

DOI: 10.31696/S086919080033177-1

ЗЛОНАМЕРЕННОЕ ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И УГРОЗЫ ИНФОРМАЦИОННО-ПСИХОЛОГИЧЕСКОЙ БЕЗОПАСНОСТИ ДЛЯ ИРАНА: МНОГОУРОВНЕВАЯ РЕАЛЬНОСТЬ ¹

© 2025

Е.Н. ПАШЕНЦЕВ ^{a, b}, П.В. КУЗНЕЦОВ ^b, В.А. ЧЕБЫКИНА ^{a, c}

^a – Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия

^b – Дипломатическая академия МИД России, Москва, Россия
ORCID: 0000-0001-5487-4457; icspsc@mail.ru

^b – ORCID: 0009-0007-5473-798X; pvkuz@yandex.ru

^c – Международный центр социально-политических исследований и консалтинга, Москва,
Россия

ORCID: 0009-0003-5031-3339; sj.vvladilena@gmail.com

Резюме: В условиях острых международных противоречий, социально-политической нестабильности и конфликтов злонамеренное использование всё более совершенных технологий искусственного интеллекта (ИИ) представляет растущую угрозу информационно-психологической безопасности общества и, как следствие, для политических, социальных, экономических, культурных процессов и деятельности государственных и негосударственных институтов. Не является исключением и Иран, где научно-технологическое развитие тесно переплетается с вопросами политической стабильности и внешнего давления, а риски, связанные со злонамеренным использованием ИИ (ЗИИИ), приобретают особую значимость. Угрозы информационно-психологической безопасности посредством ЗИИИ рассматриваются авторами сквозь призму трёхуровневой классификации: от распространения ложных представлений об ИИ и последствиях его развития, угроз критической инфраструктуре, физическому здоровью и жизни людей до информационно-психологических операций, представляющих серьёзную опасность для Ирана. Данная классификация позволяет структурировать риски и выделить направления, требующие приоритетного внимания со стороны государственных институтов и экспертов в области безопасности. Особое внимание уделяется применению алгоритмов машинного обучения и систем обработки естественного языка для создания фальшивых сообщений и изображений, которые могут быть использованы в антиправительственной пропаганде. Также анализируется злонамеренное воздействие посредством технологий ИИ на религиозные и культурные аспекты жизни иранского общества, что особенно актуально для государства с традиционными ценностями. Подчёркивается, что злонамеренное воздействие на традиционные институты может привести к подрыву общественного равновесия, что несёт риски для долгосрочной стабильности государства. Отмечаются примеры конкретных угроз и уязвимостей Ирана в условиях международной изоляции и экономических санкций. В этом контексте необходимо учитывать динамическое развитие технологий и интеграцию ИИ в глобальную инфраструктуру, что требует создания устойчивой системы мониторинга и противодействия угрозам. Это позволит не только минимизировать риски, но и укрепить долгосрочную стабильность государства в условиях стремительно меняющихся геополитических и технологических реалий.

¹ Работа Е.Н. Пашенцева выполнена при поддержке СПбГУ, шифр проекта 116471555.

Ключевые слова: искусственный интеллект, злонамеренное использование искусственного интеллекта, кибербезопасность, международная информационно-психологическая безопасность, Иран, дипфейк

Для цитирования: Пашенцев Е.Н., Кузнецов П.В., Чебыкина В.А. Злонамеренное использование искусственного интеллекта и угрозы информационно-психологической безопасности для Ирана: многоуровневая реальность. *Восток (Oriens)*. 2025. № 3. С. 125–136. DOI: 10.31696/S086919080033177-1

MALICIOUS USE OF ARTIFICIAL INTELLIGENCE AND THREATS TO IRAN'S PSYCHOLOGICAL SECURITY: A MULTI-LEVEL REALITY ²

© 2025 Evgeny N. PASHENTSEV ^{a, b}, Pavel V. KUZNETSOV ^b, Vladilena A. CHEBYKINA ^{a, c}

^a – Saint Petersburg State University, Saint Petersburg, Russia

^b – Diplomatic Academy of the Russian Foreign Ministry, Moscow, Russia
ORCID: 0000-0001-5487-4457; icspsc@mail.ru

^b – Diplomatic Academy of the Russian Foreign Ministry, Moscow, Russia
– ORCID: 0009-0007-5473-798X; pvkuz@yandex.ru

^c – International Centre for Social and Political Studies and Consulting,
Moscow, Russia
ORCID: 0009-0003-5031-3339; sj.vvladilena@gmail.com

Abstract: *In the context of international tensions, socio-political instability and conflicts, the malicious use of artificial intelligence (AI) technologies poses a mounting threat to the psychological security of society and, as a consequence, to political, social, economic, cultural processes and functioning of state and non-state institutions. Iran is no exemption. Scientific and technological development is intertwined with the issues of political stability and external pressure, and the risks associated with the malicious use of AI (MUAI) are of particular importance. MUAI is understood here as intentionally antisocial activity of state and non-state actors. The authors consider the threats of MUAI to psychological security through the prism of a three-level classification: from the spread of misconceptions about AI and the consequences of its development, threats to the critical infrastructure, physical health and human life, to psychological operations that pose a serious risk to Iran. This classification allows structuring the risks and highlighting key trends that require priority attention from government institutions and security experts. Special attention is paid to the use of machine learning algorithms and natural language processing systems to create fake messages and images that can be used in anti-government propaganda. Moreover, the authors analyze the impact of malicious use of AI technologies on religious and cultural aspects of Iranian society, especially relevant for a state based on traditional values. The authors emphasize that the malicious impact on traditional institutions can lead to the undermining of social balance, thereby posing risks to the long-term stability of the state.*

Keywords: artificial intelligence, malicious use of artificial intelligence, cybersecurity, international psychological security, Iran, deepfake technologies

² E.N. Pashentsev acknowledges Saint-Petersburg State University for a research project 116471555.

For citation: Pashentsev E.N., Kuznetsov P.V., Chebykina V.A. Malicious Use of Artificial Intelligence and Threats to Iran's Psychological Security: A Multi-Level Reality. *Vostok (Oriens)*. 2025. No. 3. Pp. 125–136. DOI: 10.31696/S086919080033177-1

ВВЕДЕНИЕ

Несмотря на более чем четыре десятилетия масштабных санкций со стороны стран Запада, Иран занимает 13-е место в рейтинге Nature Index по уровню развития технологий искусственного интеллекта (ИИ) [Iran Plans, 30.01.2022]. Рейтинговая база Scimago Университета Гранады, Испания, ранжирует страны и университеты на основе количества научных статей, опубликованных в базе данных Scopus. По рейтингу Scimago в области ИИ Иран занимал на 2024 г. 15-е место в мире и первое в Западной Азии [The SCImago Journal & Country Rank, 2024]. Однако быстрое развитие и внедрение технологий ИИ сопровождается растущим злонамеренным использованием ИИ (ЗИИИ). Под ЗИИИ авторами понимается намеренно антисоциальное применение ИИ государственными и негосударственными акторами [Pashentsev, 2023].

На фоне социального и имущественного расслоения угрозы ЗИИИ в Иране возрастают [Zadeh, Haghsheenas, 2024; Tajmazinani, 2021], усугубляясь этническими и политическими разногласиями и конфликтами [Sheikhtaheri, Firouzabadi, 2024], проявлениями коррупции [Iran International, 2023; Zamahani, 2022].

Страна при этом находится под внешним давлением со стороны, прежде всего, Соединённых Штатов Америки и Израиля. В отчёте разведывательного сообщества США от 5 февраля 2024 г., Иран вошёл в число стран, представляющих наибольшую угрозу национальной безопасности США и международной стабильности. Израиль в отчёте также упоминается, очевидно, как партнёр США на Ближнем Востоке. [Office of the Director of National Intelligence, 2024].

Непосредственно Израиль видит в Иране наиболее значимую угрозу себе и своему положению на Ближнем Востоке, практически находясь с ним в состоянии необъявленной войны. При этом и США и Израиль обладают собственными претензиями на эксклюзивное лидерство в сфере ИИ, а значит, заинтересованностью как в подрыве развития ИИ в Иране, так и в применении соответствующих технологий против этой страны [Пашенцев, Кузнецов, 2024].

В своей работе авторы опирались на официальные заявления представителей государственной власти Ирана, а также на экспертные доклады, посвящённые кибербезопасности, рискам ИИ и ЗИИИ. Исследование основано на широком круге академических трудов, затрагивающих социально-политическое развитие Ирана [Bolgov, 2023; Hajialiasgari, Khanahmadi, Atashi, 2023], роль иранских СМИ и интернет-политики [May, 2023; Conduit, 2022], роль ИИ в информационном противоборстве в современном мире [Карасев и др., 2021; Hunter et al., 2024; Reinhold, Reuter, 2024; Pauwels, Deton, 2020], а также ряд работ по проблемам ЗИИИ и информационно-психологической безопасности (ИПБ) [Howard, 2020; Pashentsev, 2023; Mantello et al., 2023; Vacarelu, 2023].

Существенную помощь при раскрытии темы имела информационно-аналитическая разработка авторов настоящей публикации [Пашенцев, Кузнецов, 2024]. Однако ни по своему характеру (оперативное выявление актуальной проблемы и ее текущее освещение на основе довольно узкого круга первичных источников), ни по основному содержанию, ни по объёму она не являлась системным научным исследованием избранной для анализа темы настоящей статьи. Между тем, за прошедший год произошло дальнейшее заметное

развитие технологий ИИ, быстрый рост ЗИИИ и соответствующих угроз ИПБ Ирана, что определяет возросшую актуальность рассмотрения темы.

Для анализа угроз ИПБ, порождаемых ЗИИИ, авторами применялась трёхступенчатая модель, предложенная Е.Н. Пашенцевым. *Первый уровень угроз* характеризуется распространением заведомо искажённой, ложной информации о сути и возможностях ИИ, последствий его развития и применения. *Второй уровень* связан с практикой ЗИИИ, когда атака на общественное сознание не является основной целью, но имеет негативные для него последствия, в результате ущерба наносимого критической инфраструктуре, жизни и здоровью людей и их имуществу. *Третий* и самый высокий *уровень угроз ИПБ* заключается в целенаправленной дезориентации общественного мнения и ухудшения психологического состояния населения посредством ЗИИИ (подробнее см.: [Pashentsev, 2023]). Данная классификация позволяет чётко структурировать угрозы и оценить их влияние на общество и государственные институты.

ПЕРВЫЙ УРОВЕНЬ УГРОЗ

Иранское общество встревожено деструктивным влиянием, которое ИИ может оказать на рынок труда – в частности, растёт обеспокоенность возможными массовыми увольнениями с заменой людей на роботов, управляемых технологиями ИИ. Представители научных кругов и профессий, связанных с грубым физическим трудом, могут чувствовать себя в безопасности. В группе риска, в первую очередь – специалисты среднего уровня квалификации. Данные опасения также могут целенаправленно усиливаться внешними акторами в целях подрыва развития технологий в стране.

На сегодняшний день общая ситуация с занятостью в Иране является не простой, но контролируемой правительством, более того, наблюдаются успехи в преодолении безработицы. Статистический центр Ирана в июне 2024 г. опубликовал отчёт, согласно которому по состоянию на 2024 г. уровень безработицы в стране является самым низким за последние 26 лет – 8,1% [Iranian Labor, 2024], однако среди молодёжи он остаётся высоким (19,4%) [Iran Jobless, 2024]; молодёжь, не получившая надлежащего образования или доступа к технологиям ИИ, рискует остаться за пределами трудового рынка, что может усилить социальное неравенство и также стимулировать попытки вовлечь молодёжь в акции протеста против автоматизации и технологий ИИ.

Использование Ираном ИИ для обеспечения общественного порядка вызывает дискуссии в странах Запада. Одним из наиболее спорных моментов стала информация о внедрении ИИ в городские системы видеонаблюдения Ирана, направленном на автоматическое выявление женщин, нарушающих правила ношения хиджаба [Mahmood, Sahraei, 2023]. Это приводило к массовым протестам и беспорядкам, в том числе с жертвами [Report of the Independent, 2024].

Иранские власти, в свою очередь, рассматривают ИИ как инструмент, который может обеспечить соблюдение общественного порядка в больших масштабах и с минимальными ресурсами. Западные СМИ и эксперты, акцентируя внимание на протестах и жертвах в связи с использованием технологий ИИ силами правопорядка [George, 2023], не рассматривают другие важные аспекты, влияющие на внутреннюю ситуацию в Иране. Например, активная деятельность террористических группировок, включая «Аль-Каиду»³, часто поддерживаемых внешними силами, остаётся в тени обсуждений, хотя именно задачами

³ Запрещена в РФ.

противодействия терроризма⁴ прежде всего может объясняться внедрение ИИ в городские системы видеонаблюдения.

Наблюдаются попытки создать искажённое представление об использовании ИИ Ираном за пределами его национальной территории. Так, в октябре 2024 г. Министерство внутренней безопасности США в своей ежегодной оценке рисков для страны заявило, что Россия, Китай и Иран используют ИИ в открытых и скрытых информационных операциях с целью воздействовать на результаты выборов и подорвать стабильность США [*Homeland Threat Assessment 2024*, р. 6–7], что получило официальное опровержение со стороны МИД Ирана [*IRNA*, 2024(1)]. Сразу после выборов, как следовало из заголовка материала, опубликованного The Associated Press 9 ноября 2024 г., Министерство юстиции США выдвинуло уголовные обвинения в связи с планом иранских наёмных убийц, нацеленных на Дональда Трампа [Tucker, Neumeister, 2024]. Таким образом, происходит связывание в едином информационном потоке многочисленных необоснованных заявлений о высокотехнологичном вмешательстве Ирана в выборы с не более обоснованными обвинениями о причастности Ирана к подготовке убийства только что победившего на президентских выборах, но ещё не вступившего в должность Д. Трампа.

Поддерживают данный нарратив и частные западные компании. Так, 30 января 2025 года подразделение анализа угроз Google опубликовало отчёт о злонамеренном использовании различными акторами большой языковой модели Google Gemini. Набор «обвиняемых» обычен – упоминаются Иран, Россия, Китай и КНДР. При этом Иран занимает первое место, и о нём сообщается, что иранские проправительственные атакующие группы явились самыми активными пользователями Gemini, применяя модель для самых разных целей, включая сбор разведанных в отношении оборонных предприятий, исследование уязвимостей и создание контента для информационных операций [Adversarial Misuse, 2025].

Умышленно односторонний, упрощённый, искажённый дискурс использования технологий ИИ в Иране, многократно повторяемый и широко распространяемый западными СМИ, является конкретным примером информационно-психологической операции на первом уровне угроз ИПБ посредством ЗИИИ. Промедление с использованием внедрения технологий ИИ может ослабить силы правопорядка в противодействии враждебным государственным акторам, террористам, другим преступным организациям (которые эти технологии в разных странах небезуспешно осваивают); а в более широком плане – позволить им сформировать менее благоприятное восприятие ИИ в иранском обществе, вызвать недоверие к использованию этих технологий властями Ирана, что может замедлить развитие страны.

ВТОРОЙ УРОВЕНЬ УГРОЗ

Серьёзные угрозы кибербезопасности Ирану стали очевидны ещё в 2010 г., когда был обнаружен Stuxnet – компьютерный червь, использованный в качестве средства несанкционированного сбора данных и диверсий и разработанный разведками США и Израиля для вывода из строя ключевой части иранской ядерной программы. В 2011 г. была создана киберполиция Ирана для борьбы с фишингом, подделкой документов, хакерством, организованной интернет-преступностью, порнографией. Киберполиция также была призвана защищать критическую инфраструктуру от электронных атак, поддерживая общественный порядок [Golkar, 2018].

⁴ 3 января 2024 г. в Кермане на мероприятии памяти генерал-лейтенанта Касема Сулеймани произошел крупный теракт, в результате которого погибли 93 человека, ещё несколько десятков получили ранения.

Угрозы второго уровня для Ирана постоянно подтверждают свою реальную опасность. В декабре 2023 г. начальник управления гражданской обороны генерал Г. Джалали заявил, что произошедший массовый сбой на заправочных станциях в стране был кибератакой с применением вредоносного программного обеспечения, и более 30% подобных кибератак на Иран осуществляются с использованием ИИ. Также уместно вспомнить, что 3 января 2020 г. в результате удара американского беспилотника, управляемого ИИ, возле международного аэропорта Багдада был убит генерал К. Сулеймани. Информация о нападении транслировалась в режиме прямого времени в Белый дом и ЦРУ [Пашенцев, Кузнецов, 2024].

Такая зрелищность, публичность «наказания» противника США и Израиля с помощью передовых технологий могла иметь как кратковременные, так и долговременные эффекты воздействия. Если в США и Израиле специальная операция среди большей части общественности скорее имела поддержку, то в Иране была воспринята как неприкрытое публичное унижение страны и вызвала у многих гнев и решимость противостоять США.

По сообщению иранских властей, ИИ, как предполагается, используется и в террористических нападениях, причём выдвигаются версии, что заинтересованной стороной может выступать Израиль [Motamedi, 2020].

По сообщению информационного агентства Tasnim, в ходе израильской операции в Секторе Газа происходит активное применение технологий ИИ. Разведывательное подразделение армии Израиля применяло и применяет в Газе технологии ИИ в целях боевого планирования, расчёта «допустимого» количества жертв среди гражданского населения при бомбовых ударах и в иных целях обработки больших массивов информации и управления ударами [The first destructive use, 2025; Israel Using AI, 2025]. Израильские официальные лица также делали заявления, свидетельствующие о намерении страны развивать военное применение ИИ в том числе в целях разведки против Ирана [Berman, 2023].

Эффекты использования ИИ на втором уровне угроз ИПБ могут иметь характер опережающего воздействия (опасения, тревога, страх перед атакой и её последствиями, например, в результате опасного воздействия на ядерные объекты страны), воздействия одновременного с атакой (страх, паника, неверный алгоритм ответных действий из-за неожиданности и/или силы воздействия) и разнообразные кратковременные и долговременные последствия ЗИИИ для ИПБ, вплоть до необратимых последствий для психики и здоровья людей.

ТРЕТИЙ УРОВЕНЬ УГРОЗ

Иранские СМИ сегодня претерпевают значительные изменения, вызванные быстрой цифровизацией, что открывает новые возможности для более мощных и разрушительных информационных атак, в том числе с использованием генеративных ИИ-моделей. Одним из примеров быстро развивающихся областей разработки ИИ являются веб-роботы, чат-боты и социальные боты [Bradshaw, Howard, 2019; Jhan et al., 2021; Marcellino et al., 2020]. Такие технологии несут не только пользу [Marcellino et al., 2020; Howard, 2020]. Злонамеренные акторы используют их для распространения дезинформации [Mustafaraj, Metaxas, 2017], создания фейковых аккаунтов в социальных сетях [Ramalingam, Chinnaiiah, 2018], сбора личных данных пользователей [Deibert, 2019], а также подрыва ИПБ [Pashentsev, Bazarkina, 2022]. Учитывая, что новые технологии формируют цифровые социальные отношения, которые воспринимаются людьми как не менее полноценные, чем традиционные, некоторые авторы предполагают, что боты будут играть всё более важную роль в радикализации экстремистских движений [Mantello et al., 2023]. В руках экстремистских группировок боты будут заменять живых пропагандистов, вербовщиков и армии троллей, что является одной из возможных угроз ИПБ Ирана.

Одним из веских доказательств злонамеренного использования ИИ стали «Файлы Twitter» – серия публикаций отдельных внутренних документов Twitter, Inc., опубликованных в Twitter с декабря 2022 г. по март 2023 г. И. Маск передал документы журналистам вскоре после того как он приобрел Twitter 27 октября 2022 г. Журналисты координировали публикацию документов с И. Маском, опубликовав подробную информацию о файлах в серии сообщений в Twitter. В частности, согласно этим документам, правительство США создало множество поддельных аккаунтов в Twitter, предназначенных для распространения проамериканских настроений на Ближнем Востоке, в России и Китае. Притом, хотя в Twitter действует чёткая политика, запрещающая правительствам делать подобное, официальные лица компании, похоже, знали, что это был план Пентагона [Loo, Khalil, 2022].

В 2017 г. представитель Центрального командования ВС США (CENTCOM) отправил в Twitter список из 52 аккаунтов на арабском языке, которые использовались «...для усиления определённых сообщений». Чиновник запросил приоритетное обслуживание для шести учётных записей и возможность внесения в «белый список», что вело к большей популярности этих аккаунтов. Они часто публиковали в Twitter информацию о военных приоритетах США на Ближнем Востоке, включая пропаганду антииранских идей и войны в Йемене, поддерживаемой Саудовской Аравией и США, и «точные» удары американских беспилотников, которые, как утверждалось, наносились только по террористам. Они обвиняли Иран в «угрозе водной безопасности Ирака и наводнении страны метамфетаминном», а также в извечении органов у афганских беженцев [McPherrin, 2023].

Согласно сообщениям в СМИ, CENTCOM активно проводил информационно-психологические операции на платформах социальных сетей. Направленные на подрыв позиций Ирана, они включали систематическое распространение пропагандистских материалов [Pentagon, 25.12.2022]. ИИ использовался для создания текстового контента и фотореалистичных дипфейков – изображений вымышленных пользователей, от имени которых публиковались твиты. Подобная тактика позволяла придать материалам видимость достоверности и усилить их влияние на аудиторию.

20 апреля 2024 г. доктор А. Маранди, глава Академии медицинских наук и врач аятолы Хаменеи, появился в программе известного иранского журналиста К. Наджафзадеи и рассказал о состоянии здоровья Верховного Лидера иранской революции. В июне на одном из каналов Telegram была опубликована аудиозапись, которая распространялась также в социальных медиа Tik Tok, YouTube и X. На данной аудиозаписи в качестве фона использовалось изображение со страницы К. Наджафзадеи в Instagram⁵ из вышеупомянутой программы, а голоса Маранди и Наджафзадеи были воссозданы с помощью технологии дипфейк. В аудиофайле голос якобы доктора Маранди заявляет, что «медицинские усилия будут гарантировать здоровье руководства страны до этой зимы, а после этого, даст Бог, молитвы людей будут целительными». Такая формулировка намекает слушателям, что аятолла Хаменеи страдает серьёзным заболеванием [IRNA, 2024(2)].

Впоследствии данная информационная акция получила определённое развитие. В различных источниках появлялись утверждения о том, что ведущий программы через несколько часов удалил этот пост по приказу властей, а Наджафзадеи обязали опровергнуть опубликованный им аудиофайл и назвать слухами озвученную информацию. Позже был обнародован якобы «подтверждённый» слух об аресте Наджафзадеи, направленный на подкрепление «правдивости» как первоначального поста, так и последовавших за ним сообщений. С опровержением впоследствии выступила Академия медицинских наук Ирана,

⁵ Организация Meta Platforms Inc. (соцсети Facebook, Instagram) признана экстремистской, ее деятельность запрещена на территории России.

объявившая вышеприведенные «новости» фейковыми [IRNA, 2024(2)]. Можно предположить, что распространялись они с целью дестабилизации обстановки в Иране.

Другим примером ЗИИИ служит чат-бот Grok, созданный на базе ИИ, который стал причиной распространения фейковых новостей на платформе «X». Так, в 2024 г. пользователи данной социальной сети увидели заголовок, в котором утверждалось, что Иран нанёс ракетный удар по Тель-Авиву [Elon Musk, 22.04.2024]. Такие эмоционально заряженные новости распространяются быстро, вопрос в том, кто несёт за это ответственность. Развитие технологий, зачастую, движется быстрее анализа возможных рисков – и в Иране это видно на примере исследования, проведённого в Тегеранском университете медицинских наук. Авторы приходят к скорее положительному выводу о необходимости внедрения ИИ в здравоохранении для обработки больших данных и в виде чат-ботов. При этом из рисков рассматриваются лишь очевидные, многократно упоминающиеся в дискуссиях – такие, как возможное разглашение врачебной тайны или психологический дискомфорт от общения с роботами [Hajialiasgari, Khanahmadi, Atashi, 2023].

Другой чувствительной сферой является внедрение ИИ в религиозную практику, особенно в Иране, где религия тесно переплетается с политикой. Религиозные чат-боты, запрограммированные для предоставления ответов на вопросы о религии, используются в образовательных целях, в миссионерской работе, а также для повседневных консультаций по религиозным вопросам. В 2023 г. Мохаммад Готби, глава Дома творчества и инноваций «Эшраг», отметил, что, хотя ИИ не способен заменить духовных лидеров, он может значительно ускорить процесс принятия религиозных решений, например, сокращая время подготовки фетвы с 50 дней до 5 часов. Тем не менее внедрение подобных технологий в религиозные процессы несёт в себе риски [Пашенцев, Кузнецов, 2024]. Так, чат-боты, разработанные за пределами Ирана, часто отражают западоцентричные политические взгляды, что может привести к конфликту с официальными нарративами страны. Зачастую такие чат-боты демонстрируют ограниченное понимание шиитских традиций и теологии. Если чат-боты обучены на данных, содержащих экстремистские взгляды, они могут стать инструментами радикальной пропаганды. Показательным является случай с GPT-3, который был обвинён в исламофобии в 2021 г. [Samuel, 2021]. Нарушение работы таких систем или их использование враждебными субъектами может вызвать негативные последствия, особенно в случае дальнейшего развития эмоционального ИИ.

ЗАКЛЮЧЕНИЕ

Человеческий разум ежедневно потребляет больше информации, чем когда-либо в истории человечества, что влияет на информационно-психологическую безопасность людей [Vacarelu, 2023]. В условиях быстрого технологического перехода Ирану крайне важно заблаговременно прорабатывать механизмы контроля, чтобы минимизировать негативные последствия. В данном контексте центральное место занимает анализ рисков, связанных с информационно-психологической безопасностью, которые могут оказывать значительное влияние на стабильность как отдельных государств, так и международного сообщества. Особое внимание следует уделить разработке «доверенных» систем, которые должны быть прозрачными, устойчивыми и соответствовать нормам права. В противном случае существует риск подрыва общественного доверия и расширения возможностей для злоупотреблений.

Так как Иран сталкивается с существенно более серьёзным внешним давлением по сравнению с внутренними вызовами, это делает угрозы со стороны ЗИИИ на втором и третьем уровнях информационно-психологической безопасности особенно опасными для

страны. Технологическая оснащённость внешних акторов позволяет им эффективно манипулировать информационным пространством, усиливая дестабилизирующее влияние.

Следует отметить, что иранские власти демонстрируют системный и стратегически обоснованный подход к развитию технологий ИИ, что свидетельствует о продуманной и дальновидной политике в области научно-технического прогресса. К примеру, Иран занял в текущем году лидирующую позицию среди исламских стран по числу опубликованных научных статей по тематике ИИ.

Это свидетельствует о том, что правительство осознаёт значимость ИИ для будущего социально-экономического развития и активно стремится использовать его возможности. Можно предположить, что, если иранское правительство уделит должное внимание угрозам, связанным с ЗИИИ, в частности в контексте информационно-психологической безопасности, можно ожидать, что стратегия противодействия этим рискам будет столь же структурной и системной, как и поступательное развитие инвестиций в технологический прогресс. Такой структурированный подход позволит Ирану эффективно защищать свои национальные интересы в условиях растущих технологических и информационных вызовов, обеспечивая долгосрочную устойчивость и безопасность.

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

Карасев П.А., Соколов В.В., Шаряпов Р.А. Теорема Томаса, методология информационных операций и приложения искусственного интеллекта. *International Journal of Open Information Technologies*. 2021. Т. 9. № 10. С. 90–98 [Karasev P.A., Sokolov V.V., Sharyapov R.S. Thomas Theorem, Methodology of Information Operations and Applications of Artificial Intelligence. *International Journal of Open Information Technologies*. 2021. Vol. 9. No. 10. Pp. 90–98 (in Russian)].

Пашенцев Е.Н., Кузнецов П.В. Злонамеренное использование искусственного интеллекта: вызовы информационно-психологической безопасности в Исламской Республике Иран. *Злонамеренное использование искусственного интеллекта и вызовы информационно-психологической безопасности в странах БРИКС*. Координатор исследовательского проекта: Е.Н. Пашенцев. М.: ООО «Сам Полиграфист», 2024. С. 44–50 [Pashentsev E.N., Kuznetsov P.V. Malicious Use of Artificial Intelligence and Challenges to Psychological Security in the Islamic Republic of Iran. *Malicious Use of Artificial Intelligence and the Challenges of Information and Psychological Security in the BRICS Countries*. Coordinator of the research project: E.N. Pashentsev. Moscow: “SAM Poligrafist”, 2024. Pp. 44–50 (in Russian)].

Adversarial Misuse of Generative AI. *Google Threat Intelligence Group*. 30.01.2025. <https://cloud.google.com/blog/topics/threat-intelligence/adversarial-misuse-generative-ai> (accessed: 23.05.2025).

Avvalin “Deep-Fake” dar Irān [The First “Deep-Fake” in Iran. *IRNA*. 02.07.2024(2) (in Persian)]. <https://www.irna.ir/news/85514109/اولين-ديپ-فيک-در-ايران> (accessed: 28.11.2024).

Baqāyi: Amerikā dar jāyghāhi nist ke Irān-rā mottaham be modāxele dar omur-e digarān kond [Baghaei: America is not in a position to accuse Iran of interfering in other people's affairs. *IRNA*. 04.10.2024(1) (in Persian)] <https://www.irna.ir/news/85617083/بقايي-امريکا-در-جاياگاهي-نيست-که-ايران-را-متهم-به-مداخله-در-امور> (accessed: 28.11.2024).

Berman L. Defence Ministry to Invest Heavily in AI in Bid to Improve Intel on Iran. *The Times of Israel*. 22.05.2023. <https://www.timesofisrael.com/defense-ministry-to-invest-heavily-in-ai-in-bid-to-improve-intel-on-iran/> (accessed: 28.11.2024).

R. Artificial Intelligence Politics and Sanctions: Comparing the Cases of Russia and Iran. *Digital International Relations*. A. Baykov, E. Zinovieva (eds.) Cham: Palgrave Macmillan, 2023. Pp. 59–72.

Bozorgmehr N. ‘Robots Can Help Issue a Fatwa’: Iran's Clerics Look to Harness AI. *Financial Times*. 24.09.2023. <https://www.ft.com/content/9c1c3fd3-4aea-40ab-977b-24fe5527300c> (accessed: 03.09.2024).

Bradshaw S., Howard P.N. *The Global Disinformation Order 2019: Global Inventory of Organised Social Media Manipulation*. Oxford Internet Institute, 2019. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf> (accessed: 27.09.2024).

Conduit D. Securing Iran in the Internet Age. *The Palgrave Handbook of National Security*. M. Clarke, A. Henschke, M. Sussex, T. Legrand (eds.) Cham: Palgrave Macmillan, 2022. Pp. 240–261.

Deibert R.J. The Road to Digital Unfreedom: Three Painful Truths About Social Media. *Journal of Democracy*. 2019. Vol. 30(1). Pp. 25–39.

Elon Musk's AI Chatbot Grok Writes Fake News Reports, Promotes Them on X. *The Hindu*. 22.04.2024. <https://www.thehindu.com/sci-tech/technology/elon-musks-ai-chatbot-grok-writes-fake-news-reports-promotes-them-on-x/article68093695.ece> (accessed: 26.11.24).

R. The AI Assault on Women: What Iran's Tech Enabled Morality Laws Indicate for Women's Rights Movements. *Council on Foreign Relations*. 07.12.2023. www.cfr.org/blog/ai-assault-women-what-irans-tech-enabled-morality-laws-indicate-womens-rights-movements (accessed: 03.09.2024).

Golkar S. The Evolution of Iran's Police Forces and Social Control in the Islamic Republic. *Middle East Brief*. Brandeis University. 07.2018. <https://www.brandeis.edu/crown/publications/middle-east-briefs/pdfs/101-200/meb120.pdf> (accessed: 28.11.2024).

Hajjaliasgari F., Khanahmadi A., Atashi A. Artificial Intelligence Chatbot in Iran Health Insurance Organization: a New Era in Service Providing. *Iranian Journal of Health Insurance*. 2023. Vol. 6(2). Pp. 91–102.

Homeland Threat Assessment 2024. https://www.dhs.gov/sites/default/files/2023-09/23_0913_ia_23-333-ia_u_homeland-threat-assessment-2024_508C_V6_13Sep23.pdf (accessed: 28.11.2024).

Howard P.N. *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. New Haven: Yale University Press, 2020.

Hunter L.Y., Albert C.D., Rutland J., Topping K., Hennigan C. Artificial Intelligence and Information Warfare in Major Power States: How the US, China, and Russia Are Using Artificial Intelligence in Their Information Warfare and Influence Operations. *Defense & Security Analysis*. 2024. Vol. 40. Pp. 235–269.

Iran Jobless Rate Down 0.4% y/y in Sep Quarter to 7.5 %: SCI. *Press TV*. 05.10.2024. <https://www.presstv.ir/Detail/2024/10/05/734629/Iran-unemployment-rate-decline-SCI-figures> (accessed: 05.10.2024).

Iran Plans to Become a Leading Country in AI. *Tehran Times*. 30.01.2022. <https://www.tehrantimes.com/news/469628/Iran-plans-to-become-a-leading-country-in-AI> (accessed: 03.09.2024).

Iranian Labor Market Getting Hotter with No Sign of Stopping. *Iran Daily*. 06.06.2024. <https://newspaper.irandaily.ir/7574/6/9194> (accessed: 05.10.2024).

Iran's Biggest Corruption Case Rattles Ruling Hardliners. *Iran International*. 07.12.2023. <https://www.iranintl.com/en/202312062449> (accessed: 03.09.2024).

Israel using AI to pinpoint Hamas leaders, find hostages in Gaza tunnels — report. *The Times of Israel*. 26.04.2025. <https://www.timesofisrael.com/israel-using-ai-to-pinpoint-hamas-leaders-find-hostages-in-gaza-tunnels-report/> (accessed: 24.05.2025).

Jhan J.H., Liu C.P., Jeng S.K., Lee H.Y. *CheerBots: Chatbots toward Empathy and Emotion Using Reinforcement Learning*. 2021. <https://doi.org/10.48550/arXiv.2110.03949> (accessed: 10.05.2025).

N., Khalil J. Twitter Aided Pentagon Influence Operations: Report. *News Nation*. 21.12.2022. <https://www.newsnationnow.com/business/tech/twitter-aided-pentagon-influence-operations-report/> (accessed: 28.11.2024).

Mahmood Z., Sahraei F. Iran Installs Cameras to Identify Women Breaking Dress Code. *CNN*. 08.04.2023. <https://edition.cnn.com/2023/04/08/middleeast/iran-cameras-unveiled-women-intl/index.html> (accessed: 26.10.2024).

Mantello P., Ho T. M., Podoletz L. Automating Extremism: Mapping the Affective Roles of Artificial Agents in Online Radicalization. *The Palgrave Handbook of Malicious Use of AI and Psychological Security*. E. Pashentsev (ed.) Cham: Palgrave Macmillan, 2023. Pp. 81–103.

Marcellino W., Magnuson M., Stickells A., Boudreaux B., Helmus T.C., Geist E., Winkelman Z. Counter-Radicalization Bot Research Using Social Bots to Fight Violent Extremism. *Rand Corp.* 2020. <https://apps.dtic.mil/sti/pdfs/AD1111251.pdf> (accessed: 27.09.2024).

May L. Protest and Activism of Women's Movements—Social Movements in Contemporary Iran. *Media and Feminist Protest in Iran*. Cham: Palgrave Macmillan, 2023. Pp. 91–115.

McPherrin J. The Twitter Files Parts 1–15: A Comprehensive Summary, Analysis, and Discussion of Ramifications for American Institutions. *Stopping Socialism*. 19.01.2023. <https://stoppingsocialism.com/2023/01/the-twitter-files-comprehensive-summary-analysis-and-discussion-of-ramifications-for-american-institutions/> (accessed: 28.11.2024).

Mohammadzadegan A. Iran Prioritizes Using AI for Cyber Defense, Says Defense Official. *IRNA*. 13.08.2023. <https://en.irna.ir/news/85197899/Iran-prioritizes-using-AI-for-cyber-defense-says-defense-official> (accessed: 03.09.2024).

Motamedi M. Iranian Official Accuses Israel of Killing Fakhrizadeh Remotely. *Al Jazeera*. 30.11.2020. <https://www.aljazeera.com/news/2020/11/30/iran-israel-killing-scientist-remotely-in-sophisticated-attack> (accessed: 03.09.2024).

Mustafaraj E., Metaxas P.T. The Fake News Spreading Plague: Was It Preventable? *Paper presented at the Proceedings of the 2017 ACM on Web Science Conference*. Troy, New York, USA. 2017. Pp. 235–239. <https://dl.acm.org/doi/10.1145/3091478.3091523> (accessed: 22.05.2025).

Office of the Director of National Intelligence. *Annual Threat Assessment of the U.S. Intelligence Community*. 2024. <https://www.dni.gov/files/ODNI/documents/assessments/ATA-2024-Unclassified-Report.pdf> (accessed: 03.09.2024).

Pargoo M. Sanctions Propel Iran in the Global Race for Terminator – Like AI. *The Atlantic Council*. 02.04.2019. <https://www.atlanticcouncil.org/blogs/iransource/sanctions-propel-iran-in-the-global-race-for-terminator-like-ai/> (accessed: 07.10.2024).

Pashentsev E. General Content and Possible Threat Classifications of the Malicious Use of Artificial Intelligence to Psychological Security. *The Palgrave Handbook of Malicious Use of AI and Psychological Security*. Ed. E. Pashentsev. Cham: Palgrave Macmillan, 2023. Pp. 23–46.

Pashentsev E., Bazarkina D. The Malicious Use of AI Against Government and Political Institutions in the Psychological Arena. *Regulating artificial intelligence in industry*. D.N. Bielicki (ed.) Cham: Routledge, 2022. Pp. 36–52.

Pauwels E., Deton S.W. Hybrid Emerging Threats and Information Warfare: The Story of the Cyber-AI Deception Machine. *21st Century Prometheus*. M. Martellini, R. Trapp (eds.) Cham: Springer, 2020. https://doi.org/10.1007/978-3-030-28285-1_6 (accessed: 10.05.2025).

Pentagon Riding the “Blue Bird” in Psychological Warfare. *Tehran Times*. 25.12.2022. <https://www.tehrantimes.com/news/480127/Pentagon-riding-the-blue-bird-in-psychological-warfare> (accessed: 03.09.2024). Pp. 107–124.

Ramalingam D., Chinnaiiah V. Fake Profile Detection Techniques in Large- Scale Online Social Networks: A Comprehensive Review. *Computers & Electrical Engineering*. 2018. Vol. 65. Pp. 165–177.

Reinhold T., Reuter C. Artificial Intelligence and Cyber Weapons. *Information Technology for Peace and Security. Technology, Peace and Security | Technologie, Frieden und Sicherheit*. C. Reuter (ed.) Cham: Springer Vieweg, Wiesbaden, 2024. Pp. 335–349.

Report of the Independent International Fact-Finding Mission on the Islamic Republic of Iran. *UN HRC*. 02.02.2024. www.ohchr.org/sites/default/files/documents/hrbodies/hrcouncil/sessions-regular/session55/advance-versions/a-hrc-55-67-aev.docx (accessed: 03.09.2024).

Sheikhtaheri M.R., Firouzabadi S.A. Conflict in Iran (Typology, Background, Causes and Dominant Factors). *Quarterly of Social Studies and Research in Iran*. 2024. Vol. 13(3). Pp. 369–390.

Tajmazinani A. Challenges of Establishing an Islamic Welfare State: A Historical Overview of Welfare System Developments in Post-Revolution Iran. *Social Policy in the Islamic World. International Series on Public Policy*. A.A. Tajmazinani (ed.) Cham: Palgrave Macmillan, 2021. Pp. 109–136.

The first destructive use of artificial intelligence in human history by Israel. *Tasnim News*. 14.01.2025. <https://www.tasnimnews.com/fa/news/1403/10/25/3237703/> (accessed: 23.05.2025).

The SCImago Journal & Country Rank. SCImagojr Country Rank 1996–2023. www.scimagojr.com/countryrank.php?category=1702 (accessed: 26.11.2024).

Tucker E., Neumeister L. Justice Department Brings Criminal Charges in Iranian Murder-For-Hire Plan Targeting Donald Trump. *The Associated Press*. 09.11.2024. <https://apnews.com/article/iran-fbi-justice-department-iran-83cff84a7d65901a058ad6f41a564bdb> (accessed: 28.11.2024).

Vacarelu M. Malicious Use of Artificial Intelligence in Political Campaigns: Challenges for International Psychological Security for the Next Decades. *The Palgrave Handbook of Malicious Use of AI and Psychological Security*. E. Pashentsev (ed.) Cham: Palgrave Macmillan, 2023. Pp. 203–230.

A.V., Haghshenas N. Housing Social Policy in Iran: Analyzing the Impact of the Maskan-I-Mehr Program on Socio-Spatial Segregation in the Metropolitan Area of Tehran. *Urban and Regional Development Planning*. 2024. Vol. 9. Pp. 199–244.

Zamahani M. Corruption, Iran. *Global Encyclopedia of Public Administration, Public Policy, and Governance*. A. Farazmand (ed.) Cham: Springer, 2022. Pp. 1–7.

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

ПАШЕНЦЕВ Евгений Николаевич – доктор исторических наук, профессор факультета международных отношений Санкт-Петербургского государственного университета, СПб., главный научный сотрудник Института актуальных международных проблем Дипломатической академии МИД России, Москва, Россия.

Evgeny N. PASHENTSEV, DSc (History), Professor, Faculty of International Relations, Saint Petersburg State University, Saint Petersburg, Russia; Chief Research Fellow, Institute of Current International Problems, Diplomatic Academy of the Russian Foreign Ministry, Moscow, Russia.

КУЗНЕЦОВ Павел Валентинович – магистрант факультета международных отношений Дипломатической академии МИД России, Москва, Россия.

Pavel V. KUZNETSOV, Master's student, Faculty of International Relations, Diplomatic Academy of the Russian Foreign Ministry, Moscow, Russia.

ЧЕБЫКИНА Владилена Александровна – магистрант факультета международных отношений Санкт-Петербургского государственного университета, СПб., стажер-исследователь Международного центра социально-политических исследований и консалтинга, Москва, Россия.

Vladilena A. CHEBYKINA, Master's student, Faculty of International Relations, Saint Petersburg State University, Saint Petersburg, Russia; Trainee Researcher, International Center for Social and Political Studies and Consulting, Moscow, Russia.