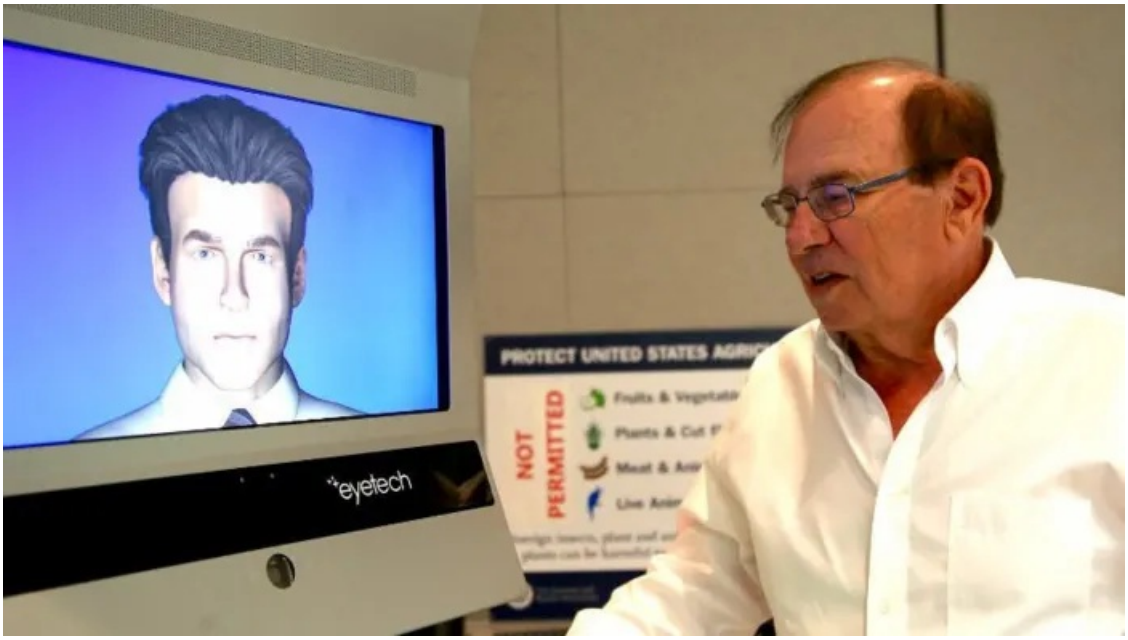
AVATAR выведет на чистую воду

Группа американских исследователей готовит к выходу на рынок детектор лжи на базе искусственного интеллекта, который планируется использовать в аэропортах и на пограничных пунктах пропуска.

ИИ-систему для определения обмана под названием AVATAR, разработал стартап Discern Science International (DSI) при Аризонском университете. Предполагается, что новинка, которая уже прошла несколько апробаций, сделает процедуру пограничного контроля более эффективной и повысит точность выявления людей с общественно опасными и противоправными намерениями. Вместе с тем, у специалистов все еще остаются вопросы по поводу того, насколько безошибочно ИИ-алгоритм сможет оценивать присущее всем людям свойство лгать.



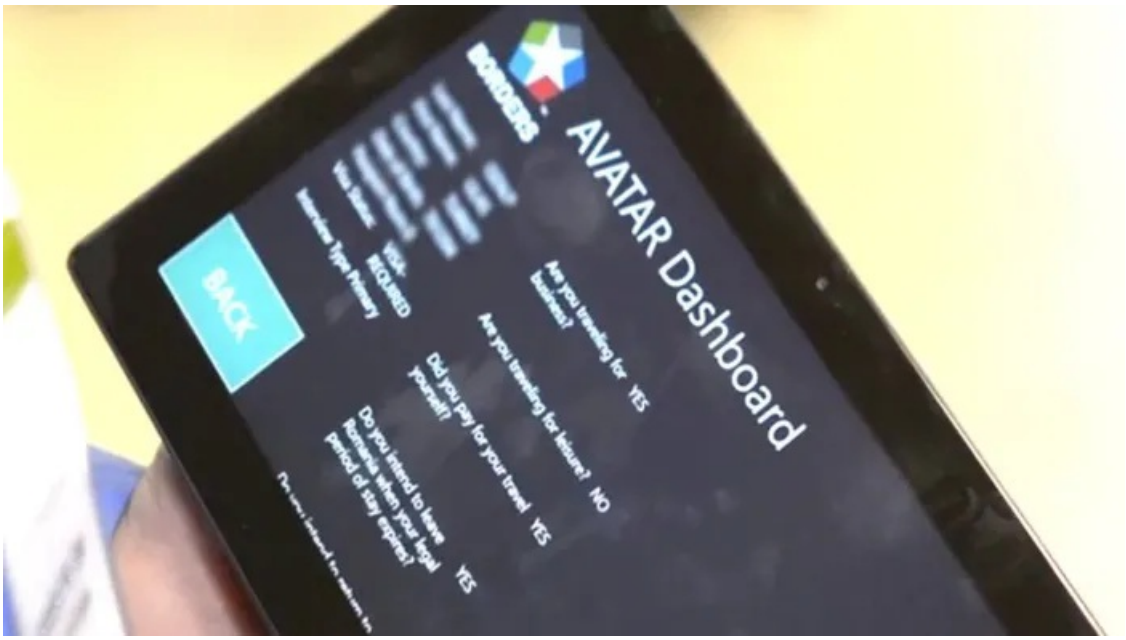
Работает AVATAR следующим образом: виртуальный "пограничник" на экране задает путешественникам серию вопросов, ответы на которые снимаются на камеру. Система анализирует всю поступающую информацию, включая выражение лиц, интонацию и вербальные ответы, выявляя признаки лжи по непроизвольным микровыражениям, которые возникают по причине когнитивного стресса при попытке обмануть.

Несмотря на сложность аналитического алгоритма, система практически мгновенно выдает результат в виде трех цветовых индикаторов. Зеленый означает, что человек говорит правду и может пройти дальше без каких-либо дополнительных проверок. Если же загорается желтый или красный, то въезжающему в страну предстоит беседа с сотрудниками пограничной службы и более тщательная проверка.

AVATAR - аббревиатура полного названия Automated Virtual Agent for Truth Assessments in Real-time, что в переводе на русский означает "автоматизированный виртуальный агент по оценке правдивости в режиме реального времени". Система разработана по результатам научных исследований, финансируемых из федерального бюджета США, и может появиться на рынке через шесть месяцев, сообщила компания изданию The Financial Times.

Discern Science International, основанная в 2018 году учеными Аризонского университета специально для монетизации технологии, в прошлом октябре заключила соглашение о партнерстве с одним из представителей авиационной отрасли. Совместно с ним разработчик планирует продавать ИИ-систему оценки правдивости аэропортам. В DSI говорят, что пока не могут раскрыть имени компаньона, но утверждают, что это хорошо известная компания.

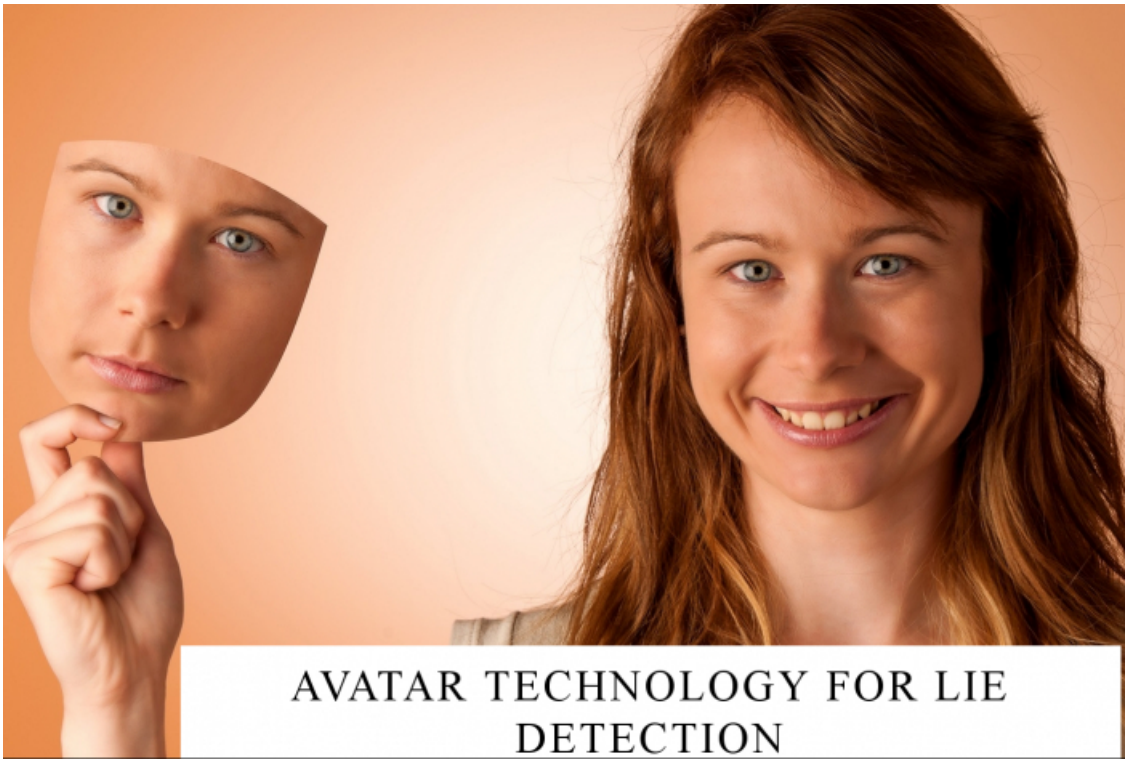
В DSI сообщили, что в настоящее время партнер компании тестирует пилотные модели системы. Также вносятся финальные изменения в формулировки вопросов, поскольку некоторые из них вызвали неожиданные ответы.



В авиационном секторе и сфере обеспечения безопасности растет интерес к поведенческому анализу. В мае 2019 года представители Управления транспортной безопасности США (Transportation Security Administration, TSA), лондонского аэропорта Гатвик (Gatwick) и Управления аэропортов Израиля (Israel Airports Authority) обсудили внедрение поведенческих технологий для повышения безопасности пассажиров и снижения рисков.

Однако остаются сомнения относительно достоверности результатов ИИ-детекторов лжи - действительно ли подобные системы выявляют обман, то есть намеренную попытку скрыть информацию, или это нечто другое? В июле аналогичный ИИ-детектор лжи под названием "Silent Talker", проходящий апробацию в ЕС, провалил тест, устроенный журналистами издания The Intercept. Система посчитала ложью четверть абсолютно правдивых ответов репортера.

Одна из проблем ИИ-полиграфов - ложноположительное срабатывание. Система может зарегистрировать подозрительное микровыражение, если человеку больно или он находится в замешательстве. Аналогичным образом, традиционный тест на детекторе лжи может выдать ложноположительный результат, если исследуемый испытывает стресс или силится что-то вспомнить. По этой причине использование результатов исследования на полиграфе не получило однозначного разрешения в практике и теории уголовного судопроизводства США.



"Люди постоянно врут. Но их не выдают носы, как это случалось с Пиноккио. Ни один из известных нам методов не позволяет гарантированно определять ложь во всех случаях", -

> заявил изданию Мирча Злотяну (Mircea Zloteanu), преподаватель психологии Тиссайдского университета (Великобритания). Он добавил, что алгоритм не может оценить и понять умысел или мотив чего-то поведения.

Однако директор по стратегическим вопросам DSI Дэвид Максталлер (David Mackstaller) уверен в ИИ-системе AVATAR. По его словам, она учитывает настолько обширный комплекс разнообразных показателей, характерных для намеренного и неосознаваемого поведения, что эффективно справляется со своей задачей.

> "Человек может управлять некоторыми своими импульсами, но не всеми одновременно, вот почему наш ИИ-алгоритм работает", - пояснил Максталлер. Он добавил, что коэффициент точности AVATAR составляет 80-85 процентов. Это существенно выше среднего показателя точности выявления обмана у людей, равного 54 процентам.

Процедуры контроля в аэропортах должны соответствовать международным стандартам, особенности которых не разглашаются. В DSI говорят, что AVATAR сможет соответствовать установленным требованиям, но отмечают, что ответственность за внедрение технологии ложится на клиентов компании.

За последние несколько лет ИИ-систему протестировали Управление пограничной службы Канады (Canada Border Services Agency, CBSA), аэропорт Бухареста (Румыния) и пограничный пункт в города Ногалес (штат Аризона, США).



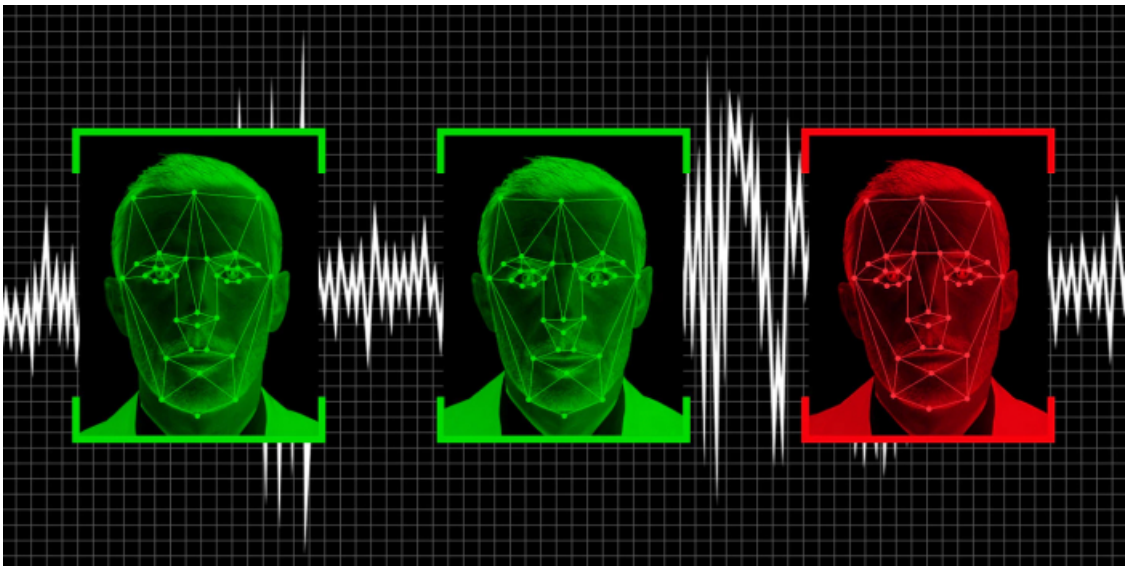
Однако в управлении канадской погранслужбы сообщили, что система прошла испытания лишь в лабораторных условиях, и что выходить за эти рамки не планируется.

> "Ряд существенных ограничений, с которыми нам пришлось столкнуться в ходе эксперимента и которые свойственны данной технологии в целом, привели нас к выводу, что испытание системы в условиях эксплуатации на данный момент не является приоритетной задачей", - заявили в CBSA.

Миграционные органы в Ногалесе сообщили, что не планируют пилотный запуск системы.

Впрочем, некоторые отрасли уже применяют технологии распознавания эмоций и поведения. Например, компания Neurodata Lab, имеющая представительства в России, Италии, Швейцарии и США, заявляет, что ее сервисы, основанные на поведенческом анализе, используются для выявления обмана в колл-центрах и банках.

Интерес к такого рода разработкам проявляют и американские госведомства. Так, Управление транспортной безопасности США профинансировало исследование на предмет того, служат ли микровыражения индикатором обмана. Кроме того, среди спонсоров научной работы, приведшей к созданию AVATAR, были Министерство обороны и Агентство Национальной безопасности США.



Хотя большинство специалистов считают перспективными подобные проекты, они советуют с осторожностью внедрять технологию, которая еще находится в зачаточном состоянии.

> "Я бы очень обеспокоился, если бы такие штуки стали использовать для принятия решений", - сказал Брайан Вуд (Brian Wood), руководитель проекта компании The SilverLogic, работающей над технологией распознавания эмоций для цифровой маркетинговой платформы.

По мере того, как технологии распознавания эмоций и поведенческого анализа совершенствуются, они могут стать более надежными в плане выявления лжи. Однако спектр применения таких технологий должен быть очень специфическим и ограниченным, убеждены в Neurodata Lab.

> "Нельзя вести речь о какой-то универсальной технологии, которая, подобно волшебной палочке, смогла бы распознать ложь во всех ее проявлениях", - подчеркнули в компании.