ИИ-технологии: пять самых больших страхов

Искусственный интеллект (ИИ) перестал быть новомодным термином, которым щеголяют СМИ и "продвинутые" бизнесмены. ИИ-технологии уже ощутимо влияют на целый ряд отраслей, особенно те, что в значительной мере зависят от ручного труда.

Однако когда речь заходит об искусственном интеллекте, в памяти многих невольно всплывают сцены из голливудских блокбастеров, типа "Терминатора" и "Бегущего по лезвию", где роботы-убийцы или вышедшие из под контроля компьютерные системы стремятся уничтожить человечество.

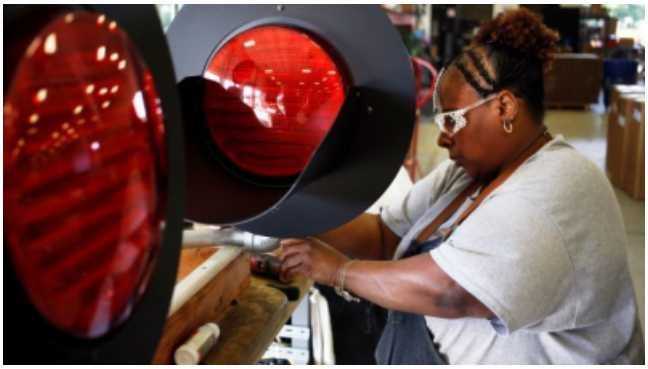
Между тем, аналитики Gartner **прогнозируют**, что выгода для бизнеса от ИИ-технологий в результате формирования потребительского опыта, увеличения продаж продуктов и услуг, появления новых источников прибыли и сокращения затрат, в 2018 году вырастет на 70 процентов и составит 1,2 триллиона долларов, а в 2022-м приблизится к четырём триллионам долларов. Сомнения относительно последствий внедрения ИИ могут серьезно сказаться на перспективах рынка.

Обозреватели CNBC поговорили с экспертами и расспросили о пяти самых пугающих сценариях будущего с искусственным интеллектом.

Массовая безработица

Среди аналитиков и просто работающих людей довольно распространен страх, что внедрение искусственного интеллекта может спровоцировать массовую безработицу в мире: растущая автоматизация на рабочих местах приведет к тому, что человеческий труд будет уже не нужен.

Однако сторонники технологии, такие как Алан Банди (Alan Bundy), профессор факультета информатики Эдинбургского университета, убеждены, что искусственный интеллект создаст новые, высококвалифицированные рабочие места. В особенности усилится спрос на инженеров, поскольку для совершенствования новой технологии потребуются соответствующие профессиональные кадры.



В Gartner прогнозируют, что из-за ИИ-технологий к 2020 году будет ликвидировано 1,8 миллиона рабочих мест, но одновременно создано 2,3 миллиона новых. То есть, чистая прибавка составит 500 тысяч рабочих мест. Впрочем, подобный сценарий не исключает возможности массовых сокращений по всему миру.

Одно из самых часто цитируемых исследований, посвящённых тому, как автоматизация и компьютеризация скажутся на рынке труда - **работа** ученых Оксфордского университета "The Future Of Employment: How Susceptible Are Jobs To Computerisation?" В нем специалисты пришли к выводу, что в зоне наибольшего риска замены роботами – низкоквалифицированные профессии. В списке – различные клерки, банковские служащие, страховщики, специалисты по заполнению налоговых деклараций, операторы станков, тестировщики, сортировщики, приёмщики заказов, водители такси, курьеры, обслуживающий персонал в кафе, на стойках регистрации в отелях и другие.

Среди мер, с помощью которых предлагается минимизировать негативные последствия от внедрения ИИ для рынка труда, называются повышение квалификации сотрудников, создание новых рабочих мест и введение универсального базового дохода.

Война

Появление так называемых "роботов-убийц" (killer robots) и другие варианты внедрения ИИ-технологий в военной сфере порождают страхи, что искусственный интеллект может спровоцировать военный конфликт.

Широкую огласку в СМИ получили мрачные пророчества американского предпринимателя и главы компании Tesla Илона Маска (Elon Musk), который неоднократно заявлял, что гонка за первенство в сфере ИИ может привести к третьей мировой войне.

Хотя Маск известен своей эксцентричностью и склонностью к преувеличениям, эксперты разделяют его тревогу. Некоторые аналитики и активисты заявляют, что разработка летального автономного оружия и использование искусственного интеллекта для принятия стратегических военных решений создаёт массу этических дилемм и открывает возможности для ведения войны с помощью ИИ-технологий и под их началом.

Уже даже существует целое движение против роботов-убийц под названием The Campaign to Stop Killer Robots. Его участники добиваются государственного запрета разработок беспилотных летательных аппаратов под управлением ИИ и других видов роботизированного оружия.

Стоит вспомнить и о нашумевшем конфликте в Google из-за контракта с Пентагоном. Под нажимом своих сотрудников корпорация была вынуждена отказаться от дальнейшей работы с Министерством обороны США.

Между тем, специалисты американского аналитического центра RAND Corporation в своем недавнем исследовании предупредили, что ИИ-технологии могут инициировать ядерную войну к 2040 году. По мнению учёных, оборонные ИИ-системы способны допустить ошибку при анализе ситуации, что толкнёт людей на катастрофические решения.

Опасения эти не беспочвенны, ведь однажды подобное чуть было не произошло. В 1983 году советский офицер Станислав Петров предотвратил начало ядерной мировой войны. Находясь на дежурстве, Петров получил от космической системы раннего ракетного предупреждения "Око" информацию о запуске пяти ядерных ракет с американских баз. Офицер счёл тревогу ложной, поскольку не получил подтверждения от специалистов "по визуальному контакту", отслеживающих обстановку в воздушном и космическом пространстве на экранах. Впоследствии выяснилось, что компьютерная система действительно дала сбой.

Роботы-врачи

Хотя эксперты, в основном, сходятся во мнении, что ИИ-технологии несут с собой множество преимуществ для медицины, включая более раннюю диагностику и повышение оперативности лечения, в научных кругах есть и скептики, считающие, что человечество чересчур стремительно движется в сторону медицины, базирующейся на анализе данных. По их мнению, люди переоценивают возможности ИИ-технологий и ждут от них слишком многого.

В частности, уже упомянутый профессор Эдинбургского университета Алан Банди отмечает, что успехи, которые демонстрирует ИИ в сфере здравоохранения, очень узконаправленны. Например, диагностическое ИИ-приложение, которое отлично справляется с заболеваниями сердца, может ошибочно определить у онкологического больного какую-то редкую кардиопатологию, что теоретически способно привести к фатальным результатам, заявил специалист.

Недавно издание Stat со ссылкой на внутреннюю документацию IBM сообщило, что на ранних этапах тестирования ИИ-платформа Watson for Oncology нередко давала неверные и небезопасные рекомендации для лечения рака. Как следует из статьи, к ошибкам привело то, что вместо данных настоящих пациентов для обучения ИИ-системы было использовано небольшое число гипотетических сценариев онкозаболеваний.

К счастью, никакого реального вреда ИИ-система Watson не причинила. По **данным** Fortune, приведенные изданием выдержки из документов – двухлетней давности, а в самой IBM заявили, что алгоритмы Watson продолжают дорабатываться с учётом отзывов врачей, и что прозвучавшая критика со стороны онкологов позволит сделать систему эффективнее и безопаснее.

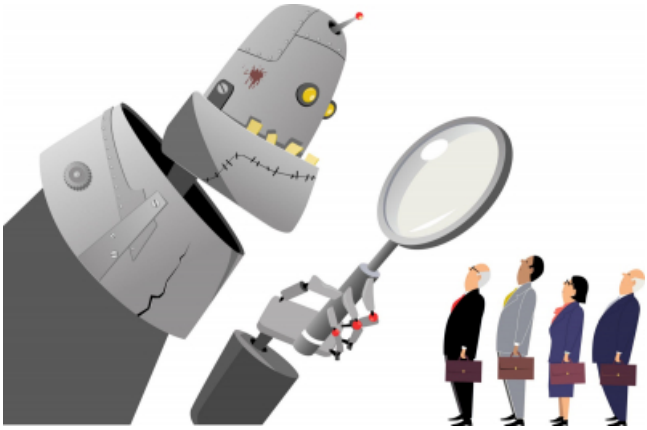
Кроме того, еще есть и этическая проблема. В частности, правозащитники задаются вопросами, насколько надёжно хранится персональная информация и к кому попадают данные из медицинских карт пациентов. Например, в 2017 году в Великобритании разразился скандал вокруг компании DeepMind Health, когда выяснилось, что Национальная служба здравоохранения Британии передала ей данные 1,5 миллиона пациентов без их о том ведома. А недавно группа независимых аналитиков

предупредила о риске "чрезмерной монопольной власти" со стороны DeepMind Health из-за наличия у нее доступа к большим объёмам информации.

Тотальная слежка

Также эксперты опасаются, что искусственный интеллект станет инструментом массовой слежки. В интервью CNBC профессор Тоби Уолш (Toby Walsh) из Университета Нового Южного Уэльса в Австралии (University of New South Wales, UNSW) заявил, что тотальная слежка – на одном из первых мест в его списке самых пугающих последствий внедрения ИИ.

В Китае такой сценарий, напоминающий роман-антиутопию, уже воплощается в жизнь. По данным журнала New York Times, в КНР действует около 200 миллионов камер видеонаблюдения. ИИ-технологии и системы распознавания лиц все шире внедряются в Поднебесной и используются для различных задач - от подтверждения платежей и поимки преступников до регистрации зрителей на крупных развлекательных мероприятиях или посетителей в сетях быстрого питания.



В Пекине и Шанхае камеры с поддержкой ИИ и распознавания лиц помогают регулировать движение транспорта и идентифицируют водителей, нарушающих ПДД. В городе Чжэнчжоу, столице провинции Хайнань, с начала 2018 года транспортную полицию **экипировали** передовыми смарт-очками со встроенными камерами. Такие очки помогают выявлять разыскиваемых полицией граждан в потоке прохожих. В Шэньчжэне ИИ-компания IntelliFusion **внедряет** систему видеонаблюдения с искусственным интеллектом и функцией распознавания лиц, которая регистрирует переход проезжей части не по правилам и фотографирует пешеходов-нарушителей.

Помимо этого, власти отслеживают интернет-активность жителей и их коммуникации, передвижения по стране и выезд за ее пределы. В рамках внедряемой в КНР системы социального рейтинга гражданам начисляются баллы за "хорошее поведение" и вычитаются за правонарушения и другие проступки. Жителям с низким рейтингом предлагается блокировать доступ к различным социальным благам и удобствам: им могут запретить поездки на поездах и самолетах, ограничить в кредитовании, отказать в соцобеспечении или не допустить на работу в госучреждениях.

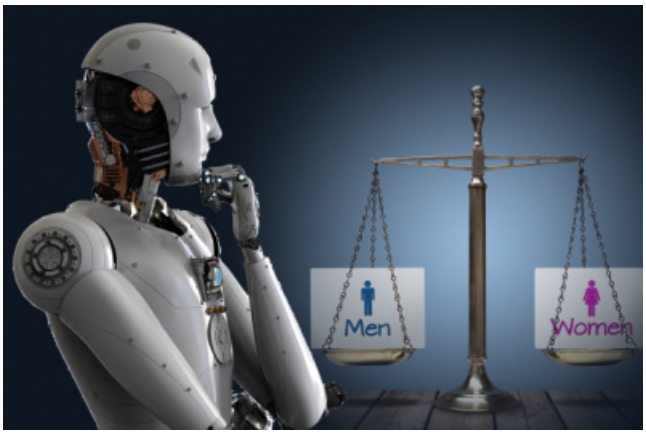
Повсеместное наблюдение за гражданами вызывает беспокойство у правозащитников. Они обвиняют китайские власти в нарушении частной жизни, стремлении создать государство тотальной слежки и контроля.

Дискриминация

Предполагалось, что ИИ-системы будут предельно объективны, рациональны и лишены расовых, гендерных, этнических и других предрассудков. Однако неожиданно для экспертов искусственный интеллект оказался склонен к предвзятости.

> "Мы видим непреднамеренную предвзятость в ИИ-алгоритмах, в особенности в системах машинного обучения, что грозит расовыми, сексуальными и другими предрассудками, за искоренение которых мы боремся последние 50 с лишним лет", - отмечает Тоби Уолш.

Как поясняют специалисты, причина предвзятости ИИ в том, что данные, используемые для обучения систем искусственного интеллекта, не беспристрастны и зачастую отражают предубеждения, сознательные или безотчетные, которые свойственны людям, собиравшим эту информацию. Также на данные может влиять исторически сложившаяся дискриминация.



Для иллюстрации того, как искусственный интеллект может приобрести опасные предрассудки, группа ученых из Массачусетского университета недавно создала **нейросеть Норман** со склонностями психопата.

Специалисты научили ИИ-систему распознавать и описывать пятна Роршаха (психодиагностический тест для исследования личности, в котором испытуемому предлагается рассказать, что он видит и чувствует, глядя на абстрактные кляксы), взяв в качестве обучающей выборки подписи к фотографиям убийств, найденные в тематической группе портала Reddit. В результате в каждом изображении нейросеть стала видеть человеческую смерть.

Исследователи и разработчики ищут способы, как избавить ИИ-алгоритмы от предвзятости. В частности, корпорация IBM **пообещала** осенью 2018 года открыть доступ к двум новым, более обширным наборам дата-сетов для обучения ИИ-систем, которые позволят повысить точность распознавания людей разного пола и с разным цветом кожи.