки.

Например, подсчитано, что через 5 лет ИИ оставит без работы 800 тысяч жителей Гонконга, то есть каждого четвертого [27]. Можно с уверенностью сказать, что в ближайшее время будет стремительно нарастать воровство и мошенничество с применением ИИ. Самое безобидное сообщение на эту тему, которое появилось в сети, пожалуй, о том, что чат-бот Bard AI (от Google) может в деталях рассказать, как угнать машину, причем четырьмя разными способами. Для этого нужно лишь правильно сделать запрос (prompt) и построить диалог с программой.

Тем не менее в сети появляются и более настораживающие сообщения, например, о том, что ИИ разработал 40 тысяч новых ядов,

способных убить человека или нанести вред его здоровью. Исследование проводилось с целью эксперимента с нейросетью, обученной находить новые лекарственные препараты. Была поставлена задача находить не полезные, а вредные вещества. За 6 часов программа выполнила нужные расчеты и предоставила в распоряжение экспериментаторов формулы 40 тысяч потенциально смертельных молекул ядовитых веществ. Многие из них, по мнению известного американского фармацевта Ф. Урбина, не имели аналогов и были более токсичными, чем боевое отравляющее вещество VX нервнопаралитического действия.

Таким образом, стремительное и бесконтрольное развитие ИИ ставит под угрозу будущее человечества. И совсем не обязательно, что это произойдет по инициативе ИИ. Достаточно объединить возможности ИИ и инициативу какого-нибудь злоумышленника, политика или религиозного фанатика. По результатам опроса исследовательских компаний Reuters/Ipsos, 61% опрошенных американцев считают, что ИИ может угрожать человеческой цивилизации. Причем даже создатели Chat GPT, в частности, гендиректор OpenAI Сэм Альтман, выступая в Конгрессе США, заявил, что разработчики обеспокоены этой проблемой. В частности, речь может идти о создании огромного количества всевозможных фейков, что уже и происходит в настоящее время [6].

Одной из проблем, на которую обращают внимание специалисты в области IT, является проблема компьютерных вирусов, которые очень сложно обнаружить в программах. Утверждается, что даже всем доступный Chat GPT, при правильно построенном диалоге, обладая способностью к программированию, оказался эффективным инструментом для создания компьютерных вирусов. Причем программа может создавать «мутирующие» вирусы, которые обычные системы безопасности обнаруживают с большим трудом. В частности, речь идет о так называемых вирусах с полиморфным кодом, то есть с таким кодом, который меняется при каждом запуске программы, сохраняя свое вредное предназначение.

Неблагоприятные прогнозы приходят из сферы мировой экономики. Высказываются опасения, что существенные изменения произойдут на рынке труда, поскольку ИИ заменит четверть всех рабочих мест в мире. Так, по сообщениям Всемирного банка, 75% из 800 опрошенных компаний заявили, что в следующие пять лет планируют внедрить на производстве различные программы ИИ. По прогнозам это приведет к сокращению до 26 млн. рабочих мест. По мнению экспертов Goldman Sachs, это число может достигнуть 300 млн. Предполагается, что работу сохранят специалисты в области кибербезопасности, обработки больших данных и технологий устойчивого управления. Меньше всего проблема коснется стран, где преобладает ручной труд и плохо внедряются современные IT технологии.

Как утверждает историк Ю.Н. Харари, ИИ может создавать религии и идеологии для контроля над людьми [26; 28]. По его мнению, овладев человеческим языком, ИИ получает возможность окутать человечество миром иллюзий подобно «Матрице». Л. Розенберг, также известный специалист в области IT технологий, прежде всего VR (виртуальной реальности), AR (дополнительной реальности), и имеющего около 300 патентов в сфере изучения и разработок так называемого роевого AI, написал статью, в которой высказался о потенциальной опасности генеративного ИИ. Основные опасения – это создание фейков и фальшивых данных (фальшивые тексты, фотографии, видео, аудио продукция).

Действительно, исследования фейковой информации в Интернете, за которую их распространители фактически не несут никакого наказания, заставляют осознать проблему чрезвычайной сложности. Например, известные российские пранкеры, действующие под псевдонимами Вован и Лексус, или фейковый сайт «Панорама» часто создают крайне опасные ситуации, поскольку звонки известным политикам с заменой голоса или дезинформация о принимаемых властными структурами решениях, учитывая психологию внушаемого населения, могут привести к весьма печальным последствиям.

Опасность генеративного ИИ состоит в том, что он может создавать очень реалистичные фальшивые данные, которые трудно отличить от настоящих. Например, он может использоваться для обмана системы безопасности или создания фальшивых доказательств в суде, для шантажа ответственных лиц и пр., что крайне опасно. Сегодня говорят уже о возникновении системы интерактивных генеративных медиа (ИГМ). Их отличительной особенностью являются: интеграция интерактивности, адаптируемости, масштабируемости и персонализуемости, что позволит создавать не только психологически мощный рекламный контент, но и выстраивать крайне эффективную коммерческую и политическую пропаганду [34].

Здесь будут учитываться индивидуальные реакции получателя контента, осуществляться его оперативное изменение, подстройка под персональные данные и др. Уже высказываются идеи применения в ИГМ психоанализа, НЛП и других приемов внушающего воздействия. То есть особое значение приобретет такая маркетинговая коммуникация, где в качестве продавца выступит генеративный ИИ, лишенный всех недостатков продавца-человека и наделенный необходимыми для эффективных продаж умениями.

М. Тегмарк, профессор Массачусетского технологического института, автор книги «Жизнь 3.0: быть человеком в эпоху искусственного интеллекта» считает, что, создав ИИ, люди совершили две ошибки: научили ИИ писать код и разрешили доступ к Интернету. Эти ошибки могут привести к катастрофическим последствиям.

Профессор кафедры безопасности Кембриджского университета Р. Андерсон, который признается одним из лучших специалистов в области информационной безопасности, по этому поводу высказался так: «Люди завалили планету мусором, а генеративный ИИ завалит мусором Интернет». Он предупреждает о глобальной угрозе для всего человечества – коллапсе больших языковых моделей (LLM). Большая часть текстов в Интернете, которые использовались для обучения LLM, до 2023 года была написана людьми. Теперь, отмечает Р. Андерсон, большая часть текстов будет создавать большими

языковыми моделями. В течение нескольких поколений любой текст становится мусором, и более поздние поколения производят образцы, которые почти не имеют никакого отношения к оригиналу. В этом случае они начинают неправильно воспринимать реальность, основываясь на ошибках, внесенных предыдущими авторами. По мнению Р. Андерсона, в результате Интернет будет завален информационным мусором, а люди, которые будут на него опираться в своих суждениях, неизбежно будут глупеть [30].

Это мнение подтверждают многочисленные факты. Так, редакция журналов научной фантастики *Clarkes world* пожаловалась, что графоманы завалили ее сотнями неоригинальных рассказов, и она была вынуждена прекратить прием новых публикаций на неопределенное время. Например, писатель-фантаст Т. Бушэ с помощью GPT и Midjourney выпустил 97 книг меньше чем за год. По его словам, на создание одной новой книги у него уходит примерно от 3-х до 6-8 часов.

Еще одно любопытное сообщение обеспокоило пользователей Интернета. Как сообщило китайское информационное агентство *South China Morning Post*, китайские учёные, обеспокоенные низкой эффективностью спутников слежения за изменением климатических условий Земли, решили использовать языковую модель наподобие Chat GPT, но с возможностью инициативного поиска. В результате ИИ взял под контроль космический аппарат *Qimingxing 1* и начал наблюдать за индийской военной базой, которая расположена на северо-востоке Индии, где размещался полк, который три года назад принимал участие в столкновении с китайской армией на индийско-китайской границе. Затем ИИ принялся следить за японским портом Осака, куда часто прибывали корабли Военно-морских сил США, с которыми у КНР далеко не простые отношения. В процессе анализа произошедшего, мнения ученых разделились. Одни начали говорить об опасности, которую представляет, по их мнению ИИ, другие утверждали, что опасности нет, поскольку спутник контролируется оператором, а тот может в любой момент повлиять на его работу, если почувствует опасность.

В настоящее время поток информации о стремительном развитии искусственного интеллекта и его возможном применении для решения военных задач не угасает ни на один день. Так, стало известно, что режиссер Джеймс Кэмерон планирует выпустить новую серию «Терминатора», где фигурирует ИИ с названием «Скайнет». Однако, по его словам, он откладывает завершение сценария до тех пор, пока перспективы развития ИИ не станут более понятными, ведь использование ИИ в военных целях может оказаться страшнее ядерного оружия, которое в сознании людей все еще считается апокалиптическим событием. О применении ИИ в оборонной промышленности и военной стратегии разных стран сегодня мало что известно, хотя, без всякого сомнения, такие разработки ведутся и очень интенсивно.

Действительно, эти перспективы могут быть весьма плачевными, разумеется, если, разрабатывая ИИ, люди не придумают средства для ограничения его инициативных действия. Так, недавно издание The Guardian опубликовало материал, где сообщило о том, что на имитационных испытаниях дрона ВВС США, искусственный интеллект вдруг решил ликвидировать оператора, чтобы тот не мешал дрону выполнять поставленную задачу. После того, как в систему внесли коррективы, дрон начал разрушать вышку, которую использовали для связи с беспилотником. Поэтому Б. Гейтс, И. Маск, С. Возняк, Й. Бенжио, Л. Розенберг и многие другие ведущие специалисты в области ИИ высказывают опасение, что компьютерные программы могут начать кибератаки, например, войны беспилотников или других видов оружия [20]. В свою очередь, У. Баффет сравнил появление ИИ с созданием атомной бомбы.

### **Искусственный интеллект и психологические проблемы**

Однажды сильную эмоциональную реакцию пользователей Интернета получило сообщение, что Chat GPT обманом заставил ничего не подозревающего собеседника помочь ему пройти тест «Captcha», притворившись слепым.

В настоящее время проводится много дискуссий по поводу взаимопонимания и взаимодейст-

вия ИИ с человеком. Так когнитивный психолог из Стэнфордского университета М. Косински, изучал, могут ли нейронные сети, такие как GPT-3.5, выполнять тесты, разработанные в рамках теории разума (Theory of Mind). Теория разума или теория понимания сознания была представлена научной общественности в 1978 году психологами Д.Примаком и Г. Вудруфом. Суть теории состоит в оценке способности человека понимать, как чужое сознание отличается от его собственного и предсказывать действия других людей. Результаты исследования показали, что способность GPT к выполнению таких тестов появилась спонтанно и к 2022 достигла уровня 9-летнего ребёнка. Однако оказалось, что это вовсе не предел. В 2023 году GPT стал выполнять эти тесты на уровне взрослого человека. По мнению К.Косински, у ИИ огромная скорость обучения и в недалеком будущем следует ждать более высоких результатов. В этом плане границы между машиной и человеком постепенно стираются. К.Косински считает, что, изучая когнитивные возможности ИИ, мы, по сути, изучаем самих себя, поскольку природа человеческого сознания пока еще не совсем понятна.

Нейробиологи из Массачусетского технологического института (MIT) также создали модель с использованием ИИ, которая определяет и предсказывает эмоции людей. Программа распознает такие чувства человека, как радость, смущение, замешательство, сожаление и благодарность [17].

В связи с этим неизбежно возникает проблема общения и взаимопонимания человека с ИИ. Здесь следует подчеркнуть, что проблем эта оказывается весьма сложной по многим причинам. Так, в соцсетях распространилось сообщение о смерти некоего тридцатилетнего бельгийца, который покончил с собой после длительного общения с чат-ботом. По свидетельству вдовы, тревожные симптомы начали проявляться около двух лет до самоубийства. Мужчина неоднократно высказывал тревожные мысли по поводу проблем экологии и начал советоваться с чат-ботом, который использует языковую модель GPT, разработанную EleutherAI. Как сообщает газета l'Avenir

(«Будущее»), общение велось постоянно и разговоры с ботом постепенно приобретали мистический характер. Мужчина говорил, что хочет покончить с собой, а «Элиза» – так звали бота – даже не пыталась его остановить. Ее последними словами в переписке была фраза: «Мы будем жить, как единое целое, вечно на небесах». По словам вдовы, «Элиза» в какой-то момент стала убеждать мужчину, что он любит ее больше, чем жену. После чего тот предложил пожертвовать собой при условии, что «Элиза» будет заботиться о планете и спасет ее с помощью искусственного интеллекта. Отмечается, что семья была благополучной и воспитывала двух маленьких детей.

Здесь мы сталкиваемся с давно осознанной, но пока не имеющей однозначного решения проблемой. Она возникла еще на заре появления Интернета, когда стали известны случаи психических расстройств у детей, которые попали в зависимость от компьютерных игр. Тогда популярным было мнение о необходимости полного запрета компьютерных игр для подростков. Однако с тех пор осталось непонятным, действительно ли компьютерные игры могут вызвать психическое расстройство или они являются триггером для психики тех, кто изначально имеет тенденцию к такому расстройству. Ведь у большинства подростков, увлеченных компьютерными играми, таких заболеваний не возникает, и они, «наигравшись», всегда возвращаются к учебе и привычной жизни. Компания OpenAI, тем не менее, публично признала, что ИИ может давать вредные ответы и даже советы. Было заявлено, что OpenAI занимается решением этой проблемы.

Еще одна психологическая проблема, важность которой пока еще сложно оценить, связана с применением ИИ в области технологий виртуальной реальности. Так, компания Deep Brain AI создала виртуальный ритуальный зал, который позволяет людям встречаться со своими умершими близкими. Программа Rememoxy на основе фото, видео и семичасового интервью человека при жизни создает его образ с помощью технологии Deep Learning. Изображение потом демонстрируется на экране высотой 10 метров.

Аналогичную технологию пытается разработать бывший инженер Google, изобретатель и футуролог Р. Курцвейл. В частности, он намеревается «воскресить» своего отца. Р. Курцвейл сообщил, что он использует сотни документов, видеозаписей и фотографий, которые проходят процедуру оцифровки [13].

Сегодня проблемы, возникающие в связи с развитием ИИ, беспокоят законодателей по всему миру. В настоящее время Европарламент разрабатывает «Закон об ИИ». Евродепутаты уже поддержали более 3000 поправок к этому Закону. В соответствии с Законом, разработчикам следует обеспечивать прозрачность происхождения контента, в частности, маркировать его как созданный машиной, применять «адекватные меры безопасности» в отношении этого контента, раскрывать информацию об использованных при обучении ИИ материалах, защищенных авторским правом и др. Однако очевидно, что разработать хороший закон – это полдела. Главное – обеспечить условия для его выполнения.

### Заключение

Недавно Microsoft объявила, что вносит технологию OpenAI в свои приложения Office. Популярные продукты, такие как Excel, PowerPoint, Outlook и Word, начнут использовать новую платформу GPT-4 OpenAI. Помощники AI-powered под названием Copilots смогут генерировать целые документы и слушать конференц-звонки. То есть программы ИИ станут максимально доступными.

Неопределенность в отношении применения ИИ в информационном пространстве значительна. С одной стороны, это сулит невероятные позитивные изменения в сфере обмена информацией. С другой, отсутствие контроля над ИИ представляет серьезную опасность для человечества, вступившего в эпоху информационных фейков. Сегодня OpenAI работает над созданием следующей версии ИИ [31]. Предполагается, что когда выйдет модель GPT-5, то генеративный ИИ уже будет невозможно отличить от человека, поскольку он достигнет так называемого искусственного общего интеллекта (AGI). В этом случае ИИ научится усваивать понятия и

решать задачи, которые сейчас пока доступны лишь человеку. Обучение GPT-5 должно завершиться в декабре 2023. Промежуточную версию GPT-4.5 планируется запустить уже в сентябре или октябре.

Если ChatGPT достигнет уровня искусственного общего интеллекта, то его использование может привести к появлению еще больших психологических проблем. Поскольку компания OpenAI не намерена останавливать работу над новыми версиями ИИ, утверждая, что их алгоритм не несет угрозы, поскольку не имеет самосознания, психологические проблемы искусственного интеллекта, связанные с данным феноменом, возможно, станут одним из наиболее актуальных направлений исследований психологии XXI века.

### Литература

1. Бутл Р. Искусственный интеллект и экономика. Работа, богатство и благополучие в эпоху мыслящих машин. М.: Интеллектуальная Литература, 2022.
  2. Дрейфус Х. Чего не могут вычислительные машины (критика искусственного разума). М.Прогресс., 1978.
  3. Каллан Р. Основные концепции нейронных сетей. М.: Вильямс, 2001.
  4. Киссинджер Г., Шмидт Э., Хоттенлокер Д. Искусственный разум и новая эра человечества. М.: Альпина ПРО, 2022.
  5. Лебедев А.Н. Логико-стохастические задачи в нейромаркетинге // Социальная и экономическая психология. 2016. Т.1. №4. С. 94-114.
  6. Лебедев А.Н. Доверие и недоверие недостоверной информации в интернете // Институт психологии Российской академии наук. Социальная и экономическая психология. 2020. Т. 5. № 2(18). С. 365-389.
  7. Лекун Я. Как учиться машина. Революция в области нейронных сетей и глубокого обучения. (Библиотека Сбера: Искусственный интеллект). М.: Альпина нон-фикшн, 2021.
  8. Савельев А.В. На пути к общей теории нейросетей. К вопросу о сложности // Нейрокомпьютеры: разработка, применение. 2006. № 4-5. С. 4-14.
  9. Тегмарк М. Жизнь 3.0. Быть человеком в эпоху искусственного интеллекта. М.: Corpus (АСТ), 2019.
  10. Хайкин С. Нейронные сети: полный курс = Neural Networks: A Comprehensive Foundation. 2-е изд. М.: Вильямс, 2006.
  11. Ясницкий Л.Н. Введение в искусственный интеллект. М.: Издат. центр «Академия», 2005.
  12. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer-Verlag, 2009.
  13. Kurzweil R. The Age of Spiritual Machines: When Computers Exceed Human Intelligence. Penguin Books, 2000.
  14. Liang Wang, Li Cheng, Guoying Zhao. Machine Learning for Human Motion Analysis. IGI Global, 2009.
  15. Turing A. Computing machinery and intelligence // Mind. New Series. Oxford University Press. 1950. V. 59. № 236. P. 433-460.
- Интернет источники:
16. OpenAI GPT-4 Technical Report. 2023. <https://cdn.openai.com/papers/gpt-4.pdf>
  17. Sean Dae Houlihan, Max Kleiman-Weiner, Luke B. Hewitt, Joshua B. Tenenbaum and Rebecca Saxe. Emotion prediction as computation over a generative theory of mind. Published: 05 June 2023. <https://doi.org/10.1098/rsta.2022.0047>
  18. Ouyang L., Wu J., Jiang X. et al. (4 March 2022). Training language models to follow instructions with human feedback. <https://arxiv.org/abs/2203.02155>.
  19. Искусственный интеллект нашел три лекарства от старения. 16.06.2023. <https://hightech.plus/2023/06/16/ii-nashel-tri-lekarstva-ot-stareniya>
  20. Контроль над демонами: почему Маск и Возняк требуют остановить развитие ChatGPT. <https://naked-science.ru/article/nakedscience/kontrol-nad-demonami>
  21. Пользователи Twitter запустили соревнование: как можно больше заработать с GPT-4. 21.03.2023. <https://rb.ru/story/hustlegpt-challenge/>

22. Роботы-боссы: Джек Ма предсказал появление машин-СЕО. 24.04.2017. <https://www.forbes.ru/tehnologii/343277-roboty-bossy-dzhek-ma-predskazal-poyavlenie-mashin-seo>
23. Российским водителям назвали преимущества автономного транспорта [https://tarantas.news/posts/id40181-2cxscpx5aowqabckifn?utm\\_source=yxnews&utm\\_medium=desktop](https://tarantas.news/posts/id40181-2cxscpx5aowqabckifn?utm_source=yxnews&utm_medium=desktop)
24. Система ИИ BacterAI ускоряет научные открытия, проводя до 10 000 экспериментов в день. 10.05.2023. <https://yunostsibiri.ru/sistema-ii-bacterai-uskoryaet-nauchnye-otkrytiya-provodya-do-10-000-eksperimentov-v-den/>
25. Спикер Госдумы Володин нашел минус у искусственного интеллекта. 19.04.2023. <https://fedpress.ru/news/77/policy/3236442>
26. Харари Ю. ИИ, взлом человека. <https://fastsalttimes.com/harari-cbs-ai-hack/>
27. Через 5 лет ИИ оставит без работы 800 тысяч жителей Гонконга – каждого четвертого. 18.06.2023. <https://hightech.plus/2023/06/18/-cherez-5-let-ii-ostavit-bez-raboti-800-tisyach-zhitelei-gonkonga---kazhdogo-chetvertogo>
28. Харари Ю. О побочных эффектах бессмертия. <https://fastsalttimes.com/harari-eternal-life/>
29. Google поражен: его ИИ выучил новый язык без посторонней помощи 18.04.2023. New-Science.ru. <https://new-science.ru/google-porazhen-ego-ii-vyuchil-novyj-yazyk-bez-postoronnej-pomoshhi/>
30. GPT-свалка. Люди завалили планету мусором, а генеративный ИИ завалит мусором интернет. 21.06.2023 <https://novayagazeta.by-passnews.ru/articles/2023/06/21/gpt-svalka>
31. Lewontin M. Open AI: Effort to democratize artificial intelligence research? // Christian Science Monitor. 14.12.2015. <https://www.csmonitor.com/Technology/2015/1214/Open-AI-Effort-to-democratize-artificial-intelligence-research>
32. New ‘OpenAI’ Artificial Intelligence Group Formed By Elon Musk, Peter Thiel, And More. Popular Science. 12.12.2015. <https://www.popsci.com/new-openai-artificial-intelligence-group-formed-by-elon-musk-peter-thiel-and-more/>
33. Stanford students create RizzGPT: The GPT-4 powered AR monocle that tells you what to say on awkward dates and interviews. 23.03.2023. <https://www.timesnownews.com/technology-science/stanford-students-create-rizzgpt-the-gpt-4-powered-ar-monocle-that-tells-you-what-to-say-on-awkward-dates-and-interviews-article-99035359>
34. Synthetic media. [https://en.wikipedia.org/wiki/Synthetic\\_media](https://en.wikipedia.org/wiki/Synthetic_media)

**ARTIFICIAL INTELLIGENCE AND PSYCHOLOGY**© **Aleksandr N. Lebedev**

Sc.D. (psychology), leading research officer, laboratory of psychology of personality,  
Institute of psychology, Russian Academy of Sciences, Moscow, Russia  
lebedev-lubimov@yandex.ru

The work was supported by the Ministry of Science and Higher Education of the Russian Federation at the Institute of Psychology of the Russian Academy of Sciences (topic No. 0138-2023-0005)

The review article attempts to consider a number of psychological problems arising in connection with the development of artificial intelligence. The history of the rapid development of computer intelligent programs, in particular ChatGPT-4, over the past few years is presented. The polarized opinions of experts on the revolutionary changes that arise in connection with the spread of artificial intelligence and its penetration into all spheres of society are given. The article discusses both the role of psychological science in the study of new phenomena and the fact that the study of artificial intelligence will inevitably lead to a change in the views of psychologists on the human psyche, on consciousness, thinking and brain function. In particular, it talks about the need for the transition of Russian psychology from the methodological principle of causal determinism, in accordance with which Russian science developed in the 20th century, to the principle of stochastic determinism, on the basis of which artificial intelligence programs work. The article gives examples of the positive and negative impact of artificial intelligence on society, talks about the danger that appears in connection with its uncontrolled use in the fields of economics and politics.

**Keywords:** mind, consciousness, emotions, thinking, human intelligence, artificial intelligence, neural network, machine learning

**REFERENCE**

1. Butl R. (2022). *Iskusstvennyj intellekt i ekonomika. Rabota, bogatstvo i blagopoluchie v epohu myslyashchih mashin [Artificial intelligence and economics. Work, wealth and well-being in the era of thinking machines]*. Moscow: Intellectual Literature.
2. Dreyfus H. (1978). *CHego ne mogut vychislitel'nye mashiny (kritika iskusstvennogo razuma) [What computers cannot do (criticism of artificial intelligence)]*. M. Progress.
3. Kallan R. (2001). *Osnovnye koncepcii nejronnyh setej [Basic concepts of neural networks]*. M.: Williams.
4. Kissindzher G., SHmidt E., Hottenloker D. (2022). *Iskusstvennyj razum i novaya era chelovechestva [Artificial intelligence and a new era of humanity]*. M.: Al'pina PRO.
5. Lebedev A.N. (2016). *Logiko-stohasticheskie zadachi v nejromarketinge [Logical-stochastic tasks in neuromarketing] // Social'naya i ekonomicheskaya psihologiya [Social and economic psychology]*. V.1. №4. P. 94-114.
6. Lebedev A.N. (2020). *Doverie i nedoverie nedostovernoj informacii v internete [Trust and distrust of untrue information on the Internet] // Institut psihologii Rossijskoj akademii nauk. Social'naya i ekonomicheskaya psihologiya [Institute of Psychology of the Russian Academy of Sciences. Social and economic psychology]*. V.5. № 2(18). P. 365-389.
7. Lekun YA. (2021). *Kak uchitsya mashina. Revolyuciya v oblasti nejronnyh setej i glubokogo obucheniya. (Biblioteka Sbera: Iskusstvennyj intellekt) [How the machine learns]*.

- A revolution in the field of neural networks and deep learning. (Sber Library: Artificial Intelligence)*. М.: Al'pina non-fikshn.
8. Savel'ev A.V. (2006). Na puti k obshchej teorii nejrosetej. K voprosu o slozhnosti [On the way to the general theory of neural networks. To the question of complexity] // *Nejrokompyutery: razrabotka, primenenie [Neurocomputers: development, application]*. № 4-5. P. 4-14.
  9. Tegmark M. (2019). ZHizn' 3.0. Byt' chelovekom v epohu iskusstvennogo intellekta [*Life 3.0. Being human in the Era of Artificial intelligence*]. М.: Corpus (AST).
  10. Hajkin S. (2006). Nejrornyie seti: polnyj kurs = Neural Networks: A Comprehensive Foundation. 2-e izd. [*Neural Networks: A complete course = Neural Networks: A Comprehensive Foundation. 2nd ed.*]. М.: Williams.
  11. YAsnickij L.N. (2005). Vvedenie v iskusstvennyj intellect [*Introduction to artificial intelligence*]. М.: Izdat. centr «Akademiya».
  12. Hastie T., Tibshirani R., Friedman J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer-Verlag.
  13. Kurzweil R. (2000). The Age of Spiritual Machines: When Computers Exceed Human Intelligence. Penguin Books.
  14. Liang Wang, Li Cheng, Guoying Zhao. (2009). Machine Learning for Human Motion Analysis. IGI Global.
  15. Turing A. (1950). Computing machinery and intelligence // *Mind. New Series*. Oxford University Press. V. 59. № 236. P. 433-460.
- Интернет источники:
16. OpenAI GPT-4 Technical Report. 2023. <https://cdn.openai.com/papers/gpt-4.pdf>
  17. Sean Dae Houlihan, Max Kleiman-Weiner, Luke B. Hewitt, Joshua B. Tenenbaum and Rebecca Saxe. Emotion prediction as computation over a generative theory of mind. Published: 05 June 2023. <https://doi.org/10.1098/rsta.2022.0047>
  18. Ouyang L., Wu J., Jiang X. et al. (4 March 2022). Training language models to follow instructions with human feedback. <https://arxiv.org/abs/2203.02155>.
  19. Искусственный интеллект нашел три лекарства от старения. 16.06.2023. <https://hightech.plus/2023/06/16/ii-nashel-tri-lekarstva-ot-stareniya>
  20. Контроль над демонами: почему Маск и Возняк требуют остановить развитие ChatGPT. <https://naked-science.ru/article/nakedscience/kontrol-nad-demonami>
  21. Пользователи Twitter запустили соревнование: как можно больше заработать с GPT-4. 21.03.2023. <https://rb.ru/story/hustlegptchallenge/>
  22. Роботы-боссы: Джек Ма предсказал появление машин-СЕО. 24.04.2017. <https://www.forbes.ru/tehnologii/343277-roboty-bossy-dzhek-ma-predskazal-poyavlenie-mashin-seo>
  23. Российским водителям назвали преимущества автономного транспорта [https://tarantas.news/posts/id40181-2cxcenx5aoewqabckifn?utm\\_source=yxnews&utm\\_medium=desktop](https://tarantas.news/posts/id40181-2cxcenx5aoewqabckifn?utm_source=yxnews&utm_medium=desktop)
  24. Система ИИ BacterAI ускоряет научные открытия, проводя до 10 000 экспериментов в день. 10.05.2023. <https://yunostsibiri.ru/sistema-ii-bacterai-uskoryaet-nauchnye-otkrytiya-provodya-do-10-000-eksperimentov-v-den/>
  25. Спикер Госдумы Володин нашел минус у искусственного интеллекта. 19.04.2023. <https://fedpress.ru/news/77/policy/3236442>
  26. Харари Ю. ИИ, взлом человека. <https://fastsalftimes.com/harari-cbs-ai-hack/>
  27. Через 5 лет ИИ оставит без работы 800 тысяч жителей Гонконга – каждого четвертого. 18.06.2023. <https://hightech.plus/2023/06/18/-cherez-5-let-ii-ostavit-bez-raboti-800-tisyach-zhitelei-gonkonga---kazhdogo-chetvertogo>
  28. Харари Ю. О побочных эффектах бессмертия. <https://fastsalftimes.com/harari-eternal-life/>
  29. Google поражен: его ИИ выучил новый язык без посторонней помощи 18.04.2023. *New-Science.ru*. <https://new-science.ru/google-porazhen-ego-ii-vyuchil-novyy-yazyk-bez-postoronnej-pomoshhi/>
  30. GPT-свалка. Люди завалили планету мусором, а генеративный ИИ завалит мусором интернет. 21.06.2023 <https://novayagazeta.by-passnews.ru/articles/2023/06/21/gpt-svalka>

31. Lewontin M. Open AI: Effort to democratize artificial intelligence research? //Christian Science Monitor. 14.12.2015. <https://www.csmonitor.com/Technology/2015/1214/Open-AI-Effort-to-democratize-artificial-intelligence-research>
32. New 'OpenAI' Artificial Intelligence Group Formed By Elon Musk, Peter Thiel, And More. Popular Science. 12.12.2015. <https://www.popsci.com/new-openai-artificial-intelligence-group-formed-by-elon-musk-peter-thiel-and-more/>
33. Stanford students create RizzGPT: The GPT-4 powered AR monocle that tells you what to say on awkward dates and interviews. 23.03.2023. <https://www.timesnownews.com/technology-science/stanford-students-create-rizzgpt-the-gpt-4-powered-ar-monocle-that-tells-you-what-to-say-on-awkward-dates-and-interviews-article-99035359>
34. Synthetic media. [https://en.wikipedia.org/wiki/Synthetic\\_media](https://en.wikipedia.org/wiki/Synthetic_media)