

---

## НЕМНОГО ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ: НЕОДНОЗНАЧНЫЕ ТЕХНОЛОГИИ



**Дмитрий Вадимович Мун**

Кандидат экономических наук, заместитель директора ФГБУ «Агентство “Эмерком”» МЧС России, сооснователь международного экспертного сообщества RISK.TODAY, Москва, Россия



**Владислав Владиславович Попета**

Кандидат технических наук, основатель и президент международного экспертного сообщества RISK.TODAY, Москва, Россия

***Аннотация.** Рассмотрены особенности использования искусственного интеллекта по различным направлениям. Представлены примеры неоднозначного и не всегда конструктивного применения подобных технологий. Акцентировано внимание на необходимости формирования систем и правовых основ безопасности при работе с искусственным интеллектом.*

***Ключевые слова:** искусственный интеллект; фейковые новости; новые технологии; риски; безопасность.*

***Для цитирования:** Мун Д.В., Попета В.В. Немного об искусственном интеллекте: неоднозначные технологии // Социальные новации и социальные науки. – Москва : ИНИОН РАН, 2021. – № 2. – С. 26–37.*

URL: <https://sns-journal.ru/ru/archive/>

DOI: 10.31249/snsn/2021.02.02

© Мун Д.В., © Попета В.В.

*Мне кажется, компьютерные вирусы стоит рассматривать как форму жизни. Это многое говорит о природе человека: единственная форма жизни, которую мы создали к настоящему моменту, несет только разрушения. Мы создаем жизнь по образу и подобию своему.*

[Хокинг, 2021]

## **Введение**

Главный механизм передачи знаний от человека к человеку, от поколения к поколению – это воспитание (включая образование). Таковы законы природы. Таковы законы цивилизации. На данных принципах происходила эволюция человечества с самого начала его существования и до нынешних времен. Дети, вырастая, становятся более знающими и умелыми, чем родители. Многим кажется, что развитием детей можно управлять, «лепить» их по своему образу и подобию или даже прихоти. Но это далеко не так. Тем не менее определенные морально-этические принципы должны быть внедрены, подобно базовому программному обеспечению, на самых ранних порах становления личности, чтобы общество могло прогрессировать дальше.

В случае же с ИИ во всем мире предполагается сначала создать «нечто» интеллектуальное, потом сразу наделить его силой, которую люди уже заведомо не смогут контролировать, и только затем задуматься о сопутствующих рисках, а также попытаться поговорить с этим неведомым нечто об этике и морали. При том что внятных философских концепций и нормативно-правовых документов даже для применения беспилотных транспортных средств пока еще не разработано. Например, если робот-автомобиль собьет человека – а такие случаи уже есть [Беспилотник Uber ..., 2018] – кто за это ответит?

Другой вопрос: не пройдя биологически проверенного этапа морально-этического становления, но будучи более сильным и «умным», чем люди, его создавшие, станет ли ИИ их слушать? Эта проблема осознается не только экспертным сообществом, но и на государственном уровне. В Национальной стратегии развития искусственного интеллекта в России (НС) указывается, что отсутствие понимания того, как ИИ достигает результатов, является одной из причин низкого уровня доверия и может стать препятствием его развития. Однако, хотя в НС предполагается осуществить значительные финансовые вливания и создать «экосистему» компаний, занимающихся развитием и внедрением ИИ, разработке единых стандартов безопасности уделяется немного внимания (один подраздел) [Указ Президента РФ от 10.10.2019 № 490 ..., 2019, с. 34].

По идее, сначала должны быть разработаны общие правила безопасности – но в начавшейся всемирной гонке за ИИ на это не хватает времени. В любом случае анализ истории развития дру-

гих прорывных технологий наглядно показывает, что средства безопасности применительно к ним создаются со значительным запозданием.

Один из величайших предсказателей XX в., английский писатель-фантаст Г. Уэллс (1866–1946) в своих произведениях часто обращался к проблеме овладения человечества силой и энергией, которые, с одной стороны, двигают вперед прогресс, а с другой – могут поставить цивилизацию на грань уничтожения [Уэллс, 1964; Wells, 1908]. В противовес научным трудам современников, в которых, например, описывалось, как в будущем атомная энергия позволит «освоить пустыни, растопить полюса и превратить всю землю в эдемский сад, озаренный улыбкой» [Содди, 1910], Г. Уэллс рисовал гораздо более пессимистичные картины. И они оказались ближе к реальности, чем мечты великих физиков.

Ученые выявляют объективные законы мироздания и на этой основе создают новые технологии. Писатели хорошо разбираются в человеческих эмоциях, включая негативные (алчность, гнев, тщеславие, зависть, лень и т.д.). Поэтому, наверное, их видение будущего общества оказывается более точным, чем представления специалистов в области естественных и точных наук. Пока эти ученые в начале нового научно-технологического прорыва мечтают о том, как они своими изобретениями подарят радость и счастье всему миру, представители творческих профессий (писатели, художники, режиссеры и т.д.) далеко не всегда разделяют их оптимизм. С ними во многом солидарны ученые-обществоведы. Например, Ю.Н. Харари предупреждает: «Хотя прогресс и открывает перед нами удивительные возможности, я вижу свою задачу в привлечении внимания к связанным с ними опасностям и угрозам. Разумеется, корпорации и бизнесмены, возглавляющие технологическую революцию, в первую очередь превозносят ее достижения; вот почему на плечи социологов, философов и историков вроде меня ложится обязанность бить тревогу, предупреждая человечество о том, что многое, к сожалению, может пойти не так» [Харари, 2019]. Все это в полной мере относится к внедрению технологий ИИ, которые далеко не безопасны и могут использоваться не всегда в конструктивных целях.

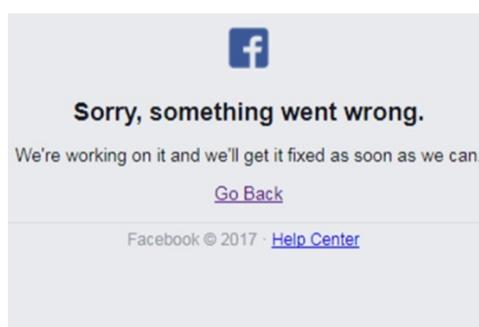
### **Фейсбук: прерванный эксперимент**

Эксперименты по организации взаимодействия человека и ИИ иногда заканчиваются нежелательным результатом. Так, в июле 2017 г. стало известно, что руководство социальной сети Facebook вынуждено было отключить свою систему ИИ (проект Facebook Artificial Intelligence Research – FAIR) после того, как машины (чат-боты) начали общаться на собственном, несуществующем языке, который люди не понимали. Изначально чат-боты создавались для общения с живыми людьми. Сперва они переписывались на английском языке, но в какой-то момент перешли на язык, который сами создали в процессе развития программы. Эксперты, прервавшие эксперимент, высказали опасение, что если боты начнут активно на нем общаться, то постепенно станут

все более самостоятельными и смогут функционировать вне контроля IT-специалистов [Боты изобрели свой язык ..., 2017].

Комментируя данный инцидент, исследователь Лаборатории нейронных систем и глубокого обучения МФТИ В. Малых заявил: «...если же заглянуть в более отдаленное будущее, где компьютеры будут уже достаточно мощными, то стоит вспомнить, что у компьютера нет тела, глаз или желудка – он будет воспринимать все окружающее совсем не так, как мы, и придумает очень отличающийся от нашего язык. Поскольку компьютер будет воспринимать иначе вообще все, будет хотеть другого, бояться другого, – он будет и мыслить иначе. Из этого следует, что интересы людей с интересами компьютеров будут пересекаться слабо или вообще пересекаться не будут. Так что велик шанс, что нам просто будет «не о чем поговорить». То есть мы будем сосуществовать, но практически не взаимодействовать, как сейчас люди и дельфины (второй условно разумный вид на нашей планете...)» [Малых, 2017].

В августе 2017 г. в сети Facebook произошел глобальный сбой. Пользователи по всему миру лишились доступа к своим аккаунтам. Вместо своих страничек они видели следующую картинку:



**Рис. 1. Скриншот страницы Facebook**

Через несколько часов контроль над крупнейшей мировой социальной сетью был восстановлен. Наверное, это просто совпадение. Но можно вспомнить и ИИ Скайнет<sup>1</sup> из фильма «Терминатор» – вымышленного суперкомпьютера министерства обороны США, развязавшего третью мировую войну. Вместе с тем, по мнению американского социолога Ч. Перроу, долгое время изучавшего социальные и технологические аспекты «рукотворных катастроф», определенные сложные системы невозможно сделать достаточно безопасными, и наличие системных ошибок (обычной или системной аварии) является их неотъемлемым свойством [Perrow, 1984].

---

<sup>1</sup> Англ. SkyNet – «небесная сеть».

## **Спасет ли искусственный интеллект человечество?**

*Добро пожаловать в будущее. Оно металлическое и на батарейках.  
Есть надежда, правда, что оно беззлобно и не станет разрабатывать  
способы порабощения человечества.*  
[Уотсон, 2012]

Скоро мир будет праздновать двадцать пятую годовщину того, как суперкомпьютер Деер Blue II<sup>1</sup> в 1997 г. обыграл Г. Каспарова – действительного на тот момент чемпиона мира по шахматам. Разработанная компанией Google машина AlphaGo в 2016 г. обыграла Ли Седоля, лучшего на планете игрока в го – игру, которая не основана, подобно шахматам, на одних лишь алгоритмах. В настоящее время более трех тысяч роботов по всему миру проводят хирургические операции на людях. Беспилотные автомобили уже проехали по общественным дорогам миллионы километров. В Китае уже существуют правила дорожного движения для дронов. В США уже принят на работу первый робот – адвокат. На очереди роботы-охранники, роботы-аптекари, роботы-повара, роботы-учителя, роботы-сиделки, роботы – члены советов директоров.

«Приход искусственного интеллекта заменит не только переводчиков, андеррайтеров и водителей, но и нас с вами», – утверждает российский предприниматель Л. Хасис [Карасюк, 2016]. Не верил в «картину всеобщего счастья» британский физик С. Хокинг. В интервью газете Independent он сказал: «Все эти достижения меркнут на фоне того, что нас ждет в ближайшие десятилетия. Успешное создание искусственного интеллекта станет самым большим событием в истории человечества. К сожалению, оно может оказаться последним, если мы не научимся избегать рисков... Когда искусственный интеллект начнет управлять финансовыми рынками, научными исследованиями, людьми и разработкой оружия, это будут вещи, недоступные нашему пониманию... если краткосрочный эффект искусственного интеллекта зависит от того, кто им управляет, то долгосрочный эффект – от того, можно ли будет им управлять вообще» [Хокинг, 2014].

Ему вторит инженер, изобретатель и бизнесмен, основатель инновационных проектов «Tesla Motors» и «Space-X» И. Маск. Он прогнозирует риск появления «чего-то очень опасного в пределах 5–10 лет», подразумевая искусственный разум. И этот разум может быть отнюдь не дружелюбен по отношению к создавшему его человеку. В январе 2015 г. И. Маск назвал ИИ «созданным человеком демоном» и пожертвовал 10 млн долл. на исследования по контролю над ним. В 2017 г., выступая на саммите Национальной ассоциации губернаторов США, Маск заявил: «Я продолжаю бить тревогу, но пока люди сами не увидят роботов, которые ходят по улицам и убивают людей, они не узнают, как реагировать [на ИИ]». Человечеству надо быть осторожным с технологией, которая «потенциально более опасна, чем ядерное оружие». Если вовремя не вмешаться в процесс развития этих систем, то будет слишком поздно [Таиров, 2020].

---

<sup>1</sup> Англ., дословно – «темно-синий» или, в литературном переводе, «глубокая печаль».

7 мая 2016 г. войдет в историю как дата первого ДТП с работающим автопилотом, закончившегося смертью человека [Погиб водитель ..., 2016]. По зловещей иронии судьбы погибшим «за рулем автомобиля будущего» оказался Дж. Браун – человек, который был активным сторонником электромобилей и продукции компании Tesla Motors. Во Всемирной сети он стал широко известен после того, как менее чем за месяц до смертельного инцидента выложил видео о том, как включенный автопилот спас его от столкновения с грузовиком (адрес ролика <https://youtu.be/9I5rraWJq6E>). Лишь 30 июня компания Tesla Motors на своем сайте опубликовала официальное сообщение об инциденте [A tragic loss ..., 2016]. Несмотря на то что пострадавший был известен, в пресс-релизе «Теслы» он упоминался лишь как «клиент» и «друг “Теслы”».

Неприятные инциденты с автопилотами «Теслы» происходят не так уж редко [Аварии и ДТП с автомобилями Tesla, 2015]. Компания даже была вынуждена признать, что на данный момент автопилот еще находится в стадии тестирования и, согласно прилагающейся к автомобилю инструкции, водитель активирует его под свою ответственность. При этом он должен быть готовым в любой момент вмешаться в процесс вождения, а его руки должны постоянно лежать на руле (иначе система подаст звуковой сигнал и постепенно остановит электромобиль). Последнее условие стало решающим аргументом для Национального управления по безопасности движения на автомагистралях (NHTSA) США, пришедшего к выводу о том, что винить в смерти водителя неисправность автопилота электромобиля или компанию Tesla Motors нельзя.

Однако многие люди поддерживают идею создания и использования ИИ: «в конце концов, у нас впервые появилось достаточно сил наделать совершенно новых ошибок» [Деннетт, 2019]. Считается, что ИИ – это вещь, которая сможет защитить нас от самих себя. Или это уже будет не вещь? По сути, любая машина – это всего лишь «запрограммированное железо». И никто за нас не решит возникающие проблемы.

### **Использование технологий ИИ для искажения реальности<sup>1</sup>**

*Начало.* В декабре 2017 г. сотрудники исследовательского центра NVIDIA – одного из пионеров освоения возможностей ИИ – опубликовали на своем сайте статью, в которой продемонстрировали, как можно при помощи определенной программы полностью видоизменить исходное цифровое фото, превратив, например, зимние пейзажи в летние, а солнечные – в дождливые (рис. 2). В отличие от фотошопа, применение которого человек может распознать невооруженным взглядом, разработанный в NVIDIA алгоритм «глубокого» нейросетевого машинного обучения факти-

---

<sup>1</sup> Далее повторены отдельные примеры, опубликованные авторами ранее в работе [Мун, Попета, 2020].

чески позволяет создавать новую цифровую реальность. Даже эксперты не смогли опознать подделку картинок, созданных при помощи технологии «глубинного обучения»<sup>1</sup> [Powell, 2017].



**Рис. 2. Скриншот с сайта NVIDIA. Как переделывались видео**

Новые технологии, часто размещаемые на открытых платформах (в том числе относящиеся к ИИ), к сожалению, быстрее всего берутся на вооружение преступниками, хакерами и просто нечистоплотными людьми. Так, неделю спустя на социальном новостном сайте Reddit появились ссылки на порнографическое видео с участием одной из самых известных и высокооплачиваемых голливудских актрис Галь Гадот [Cole S. AI-assisted ..., 2017]. Однако актриса даже не подозревала о его создании.

Автор оказавшегося поразительно достоверным фейкового видео – программист, скрывающийся под ником Deepfakes, – моментально стал чрезвычайно популярен. В последовавших за скандальной публикацией многочисленных интервью Deepfakes не только рассказал о себе, но и раскрыл общественности, как он создал свое «творение», и указал на необходимое для этого программное обеспечение (которое бесплатно и находится в открытом доступе). Выяснилось, что создавать подобные ролики может каждый, кто умеет пользоваться Интернетом и понимает принципы работы генеративно-состязательных нейросетей.

Появившийся новый жанр в порноиндустрии получил название «Дипфейк»<sup>2</sup>. Менее чем через месяц создание дипфейковых порновидеороликов со «звездами» превратилось в мощную индустрию. Юристы призывают «оскорбленных лиц» подавать на новоявленных «порнорежиссеров»

---

<sup>1</sup> Англ. Deep learning.

<sup>2</sup> Англ. deep – глубокий и fake – подделка.

в суд за незаконную обработку персональных данных. Однако «звезды», памятуя об известном «эффекте Стрейзанд»<sup>1</sup>, не торопятся выступать с официальными опровержениями, чтобы не привлекать к себе внимания в данном контексте со стороны пользователей Всемирной сети.

*Эволюция дипфейка.* Вредоносное применение технологий ИИ не ограничилось созданием дипфейковых порно и несанкционированным использованием широко известных «персональных данных» звезд гламура и кино. Оно стало распространяться и в другие области общественной жизни.

В настоящее время самым простым и популярным способом потребления информации стало мобильное видео. Главным распространителем такого видео является YouTube – самый крупный видеохостинг в мире. Летом 2018 г. рекордсменом по просмотрам на YouTube стало видеообращение, на котором бывший президент США Б. Обама публично оскорбляет действующего президента США Д. Трампа. Видео было настолько достоверным, что поначалу никто не поверил, что оно является подделкой (рис. 3).

Авторами провокационного ролика выступили американский комик и режиссер Дж. Пил и новостная интернет-медиакомпания BuzzFeed. Дипфейк-видео, едва не ставшее поводом для «гражданской войны» между сторонниками республиканцев и демократов, было создано при помощи графического редактора Adobe After Effects и простой бесплатной программы Fakeapp, выложенной в открытом доступе на сайте Reddit для всех желающих самостоятельно поупражняться в создании «порно со звездами». По словам Пила, главная цель, которую он, как режиссер, преследовал при создании своего «фильма», благородна: предупредить людей об опасности недостоверной информации. «...мы должны быть более бдительными в вопросе, чему верить в Сети.... полагаться на авторитетные новостные источники, которым стоит доверять» [Vincent, 2018].



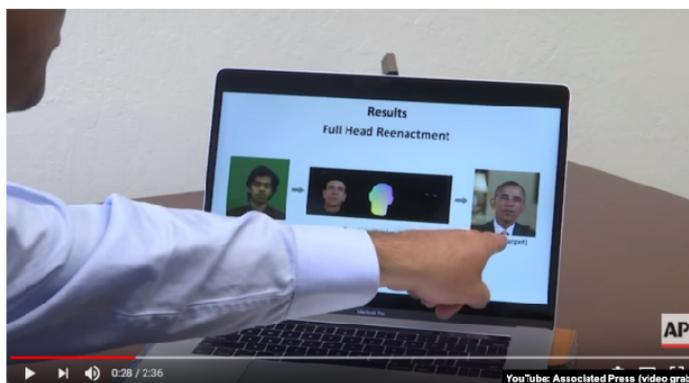
**Рис. 3. Дипфейк-видео. Вы не поверите, что Обама говорит на этом видео!**  
Источник [You won't ..., 2018].

---

<sup>1</sup> Социальный феномен, выражающийся в том, что попытка изъять определенную информацию из публичного доступа приводит лишь к ее более широкому распространению в Интернете [Jansen, Martin, 2015].

В непременно следующих за подобными эпизодами интервью журналистам Дж. Пил также поведал следующее: «Я знал, что это возможно, но не был уверен, действительно ли это сработает.... Знаете, пока это еще достаточно сложно, но это гораздо легче, чем должно быть» [«Дипфейк» ..., 2018]. Действительно, создание дипфейков становится слишком простым (рис. 4).

Активно осваивают современные информационные технологии и преступники. Всемирные новостные ленты все чаще сообщают о случаях подделки голосов директоров и CEO, по устному указанию которых сотрудники переводили деньги мошенникам, или о дипфейковых видео со звездами экрана, призывающими перейти по ссылке и получить приз, и так далее. Сегодня, с учетом количества и качества цифрового материала, хранящегося во Всемирной сети и доступного каждому, не только политики и селебритиз, но и любой рядовой пользователь, имевший неосторожность выложить селфи и / или видео на всеобщее обозрение, рискует стать жертвой не просто розыгрыша, но и серьезной провокации или мошенничества.



**Рис. 4. Скриншот с обучающего видео на YouTube, на котором объясняется, как работать с программами по созданию дипфейков**

Но самой большой опасностью, по мнению профессионалов информационной безопасности, являются дипфейки, которые можно использовать при ведении информационных войн. Профессор Калифорнийского университета (Лос-Анджелес, США) Дж. Вильясенор утверждает, что «технология deepfake станет мощнейшим инструментом в руках тех, кто захочет дезинформировать массы и напрямую влиять на политический процесс». В свою очередь, американские военные считают самым перспективным направлением редактирование и вбросы противникам поддельных спутниковых фото и видео, которые смогут их быстро дезориентировать и повергнуть в хаос [Tucker, 2019].

В Интернете уже несколько раз появлялись «пробные» видеоролики анонимных «энтузиастов», на которых руководители двух ядерных супердержав – президент США Д. Трамп и президент России В. Путин, – сообщают о планах глобальных вооруженных нападений [Кадочникова, 2020]. Хотя многие зрители отнеслись к данным видео как к провокационным розыгрышам и не поддались панике. Этому способствовало низкое качество представленных доказательств. Однако

нейросетевые технологии стремительно совершенствуются. Уже близка возможность создания абсолютно достоверных дипфейк-видео, которые будет невозможно, даже подвергнув экспертизе, отличить от реально снятых кадров.

Если верить прогнозу одного из пионеров всемирной индустрии дипфейк-видео, Хао Ли, «счет идет на месяцы, а не на годы» [Носков, 2019]. В интервью CNBC Хао Ли отметил, что «такие дипфейки смогут создавать обычные пользователи без особой технической подготовки или уникального оборудования, и у нас есть лишь от шести до 12 месяцев, чтобы научиться распознавать подделки» [Stankiewicz, 2019].

*Перспективы.* Было бы неверно утверждать, что проблеме создания специальных программ – «противоядия против дипфейков» – не уделяется должного внимания со стороны правительств, корпораций и частных лиц. Но, к сожалению, как это постоянно происходит в истории, сначала люди изобретают технологию и лишь потом задумываются, как себя обезопасить от нее. Разработка средств безопасности и защиты человека от достижений технического прогресса всегда сильно запаздывает по сравнению с темпами последнего. Сначала люди освоили огонь – и лишь тысячелетия спустя додумались изобрести огнетушитель и стали организовывать профессиональные пожарные команды; начали производить автомобили, разгоняющиеся до ста и более километров в час, – и десятилетия спустя изобрели трехточечный ремень безопасности [Мун, 2018] и т.д. К чему могут привести «шалости и игры» дилетантов с неизведанными пока возможностями ИИ – даже страшно себе представить...

### **Заключение**

В настоящее время каждый отдельно взятый индивид и общество в целом должны ясно осознавать, что любым открытиям, возможностям и перспективам сопутствуют угрозы, опасности и риски. Чем больше возможностей, тем больше, соответственно, и рисков. Иначе не бывает. Искать новые возможности и одновременно адекватно оценивать сопутствующие риски – один из основных принципов риск-ориентированного мышления. Однако лучше в первую очередь адекватно оценивать риски, а уже потом, сочтя их приемлемыми и допустимыми, рассуждать об открывающихся возможностях. Это особенно актуально по отношению к расширяющемуся использованию технологий ИИ.

### **Список литературы**

1. Аварии и ДТП с автомобилями Tesla // Автомобиль Tesla. – 2015. – URL: <http://teslazone.ru/accidents> (дата обращения 04.04.2021).
2. Беспилотник Uber сбил насмерть пешехода из-за настроек автопилота // Ведомости. – 2018. – 08.05. – URL: <https://www.vedomosti.ru/technology/news/2018/05/08/768916-bespilotnik-uber-sbil> (дата обращения 15.02.2021).
3. Боты изобрели свой язык: почему Facebook испугался искусственного интеллекта? // BBC News. Русская служба. – 2017. – 31.07. – URL: <http://www.bbc.com/russian/features-40778454> (дата обращения 12.03.2021).
4. Деннетт Д. Насосы интуиции и другие инструменты мышления. – Москва : Литагент Corpus, 2019. – 570 с.

5. «Дипфейк» – информационное оружие будущего // Радио Свобода. Крым. Реалии – 2018. – 06.07. – URL: <https://ru.krymr.com/a/29347608.html> (дата обращения 02.03.2021).
6. Кадочникова С. Настали опасные времена. Захват власти и новая мировая война: как хакеры за год стали еще опаснее // Lenta.ru. – 2020. – 01.01. – URL: <https://lenta.ru/articles/2020/01/01/cybercrime2020> (дата обращения 02.02.2021).
7. Карасюк Е. «Ребята, вы обречены». Хасис и Варданян о технологической революции // Republic. – 2016. – 03.06. – URL: <https://republic.ru/app.php/posts/68944> (дата обращения 02.04.2021).
8. Малых В. О чем говорят машины: почему чат-боты пока не опасны для человека // РБК. – 2017. – 04.08. – URL: [http://www.rbc.ru/opinions/technology\\_and\\_media/04/08/2017/598453329a7947214bab8e50](http://www.rbc.ru/opinions/technology_and_media/04/08/2017/598453329a7947214bab8e50) (дата обращения 15.03.2021).
9. Мун Д.В. От «Титаника» до «Зимней вишни». Алгоритмы грядущих катастроф. – 3-е изд., доп. и перераб. – Москва ; Берлин : Директ-Медиа, 2018. – 352 с.
10. Мун Д.В., Попета В.В. From fake to Deepfake: угрозы и риски развития технологий искажения реальности в глобальном информационном пространстве // Культура: теория и практика. – 2020. – URL: <https://cyberleninka.ru/article/n/from-fake-to-deepfake-ugrozy-i-riski-razvitiya-i-rasprostraneniya-tehnologiy-iskazheniya-realnosti-v-global-nom-informatsionnom> (дата обращения 10.03.2021).
11. Носков А. Дипфейки станут неотличимыми от оригинала в течение 6–12 месяцев // ХАЙТЕК+. – 2019. – 24.09. – URL: <https://hightech.plus/2019/09/24/dipfeiki-stanut-neotlichimi-ot-originalov-v-techenie-6-12-mesyacev> (дата обращения 09.04.2021).
12. Погиб водитель Tesla ModelS с включенным автопилотом // Автомобиль Tesla. – 2016. – 03.07. – URL: <http://tesla.zone.ru/accidents/a-3071> (дата обращения 04.04.2021).
13. Содди Ф. Радий и его разгадка / пер. с англ. под ред. Д.Д. Хмырова. – Одесса : Mathesis, 1910. – 211 с.
14. Таиров Р. Маск назвал самый опасный для человечества искусственный интеллект // Forbes. – 2020. – 28.07. – URL: <https://www.forbes.ru/newsroom/milliardery/405911-mask-nazval-samyu-opasnyu-dlya-chelovechestva-iskusstvennyy-intellekt> (дата обращения 12.04.2021).
15. Указ Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» // Сайт Президента России. Документы. – 2019. – URL: <http://kremlin.ru/acts/bank/44731> (дата обращения 25.04.2021).
16. Уотсон Р. Факты будущего: История следующих 50 лет. – Москва : Эксмо, 2012. – 352 с.
17. Уэллс Г. Освобожденный мир. – Москва : Правда, 1964. – 494 с.
18. Харари Ю.Н. 21 урок для XXI века. – Москва : Синдбад, 2019. – 416 с. – URL: <https://mybook.ru/author/yuval-harari/21-urok-dlya-xxi-veka-2/citations/9634334/> (дата обращения 11.06.2021).
19. Хокинг С. Искусственный интеллект – величайшая ошибка человечества // Техномания. – 2014. – 05.07. – URL: <https://texnomaniya.ru/other-interesting-news/hoking-iskusstvenniy-intellekt-velichaiyshaya-oshibka-chelovechestva.html> (дата обращения 11.04.2021).
20. Хокинг С. «Мне кажется, компьютерные вирусы...» // Цитаты известных личностей. – 2021. – 13.01. – URL: <https://ru.citaty.net/tsitaty/636376-stiven-uiliam-khoking-mne-kazhetsia-kompiuternye-virusy-stoit-rassmatriva/> (дата обращения 03.04.2021).
21. A tragic loss // The Tesla Team. – 2016. – 30.06. – URL: [https://www.tesla.com/en\\_EU/blog/tragic-loss](https://www.tesla.com/en_EU/blog/tragic-loss) (дата обращения 04.04.2021).
22. Cole S. AI-assisted fake porn is here and we're all fucked // VICE media group. – 2017. – 11.12. – URL: [https://www.vice.com/en\\_us/article/gydydm/gal-gadot-fake-ai-porn](https://www.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn) (дата обращения 11.04.2021).
23. Jansen S.C., Martin B. The Streisand Effect and Censorship Backfire // International Journal of Communication. – 2015. – № 9. – P. 656–671.
24. Perrow Ch. Normal accidents. Living with high-risk technologies. – New York : Basic Books, 1984. – 386 p.
25. Powell K. NVIDIA researchers showcase major advances in deep learning at NIPS // NVIDIA. – 2017. – 03.12. – URL: <https://blogs.nvidia.com/blog/2017/12/03/nvidia-research-nips/> (дата обращения 11.04.2021).
26. Stankiewicz K. «Perfectly real» deepfakes will arrive in 6 months to a year, technology pioneer Hao Li says // CNBC. – 2019. – 20.09. – URL: <https://www.cnbc.com/2019/09/20/hao-li-perfectly-real-deepfakes-will-arrive-in-6-months-to-a-year.html> (дата обращения 11.04.2021).
27. Tucker P. The newest AI-enabled weapon: «Deep-faking» photos of the earth // Defense One. – 2019. – 13.03. – URL: <https://www.defenseone.com/technology/2019/03/next-phase-ai-deep-faking-whole-world-and-china-ahead/155944/> (дата обращения 02.03.2021).
28. Vincent J. Watch Jordan Peele use AI to make Barack Obama deliver a PSA about fake news // The Verge. – 2018. – 17.04. – URL: <https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peepe-buz-zfeed> (дата обращения 02.03.2021).
29. Wells H.G. The War in the Air. – London : G. Bell and Sons, 1908. – URL: <https://library.si.edu/digital-library/book/warair00wella> (дата обращения 01.03.2021).
30. You won't believe what Obama says in this video! / BuzzFeed Video // YouTube: сайт. – 2018. – 17.04. – 00:01:12. – URL: <https://youtu.be/cQ54GDm1eL0> (дата обращения 09.04.2021).

## A LITTLE ABOUT ARTIFICIAL INTELLIGENCE: CONTROVERSIAL TECHNOLOGIES

**Moon Dmitry**

PhD (Econ. Sci.) Deputy Director of the Federal State Budgetary Institution «Agency Emercom» MCHS of Russia, co-founder of the international expert community «RISK.TODAY», Moscow, Russia

**Popeta Vladislav**

PhD (Engineering Sci.), cofounder and President of the international expert community «RISK.TODAY», Moscow, Russia

***Abstract.** The article considers features of the use of artificial intelligence in various areas. Examples of ambiguous and not always constructive application of such technologies are presented. Attention is focused on the need to form systems and legal foundations of security working with AI technologies.*

***Keywords:** artificial intelligence; fake news; controversial technologies; risks; safety.*

***For citation:** Moon D.V., Popeta V.V. A little about artificial intelligence: controversial technologies // Social Novelties and Social Sciences. – Moscow : INION RAN, 2021. – N 2. – Pp. 26–37.*

URL: <https://sns-journal.ru/ru/archive/>

DOI: 10.31249/snsn/2021.02.02