«Мы только вступаем в эпоху ИИ»

17.06.2026 ИНТЕРВЬЮ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Александр Крайнов, директор по развитию технологий искусственного интеллекта «Яндекса» объясняет, что модели «галлюцинируют» только при неправильном их применении, а предвзятости у ИИ не больше, чем у человека, а также, почему самая большая угроза ИИ — это кибербезопасность.

— Для решения тех или иных задач пользователи применяют разные модели с ИИ. При этом разница между ChatGPT, «Алисой AI» и DeepSeek простому пользователю неочевидна. Есть ли она на самом деле?

—Разница, конечно, есть. Несмотря на то, что на простые, распространенные вопросы модели отвечают примерно с одинаковым качеством, по узким темам результаты ответов могут сильно отличаться. Люди часто спорят по поводу того, какая модель лучше, при этом каждый отстаивает своего фаворита. А объяснение простое: у всех разные интересы, и для одного оказывается лучше одна модель, для другого — другая. Кроме того, под одним брендом обычно скрывается целое семейство моделей. У «Яндекса», например, есть очень быстрая модель, которая отвечает в поиске, суппаризируя найденные результаты. Там же есть модель для глубокого анализа и умных ответов. И обе эти модели ориентированы на задачи, с которыми обращаются обычные пользователи. А, например, модель, которая используется в «Нейроюристе», ориентирована на решение профессиональных юридических задач.

— Одним из рисков использования ИИ называют предвзятость моделей из-за обучающей их информации или настройки программ, которые применяют эти модели. Так, в программы для подбора кадров могут закладываться возрастное ограничение или иные предпочтения в отношении кандидатов. Какие существуют

методы обучения ИИ, позволяющие выявить незаконные или неэтичные настройки?

—У моделей ИИ предвзятость совершенно точно гораздо меньше, чем у людей. Нам крайне тяжело не опираться на собственный опыт. А сделать так, чтобы, например, модель не учитывала возраст при принятии решений, легко — достаточно убрать сведения про возраст из данных для обучения. В целом модель наследует те этические принципы, которые были в обучающей выборке. Поскольку модель в любом случае учится на примерах, которые создали или создают люди, ей свойственно наследовать этику людей.

— В последнее время в мировом юридическом сообществе активно обсуждается возможность замены искусственным интеллектом мировых судей и независимых директоров. Причем основным преимуществом такой замены называют возможность достижения независимости и объективности. Насколько реально заменить человека там, где нужно выбрать, например, наиболее компромиссное решение в условиях конфликта интересов?

—Полностью исключать человека нельзя и не нужно. При этом ИИ вполне может увеличить количество задач, решаемых одним человеком, и улучшить качество решений, а также послужить источником еще одного независимого мнения.

— Каким вам представляется дальнейшее направление развития ИИ? Какие основные угрозы вы видите в развитии ИИ?

—Думаю, мы только вступаем в эпоху ИИ. Все еще впереди. И это касается не только развития самой технологии, но и возможностей ее применения. А самая большая опасность, на мой взгляд, заключается в том, что технологии ИИ доступны и злоумышленникам. Соответственно, существенно возрастет нагрузка службы кибербезопасности. И, увы, никаким регулированием такая угроза не снимается. Единственный возможный путь минимизации рисков — усиление кибербезопасности, в частности технологиями ИИ.

— Согласно отчету Goldman Sachs, генеративный ИИ способен выполнять до 44% задач юристов. Именно рутинные задачи — анализ документов, due diligence, проверка договоров — входят в топ-5 направлений применения ИИ юристами. Но ведь именно

на таких задачах учатся младшие юристы. Выходит, мы автоматизируем то, что служило тренировочной базой. Какую замену этой базе вы можете предложить?

—На основе ИИ можно делать отличные тренажеры. Я считаю, что ИИ как инструмент для подготовки специалистов обладает невероятным потенциалом. Это же не одно и то же: брать тех, кто будет заниматься рутиной, или тех, кто будет решать сложные задачи, потренировавшись на начальном этапе на рутине.

— Большие языковые модели склонны к «галлюцинациям». С технологической точки зрения можете ли вы сегодня создать ИИ, который сигнализировал бы юристу о том, что в этом месте он может ошибаться и человеку нужно перепроверить? Если нет, то не создаете ли вы инструмент, который формирует иллюзию компетентности вместо реальных навыков?

—Уже сегодня это ошибочный штамп — завтра «галлюцинации» моделей почти исчезнут. Модели «галлюцинируют» (выдают ошибочные утверждения и ссылки на несуществующее) только при неправильном их применении. При решении большинства профессиональных задач стоит делать так, чтобы модель отвечала не из внутренней памяти, а строго на основе документов. Тогда и «галлюцинации» будут практически исключены. Или, скажем так, их вероятность будет не выше, чем у людей. Но в ответственные моменты проверка, конечно, не повредит. Впрочем, как и с людьми — тут ничего не меняется.

— Считается, что перекладывание большей части самостоятельной работы студентов на ИИ может навредить развитию у них критического мышления. Можно ли научить студентов-юристов с помощью ИИ думать как юрист? Или они навсегда останутся «операторами ИИ» с юридическим дипломом?

—Конечно, можно! Ведь не лишился же человек мышц в отсутствие потребности в физическом труде. Просто теперь для их поддержания нужно, например, ходить в тренажерный зал. Тоже самое касается и когнитивных навыков человека. ИИ дает возможность как для их развития, так и для снижения — и выбор тут за каждым человеком.

Александр Крайнов директор по развитию технологий искусственного интеллекта «Яндекса»

ИНТЕРВЬЮ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ