

Проблема ответственности в сфере технологий искусственного интеллекта.

Научный руководитель – Перов Вадим Юрьевич

Головков Владислав Владимирович

Студент (бакалавр)

Санкт-Петербургский государственный университет, Институт философии,

Санкт-Петербург, Россия

E-mail: golovkov.spb@mail.ru

Важной проблемой в этике ИИ на данный момент является проблема ответственности. Роботы становятся всё сложнее и автономнее, в связи с чем расширяется поле их деятельности. Чем больше систем ИИ в нашей жизни, тем с большей вероятностью от их работы может пострадать человек. Причём, здесь имеется в виду не только возможный физический вред, например, полученный от удара машины, но и психологический, связанный с нарушением личного пространства, конфиденциальности или с эмоциональной привязанностью человека к машине.

В настоящее время проблема решается некоей общей или распределительной ответственностью между разработчиками роботов, инженерами, программистами, производителями, инвесторами, продавцами и пользователями. Таким образом, ответственность ложится сразу на всех людей причастных к созданию или работе робота. В то же время, как замечается в докладе COMEST, такой подход имеет некую тенденцию к ослаблению самого понятия ответственности, ведь если ответственность распределяется между несколькими субъектами, то полной ответственности за действия умной машины никто не несёт. [3] Эта проблема известна как «проблема многих рук».

Одним из возможных решений является перенос всей ответственности на разработчиков роботов, но и подобный подход имеет свои подводные камни. Кто сказал, что пользователь умной машины не захочет в определённый момент использовать её для совершения преступления? Кто сказал, что абсолютно сторонний человек не сможет внедриться в систему робота и перепрограммировать его для воплощения плохих целей? В этих случаях винить разработчиков в злодеяниях машины, неправильно. Максимум, что ему можно вменить - это недостаточную степень защиты системы. Кроме всего прочего, роботов разрабатывают и создают довольно большие команды, и не всегда отдельные люди могут оценить возможные последствия их вклада в создание робота. Ну и, естественно, отследить, какой конкретно человек допустил, возможно, даже незначительную, ошибку на стадии разработки, которая привела к краху системы или её неправильной работе, довольно затруднительно. Тем не менее, предполагая, что подобная ответственность всё же может иметь место, авторы предлагают два решения для возможности введения такой ответственности. Во-первых, это создание методов максимально возможного прогнозирования влияния той или иной разработки в сфере робототехники. Во-вторых, в связи с имеющейся опасностью, внедрение роботов в жизнь человека должно производиться максимально осторожно и рассматриваться изначально как некий социальный эксперимент, который с лёгкостью и без серьёзных последствий может быть в любой момент свёрнут. [3]

Избавить разработчиков от ответственности за работу роботов и систем ИИ можно путём присвоения ответственности самим машинам. Система будет сама отвечать за свои действия. Подобные законодательные проекты уже находятся в разработке, как в Европе, так и в России. [1] Да, это избавит производителей ИИ от возможных проблем, но добавит

новых вопросов. Самый элементарный вопрос - «Как наказывать?» - уже ставит в тупик. Кроме того, некоторые исследователи, как, например, С. А. Соменков, считают, что подобные инициативы лоббируются заинтересованными в этом разработчиками и их инвесторами, ведь для них будет намного удобнее не нести ответственности за своё детище. [1] Тем не менее, подобный подход стоит продумать как минимум для будущего.

Альтернативный взгляд на проблему предлагают Розенбергер и Вербик. Он предполагает, что технологии не рассматриваются, как нечеловеческие объекты, противоположные человеческим субъектам, но являются «посредниками» между людьми и окружающей их действительностью. Данная концепция утверждает, что все технологии не являются некоторой частью мира, а частью цепочки «человек - мир». [2] Например, бортовой компьютер соединяет человека с космическим кораблём, мобильный телефон - человека с человеком, земляной бур - геолога с земляными породами и так далее. Предполагается, что робот не становится новым моральным субъектом, а лишь новым орудием в руках человека, благодаря которому изменяется возможна сама суть сферы, в которой этот инструмент работает. Медицинские блоки, боевые дроны, роботы-пылесосы и так далее меняют человеческое отношение к той или иной деятельности. И главный вопрос здесь состоит в том, как роботы влияют на изменение человеческих практик. Задача этики в этом подходе видится в анализе качества отношений между человеком и роботом и применение выводов в разработке и использовании умных машин.

Источники и литература

- 1) 1. Соменков Семен Алексеевич. Искусственный интеллект: от объекта к субъекту? // Вестник Университета имени О.Е. Кутафина. 2019. №2 (54) С. 75 - 85
- 2) 2. Rosenberger, R. and Verbeek, P. P. 2015. A Field Guide to Postphenomenology. In: Rosenberger, R., Verbeek, P. P. eds. Postphenomenological Investigations: Essays on Human-Technology Relations. London, Lexington Books, pp. 9-41.
- 3) 3) Report of COMEST on robotics ethics. 2017. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000253952>