

Леушина Влада Вячеславовна

## **Этика искусственного интеллекта в стандартах и рекомендациях**

НИЦ «Курчатовский институт»  
НИУ "Высшая школа экономики"  
e-mail: wandbpand@gmail.com

### **Аннотация**

Последние годы стали активно обсуждаться вопросы этики ИИ интеллекта. Это связано, с одной стороны, с интенсивным развитием и распространением технологий ИИ, а с другой – с повышенным интересом к этому вопросу со стороны широкой общественности (не считая политической, коммерческой и подобной заинтересованности). Очевидно, что для дальнейшего применения ИИ стали необходимы международные стандарты или рекомендации, определяющие принципы этичного ИИ, поэтому многие организации, занимающиеся вопросами стандартизации, активно принялись за их разработку. В докладе рассматриваются разрабатываемые и опубликованные нормативные документы международных и региональных организаций, в которых приводится схожее абстрактное описание этических принципов, которым должен соответствовать ИИ. Приводится краткое изложение сути стандартов и рекомендаций. Особое внимание уделено стандарту Р7000 от IEEE, который имеет существенное отличие от всех остальных предложений: он не дает строго определения этических систем, а наоборот, перекладывает эту обязанность на сами организации, таким образом учитывая их ценности и идеологию. Более подробно рассматривается и рекомендация ЮНЕСКО, в которой приводится описание универсальной модели этичного ИИ. Акцентируется внимание на присутствии в некоторых пунктах этого документа конъюнктурных и сомнительных пассажей. Также упоминается отечественный кодекс этичного ИИ, в большинстве своем состоящий из принципов похожих или аналогичных пунктам рекомендации ЮНЕСКО. Приводится сравнение предлагаемых подходов стандартизации вопросов этики ИИ и делается вывод о причинах сходства большинства разработанных нормативных документов. В заключении поднимаются вопросы, касающиеся возможности реализации и соблюдения предлагаемых норм.