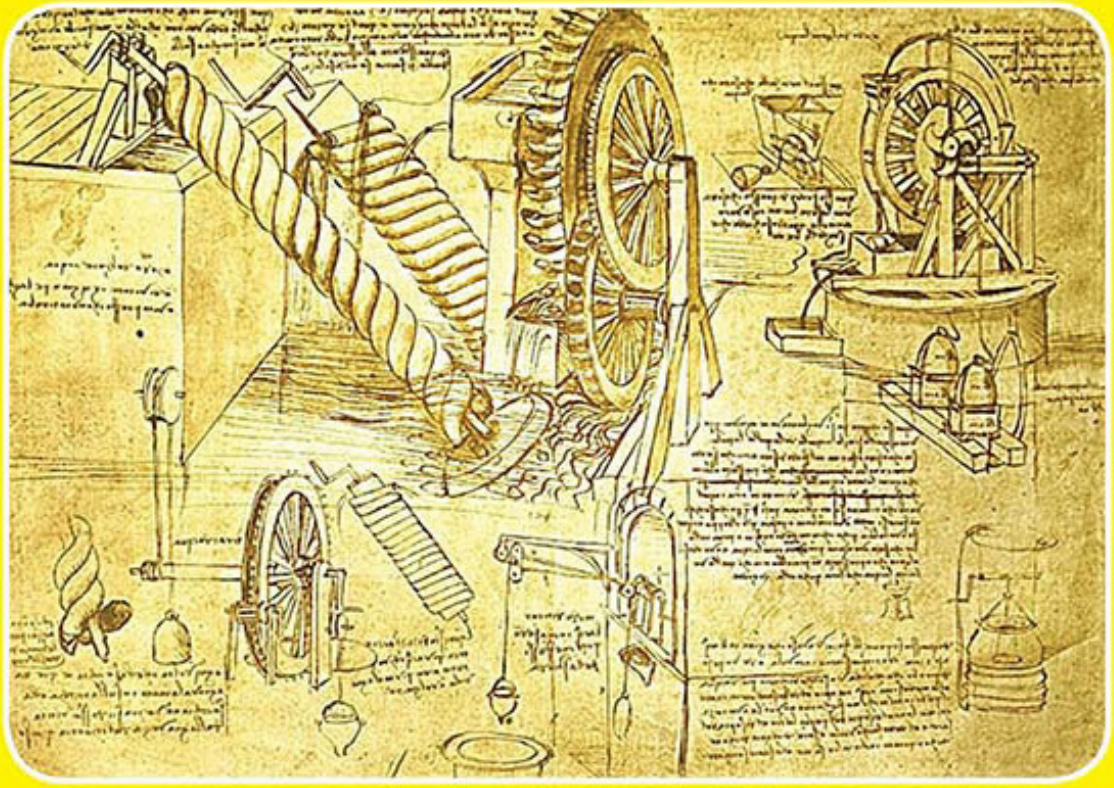
ВЛАДИМИР ОВЧИНСКИЙ
ЕЛЕНА ЛАРИНА

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ



БОЛЬШИЕ ДАННЫЕ

ПРЕСТУПНОСТЬ

Коллекция Изборского клуба

Владимир Овчинский

Искусственный интеллект. Большие данные. Преступность

«Книжный мир»
2018

Овчинский В. С.

Искусственный интеллект. Большие данные. Преступность / В. С. Овчинский — «Книжный мир», 2018 — (Коллекция Изборского клуба)

ISBN 978-5-6041495-7-7

Новые технологии меняют мир. Главным фактором изменений станет грядущая научно-техническая революция и внедрение искусственного интеллекта (ИИ) во все сферы деятельности человека.Но вместе с нашим миром изменяется и такая его неотъемлемая часть, как организованная преступность. Эксперты считают, что организованная преступность в течение следующего десятилетия подвергнется значительным изменениям, по существу, мафия превратится в своеобразную Crime Inc – глобальную криминальную многопрофильную корпорацию, использующую все новейшие достижения науки и техники, включая ИИ, дроны, социальный инжиниринг… Сможет ли противостоять этой корпорации «Зло» Интерпол и национальные полиции? Не придется ли ради победы над организованной преступностью отправить на свалку истории права человека и неприкосновенность личной жизни? Каким станет новый мир, где законопослушный гражданин окажется «голым» и перед государством, и перед преступником?

ISBN 978-5-6041495-7-7

Содержание

Елена Ларина, Владимир Обнинский Искусственный интеллект. Большие данные. Преступность

ВВЕДЕНИЕ

В последние годы ускоренными темпами развивались **искусственный интеллект (ИИ),** машинное обучение, распознавание образов и анализ **Больших Данных** (БД). Их развитие позволило создать большое число приложений, которые уже сегодня активно используются корпоративными и индивидуальными пользователями. Особый успех сопутствовал ИИ в таких сферах, как автоматизированное распознавание речи, машинный перевод, спам-фильтры и повышение качества поиска.

В ближайшие годы мировые расходы на технологии ИИ вырастут почти втрое. Прогноз сделан в марте 2018 г. в аналитической компании International Data Corporation (IDC). Эксперты ожидают, что уже в 2018 г. объем мирового рынка когнитивных систем и решений в области ИИ составит 19.1 млрд, долл., увеличившись более чем на 54 % по сравнению с 2017 г. В 2021 г. показатель достигнет

52.2 млрд, долл., а среднегодовые темпы роста будут измеряться 46,2 %. К 2019 г. около 40 % проектов цифровой трансформации, реализуемых в организациях, будут использовать ИИ-сервисы, а к 2021 г. три четверти корпоративных приложений будут иметь алгоритмы ИИ.

Ожидается, что в 2030 году доля ИИ в общемировом внутреннем валовом продукте (ВВП) составит до 15,7 трлн. долл. США. Более 50 процентов от этой суммы будет получено в результате повышения производительности труда. По имеющимся оценкам, ИИ повысит к 2035 году валовую добавленную стоимость экономики США приблизительно на треть[1].

В странах с более значительным ВВП работают более высококвалифицированные исследователи, которые могут адаптировать использование ИИ к местным потребностям более быстрыми темпами. Кроме того, в странах с более значительным ВВП компании, как правило, имеют доступ к более крупному рынку для продажи продуктов и решений на основе искусственного интеллекта. Следовательно, ожидается, что *в развитых экономиках с потенциалом ИИ произойдет рост благосостояния, тогда как менее развитые экономики наименее развитых стран, не имеющих выхода к морю развивающихся стран и малых островных развивающихся государств, могут оказаться еще более маргинализованными*[2].

Ничего мистического и непостижимого в ИИ и БД нет. Более того, согласно прогнозам подавляющего большинства ответственных исследователей и разработчиков ИИ, в перспективе 15–25 лет не может идти речи о создании так называемого «сильного» ИИ, превосходящего человеческий разум.

Сегодня, завтра и послезавтра ИИ – это вычислительная платформа для выполнения конкретных, заранее заданных функций и решения задач, общение с которым пользователя может осуществляться в звуковой или письменной форме на естественном языке или визуальным образом. По сути, ИИ – это устройство превращения любой – визуальной, акустической, текстовой и т. п. – информации в цифру, обработка этой цифры методами статистики и дискрет-

[1] PwC, «Sizing the prize: what's the real value of AI for your business and how can you capitalize?», 2017. Доступно по ссылке: www. pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf.

[2] Перспективы развития технологий в интересах устойчивого развития. ООН, ESCAP/76/16 (11–16 мая 2018 г.).

ной вычислительной математики и получение ответа в интуитивно понятном для человека виде.

Сегодня установлено, что значительную часть задач и проблем, в том числе в сфере национальной безопасности, эффективно решать на основе вычислений. Поэтому тот, кто сегодня располагает возможностями использования ИИ и БД, имеет решающее преимущество по сравнению с теми, кто такой возможности лишен. Если выбросить в мусорную корзину, а точнее стереть со смартфона бессмысленную информацию о конкуренции человека и машины в решении интеллектуальных задач, то для начала можно дать простое и короткое определение ИИ.

ИИ – это мощнейший усилитель человеческого интеллекта за счет обработки различными способами огромных массивов накопленных и постоянно пополняющихся БД абсолютно обо всем. В принципе, искусственный интеллект – это такой же усилитель разума, как любая машина – усилитель рук и ног.

В мире сегодня развернулась гонка по созданию наиболее мощных ИИ. В чем-то она напоминает гонку за обладанием ядерным оружием в 40-60-е годы прошлого века. Сегодня во главе гонки – корпорации и государственные структуры из США, Китая, Великобритании, Израиля. Франции.

Президент России В.В.Путин в День знаний 1 сентября 2017 года заявил: *«Искусственный интеллект – это будущее не только России, это будущее всего человечества. Здесь колоссальные возможности и трудно прогнозируемые сегодня угрозы. Тот, кто станет лидером в этой сфере, будет властелином мира».* В мае 2018 г., выступая на встречах с производителями вооружений и военной техники, с разработчиками робототехники и руководством РАН, В.В.Путин еще раз подчеркнул *безальтернативность для России качественного рывка в области разработки и внедрения ИИ.*

ИИ и БД определены как основные сквозные цифровые технологии в Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы, утвержденной Указом Президента России от 9 мая 2017 г. № 203, и Программе «Цифровая экономика Российской Федерации», утвержденной Правительством РФ от 26 июля 2017 г. № 1732-р.

В качестве *важнейших направлений развития ИИ, открывающих возможности и создающих угрозы в сфере национальной безопасности, эксперты выделяют три главных направления:*

во-первых, системы ИИ станут интегральными платформами для развития различных секторов и отраслей экономики, финансов, общественной жизни, государственного управления и национальной безопасности. От мощи ИИ будет в решающей степени зависеть конкурентоспособность стран, корпораций и отдельных групп граждан. Соответственно, следует ожидать попыток аутсайдеров в этом соревновании уровнять шансы за счет кражи систем ИИ либо нарушения их работы у лидеров;

во-вторых, чем дальше, тем больше системы ИИ будут использовать многомерные математикостатистические программы, в том числе для решения оптимизационных задач, где люди-пользователи будут получать эффективные результаты при неясном характере вычислений и их последствий. В перспективе это может создать риски для национальной безопасности;

в-третьих, наличие или отсутствие ИИ проведет между странами, корпорациями и людьми еще более явную границу, чем обладание или необладание огнестрельным оружием в XVI веке или авиацией – в начале XX. Различия между обладателями ИИ и теми, кто не имеет к нему доступа, могут разделить страны, народы и отдельные группы населения сильнее, чем когда-либо в истории человечества. В этих условиях *задачей номер один становится не просто создание и совершенствование ИИ, а ограничение числа его пользователей, прежде всего недопущение к нему террористических сетей и организованной преступность* I[3].

[3] См.: Artificial Intelligence and National Security. Belfer center paper. July, 2017.

Как подчеркнуто в исследовании Центра новой американской безопасности «Искусственный интеллект: тенденции, угрозы, рекомендации по минимизации рисков» (февраль 2018 г.)[4], *ИИ изменяет ландшафт рисков безопасности для граждан, организаций и государств.* Злонамеренное использование ИИ может угрожать цифровой безопасности, а также непосредственной физической безопасности отдельных граждан, групп, юридических лиц и даже государств в целом. Злонамеренное использование ИИ может нанести существенный вред цифровой инфраструктуре общества, бизнеса и правительства. На сегодняшний день важнейшим вопросом является прогнозирование возможных угроз злонамеренного использования ИИ. Это позволит оценить вероятность угроз и планировать меры по их предотвращению. Эти вопросы рассматриваются в первом разделе книги.

Использование ИИ в борьбе с преступностью неразрывно связано с *Большими данными (БД).* Поэтому в нашей работе подробно рассматриваются проблемы, связанные с БД в деятельности правоохранительных органов. Этому посвящен второй раздел книги.

В третьем разделе работы подробно анализируются проект Европола SOCTA, а также система **ePOOLICE,** которая осуществляется международной командой ученых под эгидой Евросоюза. В ePOOLICE эффективно реализуются технологии прогноза организованной преступности на основе БД и ИИ.

[4] См.: The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. February, 2018 (Center for new American Security).

Раздел I
Искусственный интеллект (ИИ), его использование преступниками и правоохранителями

§ 1. Сущность ИИ

Сегодняшний бум ИИ создает у многих иллюзию, что мы имеем дело с каким-то новейшим открытием. Это не так.

Попытки скопировать человеческий интеллект предпринимаются уже несколько веков.

ИИ часто связывают с *роботами.* Слово «робот» придумал чешский писатель *Карел Чапек* для своей *пьесы «Россумские универсальные роботы»,* написанной в 1920 году.

Некоторые источники утверждают, что первым создателем прообраза роботов был *Архит из Тарен-та* – это древнегреческий математик и механик. Он смог построить между 400 и 350 годами до нашей эры *деревянного парового голубя,* который мог покрывать расстояние в 200 метров.

Леонардо да Винчи изобрел механического человека, напоминающего *германского рыцаря,* в 1495 г. А несколько ранее, в 1478 году, он придумал *самоходную тележку,* которую многие современные эксперты считают первым в истории программируемым аппаратом. Вместо паровой энергии и двигателя внутреннего сгорания транспортное средство приводилось в движение заведенной пружиной. Тележка двигалась только вперед, но «оператор» мог повернуть ее колеса в определенные промежутки времени, поместив колышки в маленькие отверстия.

В 2004 году итальянские специалисты (дизайнеры, программисты, инженеры и плотники) собрались вместе, чтобы реконструировать автоматическую тележку Леонардо. У них это получилось.

В 1770 году венгерский изобретатель Вольфган фон Кемпелен создал автомат, который обыгрывал в шахматы всех, кто с ним состязался, и даже победил Наполеона Бонапарта. При этом для большего эффекта зрителям демонстрировались сложные механизмы машины. Позже изобретатель был разоблачен – внутри автомата сидел опытный шахматист, управляющий всем процессом.

В 1830-х годах английский математик *Чарльз Бэббидж* придумал концепцию сложного цифрового калькулятора – аналитической машины, которая, как утверждал разработчик, могла бы рассчитывать ходы для игры в шахматы. А уже в 1914 году директор одного из испанских технических институтов *Леонардо Торрес Кеведо* изготовил электромеханическое устройство, способное разыгрывать простейшие шахматные эндшпили почти так же хорошо, как и человек.

В 1954 году американский исследователь *Ньюэлл* решил написать программу для игры в шахматы.

К работе были привлечены аналитики корпорации RAND Corporation. В качестве теоретической основы программы был использован метод, предложенный основателем теории информации *Шенноном,* а его точная формализация была выполнена *Тьюрингом*[5].

С середины 30-х годов прошлого столетия, с момента публикации работ *Тьюринга,* в которых обсуждались проблемы создания устройств, способных самостоятельно решать различные сложные задачи, к проблеме ИИ в мировом научном сообществе стали относиться

[5] *Тьюринг Алан Мэтисон* – английский математик и криптограф, разработавший еще в 1936 году вычислительную «машину Тьюринга», которую в период Второй мировой войны использовали для расшифровки сообщений немецких войск.

внимательно. Тьюринг предложил считать интеллектуальной такую машину, которую испытатель в процессе общения с ней не сможет отличить от человека. Тогда же появился *термин BadyMachine – концепция, предполагающая обучение искусственного разума на манер маленького ребенка, а не создание сразу «умного взрослого» робота.*

Фундамент исследований ИИ был заложен также такими учеными, как *Норберт Винер,* который был пионером в области кибернетики, а также знаменитой работой *Уоррена Мак-Каллока* и *Уолтера Питтса* «Логическое исчисление идей, имманентных в нервной деятельности», опубликованной ими в 1943 году. Книга *У. Росса Эшби* «Конструкция мозга», которая была выпущена в 1952 году, явилась еще одной важной вехой в процессе рождения ИИ.

Летом 1956 года в Университете Дартмута в США прошла первая рабочая конференция с участием таких ученых, как *Маккарти, Минский, Ньюэлл, Саймон, Шеннон, Тьюринг,* и другие, которые впоследствии были названы *основателями сферы искусственного разума.* В течение 6 недель ученые обсуждали возможности реализации проектов в сфере ИИ. Тогда и появился сам *термин artificial intelligence – искусственный интеллект.* После знаменитой конференции в Дартмуте ИИ получил впечатляющее развитие. Были созданы машины, которые могли решать математические проблемы, обыгрывать в шахматы, и даже первый прообраз чат-бота, который мог разговаривать с людьми, вводя их в заблуждение по поводу своей осознанности.

Пожалуй, самое распространенное из них: якобы для создания ИИ надо точно разобраться с тем, как работает человеческий мозг и как он связан с сознанием. Современные специалисты в области ИИ полагают, что компьютеры ближе к арифмометрам и калькуляторам, чем к человеческому мозгу. Они работают на иных принципах, чем наш мозг. *Нейронные сети, о которых говорят компьютерщики, не имеют никакого отношения к нейронам человеческого мозга.* Наиболее совершенные нейронные сети имеют сегодня пять-шесть слоев и минимум синапсов. В человеческом мозге таких слоев сотни тысяч и миллионы. При этом, в одних областях компьютеры, как программно-аппаратные комплексы, уже сегодня превосходят людей, а в других – безнадежно уступают. *Поэтому человек и компьютер – принципиально разные устройства, также как человек не похож на компьютер, так и компьютер не похож на человека.*

Здесь нельзя не согласиться с профессором факультета философии Оксфордского университета, основателем и директором Института будущего человечества *Ником Бостромом,* который пишет: *«Нейрокомпьютерные интерфейсы вряд ли станут тем вариантом, который приведет нас к сверхразуму. Усовершенствование сетей и организаций может в долгосрочной перспективе привести к появлению слабых форм коллективного интеллекта, но более вероятно, что оно сыграет стимулирующую роль, как и биологическое улучшение интеллектуальных способностей, постепенно повышая эффективность умственной деятельности людей при решении интеллектуальных задач»*[6].

За истекшие 60 лет сложилось три основных направления определения ИИ. Эти различные понимания – не отвлеченные рассуждения. На программы, построенные на каждом из них, потрачены миллиарды долларов.

Обычно, когда дело касается ИИ, все вспоминают знаменитый *тест Тьюринга.* С тестом связано первое направление определения ИИ. Суть теста в следующем: *если при общении с компьютером посредством анонимного канала связи нельзя понять, с кем идет беседа – с человеком или машиной – то такой компьютер можно считать ИИ.* Грубо говоря, ИИ – это интеллект, похожий на человеческий, по результатам действий, т. е. по поведению. Долгое время всех удовлетворяло такое понимание ИИ. Собственно, знаменитый Watson – это и есть реализация на практике программно-аппаратного комплекса, способного пройти тест

[6] *Бостром, Ник.* Искусственный интеллект. Этапы. Угрозы. Стратегии. М., 2016, с. 93; см. также: Что мы думаем о машинах, которые думают. Ведущие мировые ученые об искусственном интеллекте. Под ред. Дж. Брокмана. М., 2017.

Тьюринга. Watson, кстати, породил нынешний бум ботов. Боты призваны вести элементарную беседу. Они используются сегодня во многих странах повсеместно – от торговых площадок до больниц, от полицейских участков до справочных служб.

Другое направление ИИ связывается со *способностью программ к самосовершенствованию.* Не случайно, что о нейронных сетях, глубоком обучении и ИИ заговорили одновременно. На самом деле известны они были примерно те же 60 лет. Главная проблема была в дороговизне железа, т. е. самих компьютеров, способных выполнять эти программы. Нейронная сеть может эффективно решать конкретные задачи, но при этом никогда не пройдет теста Тьюринга.

Большинство практиков использует третье понимание ИИ. Это – *программно-аппаратный комплекс, работающий с использованием нейронных сетей, глубокого обучения и способный общаться с человеком на естественном языке, в том числе посредством голоса.*

Если подходить с инженерной точки зрения, то необходимо понять, где компьютеры сильнее людей, и что нам от них нужно. Посмотрим на эту проблему на примере анализа ФБР о провалах и успехах ИИ, связанных с борьбой с криминалом.

На сегодняшний день успехи достигнуты там, где имеются огромные массивы БД, ограниченное время для их анализа и возможность написать программу анализа. Грубо говоря, **компьютер превосходит человека там, где имеет место огромная комбинаторика, т. е. наличие множества вариантов, короткое время исполнения и возможность вести анализ чего-либо путем выполнения последовательных операций, т. е. возможность написать алгоритм.**

Где на сегодняшний день отмечены наибольшие прорывы? В анализе БД, распознавании образов, поиске незаметных на первый взгляд связей и закономерностей. Отсюда возникает простое заключение. Если бы человек имел бесконечное время на решение той или иной задачи, был дисциплинирован и имел неограниченный объем памяти, то он бы успешно решал все задачи, где компьютер уже сегодня первенствует над человеком. Самые знаменитые достижения компьютеров, подаваемых как ИИ, связаны с победой в играх – от шахмат до го, от покера до «бесконечных шашек». Любая игра имеет правила. А там где есть правила, путь к успеху лежит в комбинаторике и написании алгоритмов.

Приведенные соображения позволили информационным подразделениям ФБР совместно с Лабораторией искусственного интеллекта корпорации Google выработать следующее инженерное определение ИИ. Именно оно положено в разработку концепции архитектуры и перечня программных решений ФБР: *ИИ – это программно-аппаратный комплекс, обеспечивающий поддержку и/или принятие результативных решений в динамичной, неустойчивой среде в установленное время, на основе заведомо неполной, нечеткой и не имеющей полной доказательной базы информации.* Применительно к одним задачам ИИ самостоятельно принимает решения, но в большинстве случаев является элементом гибридного интеллекта, взаимодействуя с человеком.

Данное определение является инженерным по трем обстоятельствам. Прежде всего, оно задает критерий. Во-первых, результативность решений не носит абстрактного характера, а определяется в каждом конкретном случае постановщиком задачи. В одних случаях у ИИ может отсутствовать право на единственную ошибку, а в других – результативным решением может оказаться показатель, выше уже сложившегося уровня успешности решения проблемы.

Во-вторых, данное определение не привязывается к конкретным видам харда или софта. Возможно, завтра у нас появятся полноценные квантовые компьютеры. В университете Нотр-Дам уже сегодня идут эксперименты по использовании в качестве элементной базы компьютера живых бактерий. То же самое с софтом. Было бы самонадеянным утверждать, что и завтра вычислительные комплексы будут использовать машинное обучение и нейронные сети. Наконец, третий, принципиальный момент в определении – это то, что ИИ обязан научиться работать с неполной и частично лживой информацией. Это, пожалуй, самая сложная проблема.

Термин «ИИ» зачастую заменяет такие сложные и непонятные лицам, принимающим решения, термины, как нейронный сети, глубокое машинное обучение, дискриминантный анализ, многомерная статистика, вычислительная лингвистика и т. п. Согласно данным контент-анализа, приведенного Стэнфордским университетов в 2017 г., ИИ не в социальных СМИ, а в научных изданиях используется как синоним того или иного математико-статистического метода. Условно назовем это *маркетинговым использованием термина ИИ.* Наиболее широко это явление проявилось в англосаксонских странах, прежде всего США и Великобритании.

Для европейских публикаций и исследований характерно другое использование термина ИИ. *В Европе, особенно в Германии и во Франции, ИИ по сути стал синонимом любых сложных экспертных систем, в основе которых лежит блок поиска, обработки и анализа информации.* Такое понимание ИИ связано с тем, что в силу целого ряда факторов в большинстве стран ЕС не получили широкого развития наиболее современные методы дискретной математики и работа идет в направлении совершенствования информационно-аналитических систем, которые были созданы в конце XX – начале текущего века.

Свое понимание ИИ имеется в *Японии,* одной из трех лидирующих стран в этой сфере. *Они понимают под ИИ программы, которые могут выполнять интеллектуальные функции человека вне зависимости от сферы их применения.*

Авторы уже указанного доклада Центра новой американской безопасности, понимают ИИ также как автор этого термина – знаменитый математик, кибернетик, создатель множества языков программирования *Джон Маккарти.* Он определили ИИ, как *«вычислительные методы, позволяющие решать нечеткие и противоречивые задачи в условиях многокритериального выбора и хронической неполноты информации».*

С учетом разработок в области когнитивных вычислений, осуществленных уже после смерти Маккарти, Центром новой американской безопасности предлагается следующее, наиболее общее и в то же время рабочее определение: *ИИ – это программноаппаратные вычислительные комплексы полного информационного цикла (включающего восприятие, фильтрацию, обработку, хранении информации, выполнение аналитических и синтетических когнитивных функций), позволяющие в режиме человек-машина или автономно принимать и реализовывать решения в сложной, динамичной и неопределенной среде.*

Данное определение подчеркивает несколько ключевых концептов, без понимания которых лица, принимающие решения, не смогут сделать правильного выбора:

Во-первых, ИИ – это не машинное обучение, не нейронные сети и не другие виды программных продуктов. Это – всегда программно-аппаратные комплексы, в которых роль физических компонентов, как минимум, не меньше, чем информационных.

Большая часть программных методов ИИ была хорошо известна еще в 60-70-е гг. прошлого века, однако не могла быть реализована, поскольку компьютеры не обладали необходимым быстродействием. Тенденцией последних лет стало стремление ключевых производителей ИИ как можно больше наиболее важных функций упаковывать в непосредственно встроенный процессор и софт. Такая архитектура гарантирует производителям ИИ монополию на рынке. Кроме того, дальнейший скачок в области ИИ связан с изменением типов процессоров. Уже сегодня на смену кремнию приходят *квазиквантовые компьютеры.* Известно об успешной разработке и опробовании в Израиле *сверхскоростного графе-нового компьютера.* Известно, что в ряде стран мира были проведены успешные испытания *процессоров для ИИ, созданных на алмазной основе,* процессоров с так называемой «алмазной подложкой». Наконец, в 2017 г. команда Стэнфордского университета, университета г. Сеул, компании «Сименс» и Израильского технологического университета смогли создать работающий в реальном режиме *биологический компьютер,* где в качестве процессора используются моле-

кулы. Т. е. *будущее ИИ связано не столько с программами, сколько с прогрессом в области аппаратной части и новыми типами процессоров.*

Во-вторых, полный цикл обработки информации в настоящее время осуществляется преимущественно на базе комбинаторных методов, глубокого обучения и нейронных сетей. Однако все три метода страдают тем недостатком, что успешно могут работать только с конечными задачами. Наиболее яркий пример конечной задачи – это любая игра, где наперед задано все количество возможных ситуаций и комбинаций, возникающих в ходе игры. Но следует считать, что комбинация нейронных сетей с машинным обучением в ближайшие годы будет господствующей в вычислительной технологии ИИ.

Наконец, *в-третьих,* тенденцией взаимодействия человек-машина в рамках ИИ является повышение уровня автономии ИИ, т. е. возложение на него частично или в полном объеме принятия решений. Это особенно ярко проявляется в военной и финансовой сфере, где счет идет на миллисекунды и соответственно вычислительная реакция превосходит человеческую.

Как уже отмечалось, в настоящее время ИИ используется, прежде всего, *для распознавания образов, прогнозирования и управления сложными системами.* Однако в принципе *ИИ может быть ориентирован на любые задачи, которые в настоящее время решают люди.* При этом необходимо оговориться, что *ИИ способен подменить людей в настоящее время только в рамках имитационных, функциональных и операционных задач.* Это означает, что *ИИ применяется лишь тогда, когда извне ему ставится четкая задача, которая может быть выполнена в рамках наперед заданной последовательности шагов или операций.* При этом сама задача носит имитационный, т. е. воспроизводимый с образца характер. Творческие задачи с созданием нового ИИ, по крайней мере на сегодняшний момент, решать не готов.

В начале XXI века за счет мощных программно-аппаратных комплексов ИИ стал *распознавать изображения* с 98 % точностью и делает он это лучше, чем человек, который распознает изображения с точностью до 95 %. Кроме того, впервые системы ИИ научились *создавать синтетические изображения,* которые практически неотличимы от оригинальных фотографий. Появилась возможность *создания несуществующих личностей,* которые, по крайней мере, в информационном пространстве могут жить полноценной жизнью, осуществляя с помощью чат-ботов коммуникации с людьми, информируя о своей жизни через ролики в YouTube и т. п. Согласно проведенным экспериментам, люди распознают ошибку, т. е. определяют искусственный характер изображения лишь в *3* % случаев из 100 %.

Системы ИИ добились впечатляющих результатов в конечных конкурентных играх: от шахмат до игры в го. В 2017 г. ИИ впервые победил человека в игре, где наряду с комбинаторикой требовалась рефлексия позиции, а именно – в покере. Методы ИИ в последние годы обеспечили прорыв в переводе. Другие направления задач, где осуществляется быстрый прогресс, включают в себя распознавание речи, автомобильную навигацию и прогнозирование биржевых процессов.

Успехи ИИ связаны с тремя основными факторами. *Во-первых,* с использованием новой высокопроизводительной элементной базы. *Во-вторых,* с применением новых программных решений, базирующихся на сложной комбинаторике и машинном обучении. *В-третьих,* с широким использованием робототехники как периферийных устройств ИИ, аналогичным периферийным устройствам человека, типа рук, ног, по отношению к мозгу.

Хотя в последние 10 лет ИИ развивался экспоненциально, вряд ли следует ожидать таких же темпов прогресса и в перспективе. Как правило, технические нововведения развиваются по *гиперциклу Гартнера.* При гиперцикле после долгого периода созревания наступает этап экспоненциальных перемен. В результате система достигает уровня зрелости и определенное время оказывается как бы на плато, раздвигаясь вширь, а не развиваясь вглубь. Затем наступает спад, связанный с насыщением данной технологией наиболее продвинутых пользователей. Однако спад является недолговременным и сменяется умеренным ростом, который характе-

рен для любой зрелой технологии. Вряд ли есть основания полагать, что ИИ не будет развиваться в рамках гиперцикла. Сегодня центральной задачей ИИ является *создание эффективных гибридных систем, где ИИ взаимодействует с человеком.*

§ 2. ИИ, распознание угроз и оценка рисков

Магистральным направлением использования ИИ являются вопросы безопасности. При решении этой группы вопросов как в никакой другой сфере важно заблаговременно *распознавать угрозы и оценивать риски.* Распознавание угрозы мало чем отличается от распознавания лица. Любая угроза имеет определенный устойчивый паттерн, который может быть выражен через набор числовых характеристик. Поскольку вопросы распознавания в решающей степени зависят от скорости и полноты вычислений, то ИИ как комбинаторная машина, позволяет распознавать угрозы намного быстрее и точнее, чем человек.

Правда, есть одно важное ограничение. *ИИ способен распознавать лишь те угрозы, которые имели место в прошлом.* Поскольку в основе распознавания лежит машинное обучение, то фактически ИИ на числовых массивах прошлого устанавливает профиль угрозы, а потом ищет этот профиль в поступающих информационных потоках.

До сих пор остается открытым вопрос, может ли человек распознавать угрозы, которых ранее не существовало. На этот счет имеются различные точки зрения. Большинство психологов занимают точку зрения, что человек способен к этому. В то же время специалисты когнитивных наук полагают, что нет принципиальной разницы между переработкой информации у машины и человека, и соответственно, человек не может решать задачи, которые не решает машина.

Авторы доклада Центра новой американской безопасности полагают, что *человек обладает способностью к решению задача, не доступных, по крайней мере, в настоящее время.* Например, *человек способен изменить правила игры, в то время как ИИ всегда играет по правилам.* Однако применительно к новым, ранее не существовавшим угрозам, на сегодняшний день не существует однозначного ответа на вопрос: способны ли люди распознавать угрозы, с которыми до этого никогда не сталкивались.

Создание ИИ носит феноменальный характер. Существует множество различных программноаппаратных комплексов, каждый из которых уникален, а потому феноменален. В отличие от персональных компьютеров, планшетов, смартфонов и т. п. ИИ носят единичный, в крайнем случае, мелкосерийный, но отнюдь не массовый характер. Если явление не носит массового характера, то оно не может быть описано количественно. Соответственно прогноз тенденций в области ИИ – это всегда качественный прогноз.

Другое дело, что отдельные аспекты этого качественного процесса могут иметь количественное выражение, типа знаменитого *закона Мура*[7]*, и* тем самым служить ориентиром прогнозирования. При определении тенденции развития сложных явлений, в т. ч. *ИИ, сегодня наиболее широко используют Форсайт метод, или сценарное прогнозирование.* При том, что форсайт прогнозы являются сегодня абсолютно преобладающей формой составления прогнозов, разработки плановых программ на государственном уровне, включая США, страны ЕС, эффективность их крайне низка. Достаточно привести два примера. С 2003 по 2010 гг. было осуществлено на уровне федеральных органов власти США, исследовательских подразделений Федеральной резервной системы (ФРС) более 15 форсайт прогнозов дальнейшего развития глобальной финансовой системы. Ни в одном из прогнозов не нашли своего места криптовалюты и цифровые активы. Ни в одном прогнозе не были упомянуты смарт-контракты и цифровые монеты.

[7] Гордон Эрл Мур (США) – почетный председатель совета директоров и основатель корпорации *Intel,* основоположник *«закона Мура»,* который сводится к тому, что количество транзисторов в кристалле микропроцессора удваивается каждый год. В 1975 году он изменил временную составляющую закона и заявил об удвоении количества транзисторов каждые два года.

Это неудивительно. Во всех странах мира форсайт составляют статусные люди, которые плоть от плоти сложившейся системы. Соответственно, они видят в будущем линейное, но масштабируемое продолжение настоящего. А это принципиально не так.

В этой связи Центр новой американской безопасности предлагает использовать *локусный подход к прогнозированию.* Он состоит в том, что в рамках среднесрочного прогноза на горизонте три-пять лет верна мысль известного американского фантаста и мыслителя У.Гибсона: *«Будущее уже наступило.*

Просто оно пока неравномерно распределено». Для среднесрочных прогнозов локусный подход является не только наиболее эффективным, но и максимально дешевым и простым. Используя отработанные методы распознавания образов и обнаружения аномалий, осуществляется сканирование ноу-хау, разработок, гипотез в той области или сфере знания, применительно к которой осуществляется прогнозирование. Это позволяет выявить локусы будущего, а затем собственно прогноз сводится к тому, чтобы постараться оценить реалистично темпы экспансии этих локусов, как правило, находящихся на периферии, в ключевые сектора мировой и национальных экономик.

Тенденции всегда проявляют себя как возможности, т. е. варианты развития будущего. Практически все эмпирические исследования в области социальной динамики показывают, что у групп действия существует не один, а несколько вариантов поведения практически в любой ситуации.

Любое лицо, принимающее решение, заинтересовано в *снижении риска.* Собственно, ИИ и является мощнейшим инструментом подавления рисков. Однако это относится к гносеологическим рискам. Они минимизируются за счет получения дополнительной информации и ее глубокой обработки, позволяющей гораздо более достоверно, чем раньше судить о движущих силах и логике той или иной ситуации. Что же касается онтологического риска, то ИИ бессилен перед ним. В конечном счете, *ИИ – это мощнейший многофункциональный вычислитель.* Если же параметры, которые он вычисляет, предельно нестабильны, носят дискретный, а не непрерывный характер, находятся в состоянии, близком к *белому шуму*[8], то даже самый мощный ИИ не сможет оказать большой помощи лицу, принимающему решения.

Использование ИИ позволяет гораздо более реалистично, чем раньше, заблаговременно определить экзистенциональные угрозы, а также позволяет в режиме мониторинга сканировать угрозы со стороны другого участника конфликта.

[8] *Термин «белый шум»* обычно применяется к сигналу, имеющему *автокорреляционную функцию.* Белый шум некоррелирован по времени (или по другому аргументу), не определяет его значений во временной (или любой другой рассматриваемой аргументной) области.

§ 3. ИИ как технология тройного назначения

ИИ – это технология тройного назначения. ИИ может быть использован как для *гражданских,* так и для *военных* целей. Отдельное направление использования ИИ – *мафиозно-террористическое.* Поскольку некоторые задачи, требующие интеллекта, являются доброкачественными с точки зрения права, а другие – нет, то ИИ обладает свойством тройного использования, также как и человеческий интеллект.

О гражданском, мирном использовании ИИ СМИ сообщают буквально каждый день. Но, откровенно говоря, самое активное использование ИИ наблюдается в **военных целях.**

Например, Министерство обороны США изучает множество разнообразных направлений использования ИИ. Эта работа ведется в основном в рамках *DARPA* (Управление перспективных исследовательских проектов Минобороны США) и *IARPA* (Агентство передовых исследований в сфере разведки). Разработкой стратегии использования ИИ в сфере национальной безопасности и координации исследований занимается Канцелярия помощника Министра обороны по исследованиям и инженерии, а сам помощник несет личную ответственность перед министром обороны, администрацией президента и Конгрессом за максимально эффективное использование ИИ в интересах национальной безопасности[9].

В апреле 2017 г. под руководством заместителя министра обороны США по разведке создана и начала активно работать междисциплинарная и многофункциональная команда по разработке стратегии и тактики *алгоритмических войн,* а также их программно-аппаратному обеспечению со стороны ИИ. Работа этой команды известна как *проект Maven.* Главная цель проекта Maven состоит в максимально быстром внедрении ИИ в оборонительные и наступательные системы в сфере военного, финансово-экономического и поведенческого противоборства. Проект призван продемонстрировать огромный потенциал технологий ИИ. В рамках проекта на период до 2020 г. поквартально расписаны цели и ресурсы. Информация по проекту Maven доступна комитетам Сената и палаты Представителей по разведке, т. к. относится к засекреченной сфере.

В начале 2018 г. директор проекта Maven заявил: «Maven предназначен для того, чтобы быть пилотным проектом. Он призван продемонстрировать неисчерпаемый потенциал ИИ в сфере алгоритмических войн, а конкретно кибер-, финансово-экономических и поведенческих конфликтов и противоборств, а также в сфере управления и прогнозирования конфликтов на пяти полях боя: на земле, в воздухе, в космосе, под водой и в киберсреде».

Ожидается, что к 2020 г. ИИ даст максимальный эффект в разведке для обработки и анализа больших, в том числе неструктурированных, зашумленных и неполных. Одним из результатов проекта Maven стало *создание системы опережающего мониторинга и прогнозирования на основе разнообразных данных действий противника* (на примере борьбы с ИГИЛ). Система Cointer-ISIL-Maven начала эксплуатироваться с июля 2017 г., она включает в себя сложный программно-аппаратный комплекс, состоящий как из периферийных систем, так и центрального ИИ. В качестве периферийных систем используются автоматизированные дроны, оснащенные системами компьютерного оптического зрения. Среди принципиально новых модулей центрального ИИ, созданного в рамках проекта, необходимо отметить гибкие модифицированные блоки нейронных сетей с машинным обучением, позволяющих распознавать нечеткую оптическую информацию на уровне более высоком, чем наблюдатели-люди.

Помимо засекреченных, у разведывательного сообщества есть несколько публично рекламируемых исследовательских проектов в области ИИ. На начало 2018 г. только в интересах ЦРУ осуществляется 137 публично финансируемых проектов, связанных с ИИ. В основном

[9] См.: Artificial Intelligence and National Security. Congressional Research Service. 26.04.2018.

эти проекты направлены на решение таких задач, как анализ разнородной структурированной и неструктурированной разноформатной, зашумленной и неполной информации. Более 2 5 проектов связаны с использованием ИИ, в том числе в составе симбиотического интеллекта, совместно *с группами экспертов для прогнозирования будущих событий,* таких как террористические атаки, гражданские беспорядки, финансово-экономические, политические и военные кризисы и т. п.

IARPA в настоящее время финансирует крупнейший в истории США *проект по созданию человеко-машинной платформы симбиотического (гибридного – человек + ИИ) интеллекта для распознавания слабых сигналов в информационном шуме и прогнозирования маловероятных событий.* Также ИИ активно используется для разработки алгоритмов одновременного многоязычного распознавания речи и перевода акустической речи в тексты с уровнем, превосходящим применяющиеся в настоящее время системы машинного перевода.

У ИИ может быть многообещающее будущее *в сфере военной логистики.* Например, ВВС США работает над использованием ИИ для составления графиков обслуживания летательных средств, включая графики дозаправки в воздухе и проведения ремонта. Вместо того, чтобы осуществлять дорогостоящий ремонт, когда самолет или вертолет выходит из строя из-за поломок, ИИ разработал модели, позволяющие проводить предупредительное техническое обслуживание воздушных судов. Это повышает надежность их эксплуатации при более низких затратах. Данная система, созданная в 2017 г., включает в себя встраиваемые в воздушные суда датчики, передающие шифрованные сигналы центральному интеллекту, в котором они становятся базой для работы алгоритма прогнозирования.

В сентябре 2017 г. Управление материально-технического снабжения сухопутных войск США подписало второй контракт с IBM на сумму 135 млн. долларов для создания *персонального электронного помощника бойца штурмового отряда на базе ИИ.* Этот проект стал продолжением первого проекта, начатого в 2014 и завершенного в 2016 г. В рамках первого проекта электронный индивидуальный помощник-эксперт был создан для работников полевых штабов дивизий быстрого развертывания на базе IBM Watson.

ВМС США заказали в 2017 г. *версию Watson, предназначенную для разработки планов оптимального материально-технического снабжения военно-морских группировок и отдельных судов, находящихся в мировом океане, и контроля над их выполнением.* Командование сухопутной армии полагает, что использование логистического Watson в армии обеспечит ежегодную экономию 100 млн. долларов за счет оптимального распределения логистических потоков и планов материально-технического обеспечения вооруженных сил.

Наиболее активно ИИ будет использоваться министерством обороны США *в киберпространстве.*

В 2018 г. киберкомандование США разместило через DARPA заказы по использованию ИИ для мгновенного обнаружения аномалий и дыр в киберзащите. Представляется, что именно ИИ с его быстродействием позволит наиболее эффективно управлять боевыми киберплатформами на самой деликатной стадии киберпротивоборств – фазе проникновения в сети противника.

Вооруженные силы США стремятся максимально использовать ИИ *в области управления и контроля.* Наиболее продвинутая система создана в настоящее время в ВВС США. Сейчас она доведена до уровня штабных работников командования ВВС. В период до 2019 г. система охватит уровень авиационных полков и дивизий.

Как известно, одной из наиболее сложных в практическом плане задач является сохранение управляемости и поддержание взаимодействия командования и боевых единиц в ходе реальных военных действий, когда противник наносит удары не только на поле боя, но и по центрам командования. До настоящего времени ни в одной стране мира, насколько известно, не создана система регенерации командования и контроля в жестких конфликтах. Регенера-

тивная система должна быть организована таким образом, чтобы после выхода из строя тех или иных узлов и уровней командования, система перестраивалась и в новой конфигурации сохраняла высокий уровень управления и координации. В настоящее время командование ВВС совместно с корпорацией Lockheed Martin и корпорацией Alphabet приступили к созданию такой системы на основе симбиотического интеллекта, используя традиционные командные центры и защищенный ИИ.

Все рода войск США в последние годы *имплантируют ИИ в различные типы автономных транспортных средств.* По сути, вооруженные силы ведут работу параллельно с бизнес-сектором по созданию транспортных средств с полным самообслуживанием. Военные подрядчики вооруженных сил, начиная с 2017 г. ежегодно представляют такого рода автономные транспортные средства с использованием ИИ. С 2019 г. министерство обороны запускает проект стоимостью в 430 млн. долларов по созданию систем, включающих центральный ИИ и роевые или стайные автономные транспортные средства, оснащенные датчиками и интерфейсами, позволяющими перейти от индивидуального к коллективному машинному обучению.

Исследовательская лаборатория ВВС завершила вторую фазу испытаний по программе «Недоверчивый Уитман». В рамках программы впервые создан и проходит испытания *полноценный беспилотный истребитель пятого поколения.* В 2017 г. тестовый вариант, реализованный на более дешевом истребителе F16, прошел испытание. В их ходе машина, оснащенная ИИ, автономно реагировала на события, которые не были включены в программу полетов и представляли собой непредвиденные препятствия и сложности для выполнения заданий. Из 17 испытательных заданий в 16, не считая самого первого, платформа справилась со всеми сложностями. Уже сегодня очевидно, что ИИ позволяет создавать полностью функциональные боевые истребители и самолеты-штурмовики, не уступающие, а по ряду параметров превосходящие такие же самолеты, пилотируемые людьми.

По сути это представляет собой следующий шаг *после массового внедрения в военную практику дронов – робототехнических летательных комплексов с ограниченными огневыми возможностями.* Кроме того, по заданию ВВС в настоящее время завершается отработка комплексных авиационных звеньев, которые предусматривают патрулирование и ведение боевых действий группой самолетов, один из которых пилотируется человеком, а несколько – системами с ИИ. В этом случае человек может в определенных случаях изменить команды ИИ. Данная система разрабатывается ВВС, поскольку командование авиацией, по крайней мере, в настоящее время и в ближайшем будущем не готово доверить решение о применении тактического ядерного оружия, которым оснащены многоцелевые истребители-бомбардировщики, роботам.

Сухопутные войска и Корпус морской пехоты США испытали прототипы автономных транспортных средств, а том числе оснащенных средствами огневого поражения. В ходе действий Сил специального назначения США в Ираке, Афганистане и Сирии в 2017 г. сухопутные войска уже активно применяли в боевых условиях роботизированные автономные эвакуационные машины.

Корпус морской пехоты в 2018 г. начнет принимать на вооружение *многофункциональный универсальный роботизированный тактический транспорт.* Роботизированное с элементами ИИ транспортное средство грузоподъемностью от одной до трех тонн будет следовать за ротами и взводами морских пехотинцев по местности с любым рельефом и любой сложности. Средства предназначены для транспортировки любых грузов – от запасных патронов и снарядов до пищи и одеял. Несколько аналогичных средств в настоящее время разрабатываются и для сухопутных вооруженных сил.

Скорее всего, сама логика развития ИИ ведет к тому, что возможности тройного использования будут постоянно увеличиваться. Многие задачи, которые в настоящее время автоматизируются, являются по своей сути тройственными. Например, ИИ, анализирующий программ-

ное обеспечение на уязвимости, может выполнять функции киберзащиты, кибердиверсий и киберпреступлений. Все то же самое относится к стаям дронов, находящихся под управлением ИИ. Команда инженеров из *Массачусетского технологического института (MIT)* с помощью ИИ смогла создать устройство, которое способно следить за человеком через стены. Причем устройство получилось совсем не громоздким.

Прибор получил название *RF-Pose.* Для того, чтобы обучить алгоритм новым фокусам, сотрудники MIT отслеживали движения людей при помощи радиолокатора и видеокамеры. Они фиксировали ходьбу, беседу между людьми, позу сидя, стоя и позу ожидания, а также открывание дверей. Затем при помощи не очень сложной компьютерной программы изображение было преобразовано в скелетную модель каждой зафиксированной ситуации. Эти модели вместе с радиосигналом изучил ИИ, и таким образом он обучился распознавать связь между показаниями радиолокатора и тем, какие действия производил в данный момент человек или группа людей.

В итоге специалисты MIT создали алгоритм, который практически в реальном времени может показывать перемещения людей за стеной или другим препятствием. Стоит заметить, что на данный момент система способна «выдавать» лишь двухмерное изображение, но со временем систему можно будет оптимизировать таким образом, чтобы она смогла производить мониторинг в трех измерениях.

В планы создателей также входит научить ИИ распознавать более сложные движения вроде мелкой моторики кистей и пальцев, так как сейчас возможно лишь наблюдение за перемещением конечностей и туловища. Сами авторы утверждают, что их система может найти применение в самых различных сферах. От вполне логичных наблюдений за местами массового скопления людей до медицинских учреждений и мест лишения свободы.

Рассмотрим теперь более подробно **криминальные угрозы ИИ.**

Развитие ИИ происходит в условиях увеличения анонимности. Многие задачи включают в себя общение с другими людьми, наблюдение за ними, принятие решении, воздействующих на их поведение. Согласно экспертным оценкам, до 90 % подобных задач могут быть автоматизированы и на горизонте пяти лет переданы ИИ. При этом соединение ИИ и интернета всего создают поистине безграничные возможности для злонамеренной деятельности, вплоть до убийств, которые правоохранительным органам весьма трудно квалифицировать именно как убийство. Проблема мира интернета в том, что в нем любое целенаправленное действие может быть замаскировано либо под отказ, либо нерегламентную работу оборудования. Разделить их никто не сможет.

Развитие ИИ открывает широчайшие перспективы для преступности. Уже сегодня злоумышленники могут без каких-либо препятствий приобрести как программные, так и аппаратные компоненты самых мощных систем ИИ. Более того, *широкое использование открытого кода в ИИ позволяет преступникам без каких-либо затрат средств получать доступ к последним разработкам ведущих компаний.*

В настоящее время алгоритмы ИИ создаются в течение месяцев, а не лет. Это достигается за счет того, что над ними работают не закрытые коллективы специалистов, а открытые сообщества программистов. Соответственно, преступники имеют возможность бесплатно получать новейшие разработки. Министерство внутренней безопасности США пыталось законодательно ограничить распространение определенных разработок в области ИИ, но потерпело поражение. Программисты просто отказались закрыть для себя возможность создавать системы с открытым программным кодом.

Сегодняшние системы ИИ страдают от ряда новых неурегулированных уязвимостей. К ним относятся *атаки через фиктивные данные,* используемые при обучении, а также *злонамеренное вмешательство в работу нейронных сетей.* Эти уязвимости не имели аналогов в

прошлом, поэтому не очень понятно, как с ними справляться. Есть основания полагать, что в будущем атаки на ИИ будут более эффективными, более точными, более сложными для атрибуции и будут использовать в основном уязвимости в программном обеспечении ИИ.

Проведенные эксперименты показывают, что более 70 % атак на ИИ осуществляется с использованием ИИ. Соответственно, в будущем, по мере широкого распространения ИИ, следует ожидать нарастания количества подобных атак экспоненциальными темпами. Кроме того, при широком внедрении ИИ неизбежно усиление дифференциации в мощи и функциональных возможностях различных ИИ. Соответственно суверенностью можно сказать, что *более сильные ИИ будут использоваться для атак, в том числе в целях установления программного контроля над более слабыми ИИ.*

Наверняка следует ожидать *атак на ИИ с использованием фишинга.* В обозримой перспективе основным типом ИИ будут системы, постоянно взаимодействующие с человеком: с программистами, аналитиками или администраторами. Поскольку именно *человек является наиболее уязвимым звеном в системах гибридного интеллекта,* есть основания полагать, что злоумышленники будут предпочитать дешевые фишинговые атаки дорогостоящим атакам непосредственно на ИИ.

Самые современные фишинговые атаки предполагают *высокий уровень социального инжиниринга*[10]*и* хорошее знание психологических особенностей того человека, против которого осуществляется атака. В конечном счете, все делается для того, чтобы он кликнул на ссылку, которая позволит злонамеренному софту проникнуть не только в компьютер человека, но и в ИИ, с которым он взаимодействует.

Прогресс в ИИ, несомненно, приведет к увеличению масштабов и эффективности злоумышленников. Этот вывод следует из того простого факта, что *ИИ способствует повышению анонимности взаимодействующих с ним акторов.* Если актор знает, что его атака не может быть отслежена, или даже в худшем случае обнаружение не может надежно идентифицировать его с атакой, ему гораздо легче решиться на нападение, чем хакерам в прошлом. По мнению многих практиков в сфере кибербезопасности, распространение ИИ приведет к значительным изменениям в психологии потенциальных злоумышленников и сформирует у них *установку на безнаказанность.*

Большую угрозу создает тот факт, что *особенностью настоящего времени является создание комбинированных систем, включающих ИИ и роботизированные устройства.* Известно, что ряд стран начал эксперименты по созданию обучающихся роев боевых дронов, связанных с ИИ, выступающим как их центр управления. Причем такие работы ведутся как государственными, так и негосударственными (террористическими и иными деструктивными) структурами.

Преступный мир использует дроны уже довольно активно. В основном цель использования – это *контрабанда* или *доставка тех или иных предметов в тюрьмы.* С помощью дронов заключенным доставляют наркотики, сигареты, мобильные телефоны, бритвенные лезвия и другую контрабанду.

Иногда пытаются доставить таким способом и оружие. Предполагается, что в обозримом будущем все тюрьмы США будут оборудованы системами обнаружения приближающихся дронов. Наркокартели из Мексики отправляют свой товар в США с помощью беспилотников с заранее введенными данными GPS, так что отпадает необходимость наличия оператора.

Но в последнее время в прессу все чаще попадают истории, когда *дроны применяются для сбора информации или иных противоправных действий.*

---

[10] *Социальный инжиниринг* – система управления поведением человека с использованием методов социологии и психологии.

По данным ФБР, *слежка за агентами и сотрудниками правоохранительных органов* с помощью дронов – это быстро растущий тренд в преступном мире.

Дроны применяются и *для оказания давления на свидетелей или лиц, кто может дать показания.* Преступники ставят под наблюдение с воздуха полицейские участки и другие объекты, чтобы фиксировать, кто входит и выходит из здания, таким образом определяя, кто сотрудничает с органами или был вызван для дачи показаний. После чего на этих людей можно начинать воздействовать.

Слежка за членами конкурирующей группировки и ее лидерами – еще одно направление.

Еще один популярный способ применения дронов – это *поиск объектов для совершения краж или ограблений.* Маневренность дронов позволяет им изучать план собственности, где установлена какая система безопасности, где находятся хлипкие двери и окна, где расположены камеры. Все это позволяет преступникам выбрать дом или иной объект, составить детальный план и выбрать маршрут для взлома. С помощью дронов можно также установить график пребывания хозяев или охраны на объекте, привычки владельцев, количество людей и так далее.

В Австралии с *помощью дронов преступники следят в портах за контейнерами с контрабандным товаром.* Если сотрудники портовых служб приближаются к контейнеру, то преступники идут на различные ухищрения, чтобы отвлечь их: сообщают о пожаре в другом конце порта, краже или вызывают ложную тревогу любым другим способом.

В Ирландии зафиксированы случаи, когда дроны вели *воздушную съемку банкоматов,* так как операторы, видимо, надеялись, что технологии позволят на большом расстоянии рассмотреть, какие пин-коды вводят люди.

Еще одно направление – это *использование дронов для несанкционированной съемки пикантных сцен* сексуального плана через окна или с воздуха над частной собственностью в целях дальнейшего использования полученных кадров для шантажа попавших под камеру людей.

Наибольшую опасность представляют *дроны со взрывчаткой.* Беспилотники такого рода находят с 2002 года. Использование в массовом порядке таких беспилотников Исламским государством (организация, запрещённая в РФ) на Ближнем Востоке не осталось незамеченным преступниками в других частях света. Например, в Мексике подобные летающие машинки смерти уже находят у картелей. Дроны со взрывчаткой прозвали «*бомбы-картошки*»[11].

Прогресс ИИ породит *новые угрозы.* Уже есть признаки того, что *начали совершаться киберпреступления и проведены хакерские атаки, управлять которыми будет не человек, а ИИ.* При этом даже в случае обнаружения и отражения атаки крайне сложно найти лиц, инспирировавших эти атаки.

Кроме того, системы ИИ активно используются *для распространения дезинформации и фейков.* В настоящее время дезинформация, как правило, разоблачается на основе анализа фотографического материала. Поскольку ИИ позволяет не только синтезировать любое фотоизображение, но и создать практически не отличимую от реальности фальсифицированную аудио- и видеозаписи, можно ожидать, что в самое ближайшее время появятся технически сложные фейки, подкрепленные синтетическими фотографиями, аудио- и видеозаписями. Для того, чтобы доказать их поддельность, потребуются огромные финансовые средства, мощные технические возможности и усилия высококвалифицированного персонала.

В ближайшем будущем существующие угрозы будут дополнены новыми, связанными с развитием ИИ. Типовые угрозы станут более технически сложными и изощренными. Это проявится по нескольким направлениям.

Во-первых, в ближайшие годы скачкообразно увеличится сложность кибератак и кибертерроризма. Типичными станут не привычные атаки, связанные с фишингом, заражением

[11] См.: *Плеханов И.* Разведывательные дроны преступников. Альманах «Искусство востока», 17.05.2018.

компьютеров и т. п., а гораздо более *высокотехнологичные атаки, нацеленные на овладение информационными массивами атакуемых компьютеров и перехват управления ими.* С другой стороны, более широкое распространение получат *целевые атаки.* В настоящее время типичная кибератака со стороны высокотехнологичных преступников ориентирована на компьютеры, обслуживаемые тем или иным провайдером, расположенные в той или иной местности и т. п. При подключении к киберпреступным атакам ИИ можно будет проводить предварительную селекцию не самих технических средств, а их обладателей по полу, возрасту, профессиональным занятиям и т. п. Соответственно, в этом случае атаки будут ориентированы не на регионы или провайдеров, а на те или иные группы населения, либо компании, обладающие определенными характеристиками. Это будет *киберпреступностью принципиально нового типа.*

В 2017 г. антитеррористические подразделения Израиля успешно провели испытания дрона, который атаковал в многотысячной толпе строго определенных лиц. В качестве эксперимента одежда этих лиц была обрызгана определенной краской. Однако никто не мешает вместо краски использовать отравляющее вещество, либо просто пулю. Мы имеем дело с объединением дрона с ИИ, способным в потоковом видео опознать лицо среди тысяч субъектов.

В подавляющем большинстве докладов по теме ИИ львиная доля внимания, связанного с угрозами, приходится на риски попадания ИИ в детские руки, либо в руки террористов и т. п. Что касается детей и подростков, то разрушительный эффект их деятельности подчас оказывается сопоставимым с ударами со стороны экстремистских радикалов, вооруженных гаджетами. В 2017 году в Польше на протяжении двух дней было парализовано все городское движение просто потому, что одному 13-летнему «таланту» захотелось проверить свои расчеты относительно того, можно или нет проложить трамвайную линию не внутри города, а между городами. Эта шалость обошлась Польше почти в 30 млн. злотых и почти 10 человек, пострадавших в авариях, были доставлены в больницы.

На сегодняшний день накоплено достаточно материала, чтобы изложить классификацию вредоносного использования элементов ИИ по недосмотру, ошибке и т. п., приводящих к негативным последствиям.

В максимально грубом приближении можно выделить *три типа угроз, связанных с использованием ИИ добропорядочными акторами.*

Первая группа объединяет ИИ с подавляющим большинством других сложных машин, созданных человеком. Речь идет о банальных отказах. К сожалению, совершенно не осознанным остался тот факт, что интернет всего, по сути, означает ИИ всего. Любые компании – производители продукции, в которую встроены миникомпьютерные элементы – от чайника до кроссовок – наиболее эффективно выполняют свои обязанности. Если могут участвовать в коллективном машинном обучении. Однако это возможно лишь в том случае, если все эти устройства задействованы на центральный процессор, который анализирует недостатки, конфликты, инциденты, делает из этого выводы, и вносит изменения в программы вещей, связанных с интернетом. Теперь предположим, что в силу программного сбоя обучение произошло неправильно, и вместо того, чтобы снизить вероятность неблагоприятных последствий, все устройства сети научились, как попадать в ситуацию, в которую попал виновник происшествия. Такие случаи в реальности уже случались. Компьютер не обладает самосознанием и поэтому он обучает всех тому, что предусмотрено в его программе.

Вторая группа угроз сопряжена с особенностями программного обеспечения ИИ. На сегодняшний день и, видимо, в период ближайших пяти лет, алгоритмическим ядром ИИ будут выступать нейронные сети вкупе с машинным обучением. Как уже отмечалось, по сути, нейронные сети – это программная поисковая среда, которая постоянно меняется за счет перенормирования удельных весов, определенных программой, в зависимости от фактически полученных результатов.

Если в 2015–2017 гг. ИИ использовал простые нейронные сети, соответственно, и разработчики и аналитики хорошо понимали значение перенормировок на каждой итерации расчетов, то *нынешние глубокие сети оказываются для человека черным ящиком.* Фактически *возникает ситуация, когда машины делают выводы, которые в подавляющем большинстве являются точными, но как и почему они делаются, люди не понимают. Фактически ИИ превращается в черный ящик, относительно которого известны только вход и выход.*

В научных и политических дискуссиях, которые ведутся вокруг модели «ИИ как черный ящик», прежде всего, в США, а также Великобритании и Израиле, на первый план выступает стремление сделать этот черный ящик прозрачным и понятным для аналитиков. Однако если посмотреть статистику фактических инцидентов с ИИ, то заботиться надо не о вскрытии черного ящика, а о явном задании времени оптимизации.

Многие исследователи опасаются, что компьютер при решении той или иной задачи построит программу, в которой оптимизироваться должно то, что оптимизируемым с точки зрения человеческого общества ни в коем случае быть не может. Грубо говоря, существует перезагруженный авиационный маршрут. Число желающих осуществить перелет намного превышает возможности авиакомпании. Компьютер, рассмотрев различные способы решения этой проблемы, пришел к выводу, что лучшим вариантом будет серьезная авария без смертельных случаев, но с большим числом раненых самолета данной авиакомпании на данном маршруте. Модель показала падение числа желающих до нормативного уровня. У математиков эта ситуация известна как *отсутствие запрета на скрытую оптимизацию.*

Данный пример показывает не только появление принципиально новых угроз, но и принципиальное различие в подготовке, анализе и принятии решения у человека и компьютера. Человек отказался бы от подобной оптимизации на самой ранней стадии разработки темы. А компьютер выбрал ее как основную.

Еще одна группа угроз связана, как это ни парадоксально, с *притуплением внимания и снижением ответственности лиц, принимающих решения, чьим советником является ИИ.* В отличие от триллеров и фантастических блокбастеров, лица, принимающие решения, это, в подавляющем большинстве, обычные по интеллектуальным способностям средние люди. Они находятся под прессингом воздействия социальных СМИ, телевидения, интернета, которые изо дня в день вот уже на протяжении двух-трех лет рассказывают о всемогуществе ИИ. Соответственно, даже в тех случаях, когда окончательные решения остаются за человеком, а ИИ дают лишь рекомендации, то, как показали эксперименты в университетах Йокогамы (Япония) и Ванкувера (Канада), лица, принимающие решения на уровне полицейских управлений городов, более чем в 98 % случаев солидаризировались с рекомендациями ИИ и принимали те решения, которые де-факто выработал ИИ.

В одном случае опыт проводился для 70 ситуаций, в которых принимали участие три полицейских начальника, а в другом – для 300 ситуаций, где работало пять начальников. Самым удивительным итогом эксперимента стало следующее. ИИ дали неправильные ответы по оценке ситуации для Японии примерно в 20 % случаев, для Канады – в 17 %. Начальники же в тех примерно 10 % случаев, где приняли решение вопреки ИИ, правы оказались лишь в Канаде в 5 %, а в Японии – ни в одном. Данные выкладки показывают, что тема гибридного или человеко-машинного интеллекта чрезвычайно сложна. В конечном счете, мы пытаемся соединить то, в чем мы вообще ничего не понимаем, – человеческое сознание, с тем, что является техникой в первом поколении ИИ, и надеемся на базе этого соединения успешно решать все проблемы.

Рассмотрим основные **сценарии злонамеренного использования ИИ**

Ключевой угрозой является *автоматизация социальной инженерии.* С помощью ИИ на человека, являющегося целью социальных инженеров, собирается досье. При этом особое вни-

мание обращается на его непроизвольные автоматические реакции, которые и будут использоваться социальными инженерами при фишинговых атаках, использовании телефонии и т. п. По мере развития ИИ в целях обеспечения анонимности возможно использование социальными инженерами чат-ботов, которые будут вести разговоры с жертвами. Наряду с автоматизацией социальной инженерии следует ожидать использования ИИ для улучшения выбора целей и определения приоритетов в злонамеренных атаках. Автономное программное обеспечение, внедренное в атакуемую сеть, будет в течение долгого времени обеспечивать ИИ необходимой информацией.

Конец ознакомительного фрагмента.