

[Весь мир охвачен беспрецедентным интересом к технологиям ИИ. Какие риски, опасности и угрозы они несут человеку, обществу и национальной безопасности России?](#)

[Анатолий Смирнов](#) , главный научный сотрудник ЦМИБ и НТП Института международных исследований, профессор МГИМО МИД России, доктор исторических наук

В Концепции внешней политики Российской Федерации (утверждена Президентом Российской Федерации В.В.Путиным 31 марта 2023 г.) отмечено, [\[2\]](#) что человечество переживает эпоху революционных перемен, продолжается формирование более справедливого, многополярного мира.

США и их сателлиты рассматривают укрепление России в качестве одной из угроз западной гегемонии. Принятые Россией меры по защите своих национальных интересов на украинском направлении «коллективный запад» использовал как предлог для обострения многолетней антироссийской политики и развязал гибридную войну нового типа. Резко усилилось политико-дипломатическое, экономическое, военное и информационное сдерживание России и её партнеров.

С появлением технологий ИИ изречение «история войн – это история великих открытий» наполняется принципиально новым, более грозным смыслом. Действительно, наряду с несомненным позитивом технологий ИИ для социально-экономического прогресса человечества, эксперты ставят ИИ в перечень технологий, навсегда меняющих облик войны и мира.

По оценкам ряда экспертов ИИ – это не очередной (седьмой) технологический уклад [\[3\]](#) , а фазовый переход человечества, сопоставимый, например, с изобретением электричества.

За период с 2017 г. более 60 стран разработали и приняли национальные стратегии развития ИИ. В этих условиях Указом Президента РФ от 10.10.2019 №490 была утверждена Национальная стратегия развития ИИ на период до 2030 года. [\[4\]](#) В Страте

гии определено, что Россия должна стать одним из международных лидеров в развитии ИИ.

В ходе прямой линии «Итоги года с Владимиром Путиным» 14.12.2023г. президент, отвечая на вопрос дипфейкового двойника об ИИ, подчеркнул: «Предотвратить развитие искусственного интеллекта невозможно, а значит, это направление нужно возглавить.» [\[5\]](#) В этих целях президент поручил обновить Национальную стратегию развития ИИ до 2030 года.

### Гонка ИИ вооружений

Следует отметить, что за две трети века технологии ИИ прошли огромный путь от фантастики до глобального охвата человечества. При этом ИИ, как технологии двойного назначения, создают и многофакторные угрозы.

В вышедшей в июне 2022 г. книге «Искусственный разум и новая эра человечества» [\[6\]](#) три автора: экс-госсекретарь США и лауреат Нобелевской премии мира 1973 года Генри Киссинджер, бывший глава Google, экс-председатель совета директоров Alphabet, Эрик Шмидт и декан Колледжа компьютерных наук имени Шварцмана в Массачусетском технологическом институте, экс-директор Amazon Дэниел Хаттенлокер попытались ответить на вопрос, как трансформируется человечество и каким образом на него влияют технологии ИИ, в том числе в области политики и безопасности?

Вердикт политических тяжеловесов не радует. Контроль над ИИ вооружениями невозможен. Кроме того, опасность ИИ на службе политиков в том, что у него нет ни психологии, ни морали, а есть только предустановленные цели и задачи.

Данные опасения не были случайны. Через месяц после появления самого мощного генеративного ИИ ChatGPT4 в апреле 2023 года в сети появился Chaos-GPT – инструмент ИИ, целью которого было заявлено уничтожение человечества.

[\[7\]](#)

Не без влияния данного факта в мае 2023 года группа лидеров отрасли ИИ в заявлении (из одного предложения!) предупредила, что «Снижение риска вымирания из-за искусственного интеллекта должно быть глобальным приоритетом наряду с другими рисками общественного масштаба, такими как пандемии и

ядерная война»

[\[8\]](#)

.  
Заявление подписали более 350 экспертов в сфере ИИ. Среди подписантов были топ-менеджеры ведущих компаний в области ИИ: OpenAI, Google DeepMind, Anthropic, которые в мае 2023 г. встретились с президентом Байденом, чтобы обсудить вопросы регулирования ИИ. По состоянию на 18 декабря 2023 года общее количество подписей составляло более 35 тыс.

[\[9\]](#)

### **Инициативы Великобритании по проблеме безопасности ИИ**

В ходе визита премьер-министра Великобритании Риши Сунака в США 7 июня 2023 года он обсудил и согласовал с президентом США Байденом в том числе вопрос о внесении темы ИИ в ООН, о проведении глобального саммита по безопасности ИИ в Лондоне. [\[10\]](#)

В соответствии с Уставом ООН главная ответственность за поддержание международного мира и безопасности возложена на Совет Безопасности ООН. Великобритания как действующий в июле 2023 года председатель Совбеза ООН приняла решение провести 18 июля 2023 года заседание на тему «Искусственный интеллект: возможности и риски для международного мира и безопасности»

.  
Мининдел Великобритании Клеверли на заседании Совбеза обратился к генсеку ООН Антониу Гутерришу «Ваше присутствие сегодня подчеркивает важность рассматриваемой темы». [\[11\]](#)

Генсек ООН, по его словам, уже шесть лет пристально наблюдает за развитием ИИ. Он подчеркнул, что не ожидал столь сильного влияния ИИ на устойчивое развитие и социальный порядок и признался, что «был поражен и впечатлен» радикальным прорывом в возможностях ИИ, который проявился в ChatGPT, а также других подобных сервисах. Согласно прогнозам к 2030 году ИИ может принести мировой экономике от \$10 трлн до \$15 трлн дохода.

Относительно «возможностей и рисков» Гутерриш заявил, что даже сами разработчики не знают, куда приведет их этот потрясающий технологический прорыв.

Генсек ООН отметил, что злонамеренное использование систем ИИ преступными и террористическими организациями вплоть до государственного уровня может вызвать смерти и разрушения ужасающих масштабов.

В этом контексте он поддержал призывы ряда стран к созданию нового органа ООН по вопросам ИИ, по аналогии с МАГАТЭ или ИКАО.

Клеверли согласился и предложил провести в Лондоне осенью первый глобальный саммит, посвященный безопасности и регулированию ИИ.

В канцелярии премьер-министра Риши Сунака сообщили, что в работе саммита 1-2 ноября 2023 г. примут участие ученые, а также представители руководства стран-лидеров на рынке технологических компаний и «ключевых государств». Их список оглашать не стали, но издание The Daily Telegraph сообщило, что речь идет о встрече глав «государств-единомышленников». Россия, естественно, в список «государств-единомышленников» не вошла.

30 октября, накануне саммита, на встрече «семерки», проходившей в японской Хиросиме, было опубликовано совместное заявление стран G7 об основных принципах регулирования индустрии ИИ

. 11 положений заявления представляют собой основные проблемы и риски — с предложениями о том, как можно их избежать или минимизировать.

На саммите по безопасности ИИ в Блетчли-парке 1-2 ноября 2023 года, приняли участие около 100 представителей правительств, ведущих компаний, занимающихся ИИ, гражданского общества и научных кругов 28 стран и ЕС.

Представители правительств 12 стран подписали достаточно обтекаемую Декларацию Блетчли, а также провели обсуждение на восьми круглых столах дежурных тем. [\[12\]](#) Важным для Великобритании стал запуск первого в мире Института безопасности ИИ, задачей которого является тестирование безопасности новых типов ИИ.

[\[13\]](#)

Институт призван тестировать новые типы ИИ до и после их выпуска. Цель - выявление потенциально опасных возможностей у моделей ИИ, включая изучение всех рисков, от социального вреда, такого как предвзятость и дезинформация, до самой экстремальной угрозы: полной потери человечеством контроля над ИИ.

Запустив Институт безопасности ИИ, Великобритания пытается укреплять свои позиции как лидера в области безопасности ИИ и уже договорилась о сотрудничестве с американским институтом безопасности ИИ и с правительством Сингапура о взаимодействии в области тестирования безопасности ИИ.

При этом Институт безопасности ИИ станет носителем международного стандарта. В преддверии выхода новых мощных моделей ИИ, возможности которых, возможно, еще не полностью поняты, его первой задачей будет быстрое внедрение процессов и систем для их тестирования перед запуском, включая модели с открытым исходным кодом. На саммите достигнута договоренность, что переговоры по безопасности ИИ продолжатся в Южной Корее через полгода, а через год - на следующем саммите во Франции.

Резюмируя, нетрудно видеть, что «коллективный запад» использует уже апробированную методику обхода универсальных международных организаций, в первую очередь ООН. В данном случае представители 12 государств-подписантов декларации приняли «правила» глобальной безопасности технологий ИИ в целях обеспечения технологического доминирования с далеко идущими геополитическими амбициями. При этом нарушается один из основополагающих принципов Устава ООН – принцип суверенного равенства государств

### **Указ президента США Байдена о безопасной разработке и использовании ИИ как платформа обеспечения преимущества в предвыборной гонке**

США – один из лидеров создания технологий ИИ. Наряду с государственным участием самое активное участие в развитии ИИ играют ИТ-гиганты, тесно сотрудничающие с ВПК, а также ряд научных и экспертных сообществ.

30 октября 2023 года президент США Байден издал беспрецедентный Исполнительный указ о безопасной разработке и использовании заслуживающего доверия ИИ

[14]

Документ поражает своей масштабностью и формулирует политику, принципы, а также новые задачи для США, чтобы оставаться мировым лидером в технологиях ИИ и в целях обеспечения национальной безопасности.

Особое внимание уделено роли силового блока министерств, в том числе разведсообщества, а также Госдепартамента, Министерства торговли, Министерства энергетики и иных ведомств как в создании ИИ, так и в системе защиты, в том числе критически важной инфраструктуры.

Одним из ключевых пунктов указа является разработка и представление Президенту Меморандума о национальной безопасности по ИИ

(в течение 270 дней после даты издания указа). Меморандум должен быть посвящен управлению ИИ, используемым в качестве компонента системы национальной безопасности или в военных и разведывательных целях. В частности, меморандум должен:

- предоставить рекомендации Минобороны, другим агентствам и разведывательному сообществу рекомендации по внедрению возможностей ИИ в интересах национальной безопасности США;

- руководить дальнейшими действиями, в зависимости от обстоятельств и в соответствии с применимым законодательством, для предотвращения потенциального использования систем ИИ противником и другими иностранными субъектами способами, которые угрожают возможностям или целям Минобороны или разведывательного сообщества или иным образом представляют угрозу безопасности США или их союзников и партнеров.

Заслуживает внимания также положение указа о том, что Министр внутренней безопасности должен учредить Совет по безопасности и защите ИИ в качестве консультативного комитета.

В его состав должны входить эксперты ИИ из частного сектора, научных кругов и правительства, для подготовки рекомендаций в целях повышения безопасности, устойчивости и

реагированию на инциденты, связанные с использованием ИИ в критической инфраструктуре.

Следует особо подчеркнуть, что в указе имеется специальный пункт о снижении рисков на стыке угроз ИИ, химического, биологического, радиоактивного и ядерного оружия

Важно отметить, что указом предписывается ИТ-гигантам: Microsoft, Amazon, Google и др. делиться результатами тестов разработки новых систем ИИ с федеральным правительством. При этом следует осознавать, что именно мощные системы ИИ позволяют креативно создавать абсолютно правдоподобные дипфейки в виде роликов, изображений, мемов и т.д. Другими словами, это технологии когнитивно-ментального воздействия на человека и манипулирование его поступками на подсознательном уровне, в том числе и в электоральном процессе.

Одиннадцать ключевых пунктов указа имеют срок выполнения 270 дней, т.е. к 1 августа 2024 года.

Таким образом не вызывает никакого сомнения, что указ Байдена приурочен к финишному этапу президентской гонки 2024 г. Естественно, в интересах действующего президента США или его однопартийца.

В условиях подготовки к выборам Президента России в марте 2014 г. стоило бы учитывать огромные возможности ИИ манипулирования электоральным процессом, а также в условиях информационной агрессии «коллективного запада» против России поднять с помощью ИИ «четвертую волну демократизации» (по Хантингтону) путем раскола нашего общества.

С учетом предстоящего председательства России в БРИКС+ продвигать предложение председателя КНР Си Цзиньпина о создании рабочей группы по регулированию ИИ.

*Мнение автора может не совпадать с позицией редакции*

[1] Национальная Ассоциация международной информационной безопасности <https://namb.online/>

[2] <http://www.kremlin.ru/events/president/news/70811>

[3] Глазьев С.Ю. Уроки современной революции: крах либеральной утопии и шанс на «экономическое чудо» / С.Ю. Глазьев. — М.: Издательский дом «Экономическая газета», 2011. — С. 330.

[4] <http://www.kremlin.ru/acts/bank/44731>

[5] <http://www.kremlin.ru/events/president/news/72994>

[6] <https://alpinabook.ru/catalog/book-iskusstvennyy-razum-i-novaya-era-chelovechestva/>

[7] <https://decrypt.co/126122/meet-chaos-gpt-ai-tool-destroy-humanity>

[8] THE CENTER FOR AI SAFETY (safe.ai) Statement on AI Risk <https://www.safe.ai/statement-on-ai-risk>

[9] <https://www.safe.ai/statement-on-ai-risk>

[10] <https://www.gov.uk/government/news/uk-to-host-first-global-summit-on-artificial-intelligence>

[11] <https://media.un.org/ru/asset/k1j/k1ji81po8p>



[12] <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>

[13] <https://www.gov.uk/government/news/prime-minister-launches-new-ai-safety-institute>

[14] <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>

Источник: <https://interaffairs.ru/news/show/43847>