

**Философия Ника Бострома: экзистенциальные риски развития
искусственного интеллекта.**

Научный руководитель – Кузнецов Антон Викторович

Сергеева Татьяна Александровна

Студент (магистр)

Московский государственный университет имени М.В.Ломоносова, Философский факультет, Кафедра истории зарубежной философии, Москва, Россия

E-mail: tat.ser2011@yandex.ru

На сегодняшний день искусственному интеллекту отдается большая роль в жизнь науки и общества. Машинный разум может служить как панацея от многих экзистенциальных угроз человечеству, но что делать если он сам является одной из этих угроз? Современный шведский философ, Ник Бостром, в своем труде «Искусственный интеллект. Этапы. Угрозы. Стратегии» рассуждает на эту тему.

По мнению Бострома, человечество не осознает опасность таящую в себе стремительное развитие искусственного интеллекта, который человечество не сможет контролировать. Слабое развитие искусственного разума на ранних этапах может оказаться обманчивым. На первый взгляд и большое отставание в развитии от уровня человеческого интеллекта может быть преодолено очень быстро в одном из сценариев стремительного развития искусственного интеллекта. На этапе, когда искусственный разум преодолеет черту уровня развития человеческого интеллекта, мы уже не сможем контролировать его, подобно тому, как животные, находящиеся на более низком уровне умственного развития, не могут контролировать человека. Таким образом, человечеству необходимо заранее продумать сценарии более плавного развития искусственного интеллекта. Прежде чем это произойдет нам также предстоит решить несколько наиболее важных этических проблем. Проблема переноса ценностей, которую нужно решить прежде чем искусственный разум можно будет подпустить к решению важных экономических или политических задач. Также нужно решить каких критериев разум должен придерживаться в решении этих задач.

Не смотря на разумные опасения, Бостром не считает, что все исследования в отношении искусственного интеллекта необходимо заморозить или прекратить вести разработки в этой сфере, но полагает, что требуется провести большую подготовительную работу, прежде чем запускать в работу сверхразумный интеллект.

Источники и литература

- 1) Бостром, Ник Искусственный интеллект. Этапы. Угрозы. Стратегии / Ник Бостром ; пер. с англ. С. Филина. — М. : Манн, Иванов и Фербер, 2016.
- 2) Юдковский Э. — Искусственный интеллект как позитивный и негативный фактор глобального риска // Singularity Institute for Artificial Intelligence Palo Alto, CA Перевод: Алексей Валерьевич Турчин