

Социально-психологические аспекты этических проблем, связанных с массовым
внедрением систем искусственного интеллекта

Т.А. Нестик

nestik@gmail.com

Институт психологии РАН, МГУ им. М.В. Ломоносова

В докладе рассматриваются социально-психологические механизмы, влияющие на отношение человека к искусственному интеллекту (ИИ) и его использованию в этически спорных ситуациях. Анализируются особенности доверия человека системам ИИ, ожидания и страхи, связанные с внедрением ИИ в повседневную жизнь, а также психологические предпосылки готовности личности к использованию ИИ в различных сферах личной и общественной жизни. На основе эмпирических исследований рассматриваются социально-психологические предикторы представлений российской молодежи об этических принципах, которым должен следовать ИИ, а также допустимой автономности ИИ в этически спорных ситуациях. Делается вывод о ключевой роли межличностного, межгруппового и институционального доверия в формировании доверия к технологиям ИИ. На основании проведенных исследований формулируются социально-психологические условия эффективности этически ориентированного дизайна систем ИИ. Рассматривается роль ИИ как посредника, поддерживающего интерсубъективность и этику диалога.

Намечены перспективные направления психологических исследований, связанных с внедрением ИИ в повседневную жизнь российского общества. Обращается внимание на необходимость исследования последствий использования алгоритмов для когнитивного и эмоционального развития личности; влияние культурных различий на разработку и подходы к использованию ИИ; возможности, открываемые ИИ для увеличения осознанности, конструирования личностью своей идентичности, образа Я и временной перспективы, для целенаправленного управления самопрезентацией; влияние программируемости мира на процессы каузальной атрибуции и доверие к социальным институтам; последствия использования машинного обучения в системах поддержки группового принятия решений; влияние ИИ на правовое сознание и способы мобилизации права.

Обсуждаются причины и последствия технодетерминистского подхода к внедрению технологий ИИ как «костылей» для плохо работающих социальных институтов.

Необходимость обновления принципов этически ориентированного дизайна по мере развития технологий ИИ требует высокого уровня рефлексивности как со стороны разработчиков, так и со стороны пользователей. В докладе обращается внимание на риск снижения рефлексивности общества под влиянием цифровой «архитектуры выбора», а также на увеличение культурного разрыва между теми, кто готов к неопределенности и выбору, а теми, кто стремится избежать необходимости что-либо выбирать, перекладывая ответственность на обезличенные алгоритмы. Делается вывод о необходимости в разработке гуманитарных цифровых технологий, поддерживающих этику диалога и осознанность выбора граждан «умного города» при использовании систем ИИ.