

15 октября 2023

# ИИ в законе: что такое этика искусственного интеллекта

Искусственный интеллект (ИИ) стал неотъемлемой частью продуктов и услуг, вследствие чего возникла потребность в разработке принципов этики. Этика искусственного интеллекта — это единая система рекомендательных принципов и правил, предназначенных для создания среды доверенного развития технологий ИИ. Как регулировать использование нейросетей и какие проблемы это может повлечь за собой — разбирались «Известия».

## ЭТИКА МАШИН

Регламентирование использования искусственного интеллекта с законодательной и этической стороны постоянно обсуждается как в экспертном сообществе, так и в массовой культуре. И чем активнее развивается искусственный интеллект, тем этот вопрос становится острее. В первую очередь необходимо понимать, что ответственность за любые действия нейросети всегда должна лежать в плоскости человеческих отношений, у машинного алгоритма нет фактора ответственности. Но при этом в алгоритме можно допустить ошибку. Об этом изданию рассказал основатель и CEO Mirey Robotics Андрей Наташкин.



— Если мы рассматриваем аргументы экспертов, критикующих бурное развитие технологий искусственного интеллекта, то лидирующим аспектом в этой полемике почти всегда будет этический вопрос. Эта тема настолько объемна и сложна, что некоторые ученые предлагают изучать этику искусственного интеллекта в разрезе современных цивилизационных процессов, — отметил эксперт.

В сентябре 2023 года журнал Time опубликовал список из 100 самых влиятельных людей в сфере искусственного интеллекта. Интересным фактом является то, что в этом списке можно найти достаточное количество имен, которые резко критикуют искусственный интеллект с точки зрения этики. По словам Наташкина, американский писатель Тед Чан считает, что нейросети провоцируют пользователей на то, что они не могут отличить правду от лжи. Потому что нейросети убивают в нас критическое мышление.

Еще один представитель из топ-100 — это британский ученый Джеффри Хинтон. Фактически Хинтон разработал всю систему искусственного интеллекта для компании Google. Но в 2023 году он покинул свое кресло по причине того, что хотел открыто высказывать свою точку зрения о развитии технологий. Хинтон уверен, что искусственный интеллект во много раз превосходит мыслительный процесс человека и это несет за собой значительные риски для всего человечества в целом.

— Мы можем отпустить ситуацию, не мешать вообще развитию технологий. Есть опыт Израиля, когда мы общались с израильскими партнерами, они говорили о том, что у них так мало законодательной базы по той причине, что

они не хотят лишний раз вмешиваться в развитие технологий.

Соответственно, это не наш путь, потому что у нас большая страна, и мы выбираем сбалансированный подход между правовым регулированием и этическим, — рассказала менеджер Центра регулирования AI «Сбера» Елена Мимоглядова в рамках форума «Лидеры цифрового развития», организованного компанией «АктивБизнесКонсалт».

По словам эксперта, этическое регулирование заключается в том, что это саморегулирование, а не закон, это рекомендательные нормы, которые принимает на себя бизнес, государственные и некоммерческие организации. Первые шаги к правилам регулирования и правилам взаимодействия с этими технологиями были предприняты крупными корпорациями. Например, «Десять законов для искусственного интеллекта» от Microsoft, «Семь принципов ИИ» от Google.

— Этика ИИ — это не про этику машины, компьютера, на котором он [ИИ] крутится, а именно этика отношения разработчиков и разработки этих технологий. Допустим, при обучении мы проверяем и валидируем data-сети и несем за это ответственность, гарантируем, что наши технологии не навредят человеку, — рассказала Мимоглядова.

При помощи технологий ИИ создаются различные инструменты, эксплуатация которых может быть связана с теми или иными рисками. Сложности могут быть связаны как со злонамеренным использованием знаний и технологий, так и с ошибками в их применении.

— Можно ли сказать, что технология при этом вышла из-под контроля? В какой-то степени да. Именно поэтому в тех областях, где применение ИИ может быть связано с теми или иными рисками, разрабатываются специальные меры по их снижению и обеспечению безопасности. При этом важно понимать, что системы ИИ, создаваемые нами, не обладают ни собственным сознанием, ни собственными целями, ни собственной волей — всё это инструменты в руках людей. И то, принесут они пользу или вред, зависит в первую очередь от тех, кто эти инструменты использует, — рассказал в беседе с «Известиями» руководитель управления экспериментальных систем машинного обучения SberDevices Сергей Марков в рамках форума «Лидеры цифрового развития».

## РЕГУЛЯЦИЯ

Первой попыткой международного урегулирования деятельности ИИ стала конференция в Азиломаре (США), в ходе которой были приняты основные положения, касающиеся искусственного интеллекта, которые получили название «Азиломарские принципы». Из них следует, что команды, разрабатывающие системы ИИ, должны активно сотрудничать между собой и не пытаться победить за счет игнорирования стандартов безопасности. Системы ИИ должны разрабатываться и работать таким образом, чтобы быть

совместимыми с идеалами человеческого достоинства, его прав и свобод, многообразия культур. Кроме того, экономическое процветание, достигнутое благодаря ИИ, должно широко использоваться в интересах всего человечества.

В России 26 октября 2021 года на площадке международного форума «Этика искусственного интеллекта: начало доверия» был принят российский Кодекс этики в сфере ИИ. Документ был подписан крупнейшими российскими компаниями. Андрей Наташкин отмечает, что это соглашение носит рекомендательный характер и распространяется только на гражданские разработки. В 2023 году к нему присоединились 10 стран Африканского континента.

— Если говорить о мировой практике, то в одно время с Россией в Китае приняли Кодекс этических принципов для искусственного интеллекта, в котором особое внимание уделено контролю над ним. В США всегда акцент был смещен на использование искусственного интеллекта в военных целях, поэтому в 2020 году они инициировали создание этических принципов использования искусственного интеллекта в вооружениях, — рассказал основатель Mirey Robotics.

## **ВОЗНИКШИЕ ПРОБЛЕМЫ**

При отсутствии регулирования ИИ на уровне этических принципов можно столкнуться с «предвзятостью» информационных систем. Например, разработчик создает рекомендательную систему на базе ИИ и может настроить ее так, чтобы в первую очередь предлагать товары определенного производителя. Такой подход не дает равных прав для всех игроков рынка и может привести к нездоровой конкуренции и рискам для бизнеса. Другой пример — «предвзятая» система распознавания лиц, которая изначально может идентифицировать людей определенных национальностей как потенциальных нарушителей закона.

Светлана Вронская,  
эксперт департамента аналитических решений ГК «КОРУС  
Консалтинг»

Этические проблемы использования ИИ уже возникают. Так, в 2018 году инструмент рекрутинга на основе ИИ дискриминировал женщин. Около 60% кандидатов, выбранных ИИ, были мужчинами, что было связано с закономерностями в данных об историческом наборе персонала конкретного разработчика. В этом году в США искусственный интеллект полгода преподавал в университете, но никто из учеников не понял, что взаимодействует не с человеком, а с системой ИИ. В 2019 году искусственный интеллект, определяющий, кому из пациентов необходима дополнительная медицинская помощь, стал принимать решения, в первую очередь

основываясь на уровне материального положения пациента.

