



**ПЕРСПЕКТИВЫ ЧЕЛОВЕЧЕСТВА.
ФИЛОСОФИЯ ГУМАНИТАРНО-
ТЕХНОЛОГИЧЕСКОГО РАЗВИТИЯ**



**Философия
искусственного интеллекта**



DOI: 10.30727/0235-1188-2021-64-1-7-12

Введение в рубрику

Section introduction

**О философских проблемах искусственного интеллекта
и когнитивных исследований**

В.А. Лекторский

Институт философии РАН, Москва, Россия

Лекторский Владислав Александрович – доктор философских наук, академик РАН, главный научный сотрудник сектора теории познания Института философии РАН.

v.a.lektorski@gmail.com

<https://orcid.org/0000-0001-8496-7814>

Для цитирования: *Лекторский В.А.* О философских проблемах искусственного интеллекта и когнитивных исследований // Философские науки. 2021. Т. 64. № 1. С. 7–12. DOI: 10.30727/0235-1188-2021-64-1-7-12

**On the Philosophical Issues of Artificial Intelligence
and Cognitive Studies**

V.A. Lektorsky

Institute of Philosophy, Russian Academy of Sciences, Moscow, Russia

Vladislav A. Lektorsky – D.Sc. in Philosophy, Full Member of the Russian Academy of Sciences, Chief Research Fellow, Department of the Theory of Knowledge, Institute of Philosophy, Russian Academy of Sciences.

v.a.lektorski@gmail.com

<https://orcid.org/0000-0001-8496-7814>

For citation: Lektorsky V.A. (2021) On the Philosophical Issues of Artificial Intelligence and Cognitive Studies. *Russian Journal of*

В январе 2005 года основан Совет по методологии искусственного интеллекта при Отделении общественных наук Российской академии наук, впоследствии переименованный в Научный совет по методологии искусственного интеллекта и когнитивных исследований при Президиуме РАН.

Вначале я занимал должность заместителя председателя этого Совета, в настоящее время являюсь председателем. В первые годы деятельности нашего Совета многие смотрели на нас как на чудаков, занявшихся экзотической, малоперспективной и далекой от практики проблематикой. Сегодня становится очевидным, что вопросы, которым посвящена работа Совета, – это передовой край развития всех наук, как естественных, так и социально-гуманитарных. Вместе с тем речь идет о принципиально новых, интеллектуальных технологиях, в корне изменяющих экономику, социальную жизнь, человека. Когда с высокой трибуны было сказано о том, что тот, кто оседлает искусственный интеллект, будет определять мировые процессы, повсюду стали говорить о значимости разработки проблематики интеллектуальных систем, когнитивных наук и цифровизации.

За 16 лет интенсивной работы Совета создано несколько десятков отделений в России, десятки регулярно заседающих секций, охватывающих специальную проблематику исследований. С нами взаимодействуют разработчики искусственных интеллектуальных систем, математики, психологи, нейрофизиологи, лингвисты. Важнейшую роль при этом играют философы. В указанный период проведено 100 заседаний семинара Совета, несколько сотен заседаний секций. В процессе и по итогам работы Совета издано несколько книг, опубликовано множество статей.

В контексте изложенного хотелось бы высказать ряд соображений о философско-методологическом смысле проблематики искусственного интеллекта и когнитивных исследований. 70 лет тому назад английский математик А. Тьюринг написал статью «Вычислительные машины и разум», которая была

размещена на страницах ведущего британского философского журнала «Mind». Это не случайно. Текст Тьюринга построен на его математических идеях, однако данная статья имеет прежде всего философский смысл. Речь идет о том, можно ли допустить ситуацию, когда искусственный интеллект достигнет такой степени развития, что с помощью некоторых текстов невозможно будет отличить, с кем мы имеем дело – с живым человеком, т.е. естественным существом, или с искусственным интеллектом. Статья Тьюринга породила массу продолжающихся до сих пор философских дискуссий: можно ли искусственно имитировать все познавательные процессы человека, возможно ли машинное творчество, может ли искусственная интеллектуальная система понимать другого и саму себя и др. Интересно, что Тьюринг сформулировал эти вопросы, когда еще не было компьютеров. Но именно он заложил фундамент компьютерных наук, а также выявил многое из философской проблематики, которая обсуждается ныне в связи с разработками в области искусственного интеллекта.

После публикации Тьюринга появились первые компьютеры, а спустя некоторое время возникла новая дисциплина – когнитивная наука. Это междисциплинарное движение включало в себя разработчиков искусственных интеллектуальных систем, философов, специалистов в области когнитивной психологии, нейронауки, когнитивной лингвистики. Центральная идея когнитивной науки заключалась в том, что наше сознание, мышление работает по типу гигантского компьютера. Согласно мнению основателей когнитивной науки, человек – несовершенная интеллектуальная система, а искусственный интеллект может стать совершенной системой такого рода.

Со временем в системе когнитивных исследований произошли существенные изменения. Сегодня приоритетными в этих разработках стали т.н. когнитивные нейронауки – попытки понять работу нашего мозга с помощью вычислительных моделей. В итоге получены огромные результаты, сделаны интересные открытия (например, зеркальных нейронов), предложены различные модели работы сознания. Большинство ис-

следователей в этой области, если и не разрабатывают модели искусственного интеллекта, то пытаются их применять.

В качестве наиболее важного аспекта исследований в области искусственного интеллекта выступает не просто понимание сознания, мышления, человека, а мысль о том, что искусственные интеллектуальные системы всё активнее входят в нашу жизнь. Это – уже реальность, а не проект или тем более фантазия. И эта реальность изменяется катастрофически быстро. В этой связи возникает много новых проблем, начиная от экономических, социальных, оборонных, политических и заканчивая вопросами о том, каково место человека в новой реальности, которую не так давно называли «электронным обществом», а сегодня – «цифровым». Необходимо учитывать, что цифровизация основана на использовании искусственных интеллектуальных систем, без которых она была бы невозможна. Цифровизация – наша жизнь, и в этой реальности мы существуем.

Кажется, что на пути цифровизации человечеству под силу реализовать то, о чем люди мечтали на протяжении столетий. Однако в этой связи возникает и множество вопросов, решение которых не всегда ясно. Например, в условиях цифровизации слишком удешевляется производство товаров, и ряд специалистов даже говорят о возможности того, что раньше считалось немыслимым: о возникновении ситуации, когда некоторые товары будут раздавать бесплатно. При глобальной цифровизации изменяется рынок как таковой, поскольку можно просчитывать результаты, которые раньше просчитать было нельзя. Ряд привычных профессий исчезнут. Распространение социальных сетей изменяет границу приватного и публичного пространства, что ставит под вопрос привычные представления об индивидуальной автономии и идентичности. Появляется возможность онлайн-образования. Но как мы смогли убедиться в условиях пандемии коронавируса, обучение в режиме онлайн не заменяет личного общения педагога с учащимися в реальном, а не виртуальном мире. Как показывают исследования, чтение лишь электронных учебников («цифровое чтение») и коммуникация через интернет не дают полноценного образования.

Но главный вопрос состоит в том, что происходит с самим человеком. Великий мыслитель И. Кант считал, что философия отвечает на три ключевых вопроса. Что я могу знать? Что я должен делать? На что я могу надеяться? Ответить на них означает ответить на главный вопрос о том, что такое человек. Данные вопросы требуют новых ответов. Попытаемся ответить на каждый из них.

Что я могу знать? По мнению адептов и фанатиков глобальной цифровизации, можно знать только то, что можно просчитать. С этой точки зрения полноценным знанием о чем-то могут обладать только интеллектуальные технические устройства, умные машины. Именно на такую машину нужно теперь полагаться даже в простейших жизненных ситуациях. Между тем индивидуальная идентичность конструируется с помощью знания о своем прошлом. Но происходящая цифровизация памяти создает принципиально новые возможности для вмешательства в нашу жизнь. Необходимо помнить о том, что наши действия оставляют цифровые следы, которые доступны внешним инстанциям, способным нас контролировать. В итоге мы перестаем быть владельцем информации о собственной жизни и ее хозяином. Личное пространство любого может быть взломанным, и в этом случае можно стать предметом управления со стороны других людей.

Что я должен делать в эпоху глобальной цифровизации? Очевидно, что в условиях, когда происходит утрата индивидуальной автономии, не мы будем решать, что нужно делать: умные технические устройства будут предписывать нам эти решения. Человек, который считал, что с помощью искусственного интеллекта он будет управлять миром, на самом деле становится управляемым с помощью этого интеллекта. Тем самым такой человек превращается из Бога, демиурга Вселенной, в раба, поскольку искусственный интеллект лучше знает, что нужно такому человеку.

На что я могу надеяться? Вроде бы надеяться и не на что, хотя надежду подают некоторые футурологи. В частности, на сайте Российского футурологического общества размещен прогноз о том, что к середине XXI века возникнет цифровое бессмертие,

т.е. человек получит возможность существовать вечно в виде электронного двойника. В современном мире немало людей (физики, нейрочеловеки, психологи, философы, футурологи) верят в подобную возможность и пытаются вести практическую работу в данном направлении. Такая идея представляется мне сомнительной.

В ряде работ я обосновал мысль о том, что, если бы человек стал бессмертным, он перестал бы быть человеком. И в этом обществе потеряли бы смысл особенно значимые ценности, делающие человека человеком: забота о ближних, самопожертвование, мужество, сострадание. Утрата смысла смерти ведет к утрате смысла жизни. И не следует забывать о том, что даже если бы удалось осуществить «цифровое бессмертие», о чем мечтают некоторые футурологи, в ситуации глобальной цифровизации жизни человек в любом случае оказался бы управляемым и контролируемым сверхмощным искусственным интеллектом. Если бы при этом у него осталось что-то человеческое, вряд ли он захотел бы такого бессмертия.

Человечество стоит перед лицом экзистенциальных вызовов. Один из них связан с глобальной цифровизацией на основе использования искусственного интеллекта и разработок в области когнитивных наук. Цифровизация приносит немало благ, решает ряд проблем, но она способна создать и угрозу в отношении бытия человека. Это означает, что предлагаемые в рамках цифровизации проекты нуждаются в философской экспертизе, направленной на гуманизацию информационных технологий и культивирование высших человеческих ценностей: свободы, личной автономии, достоинства, идентичности, творчества, понимания, взаимопонимания. Цифровизация может и должна вывести человека на новый уровень творческой жизни, а не привести его к самоуничтожению. Философия бывает максимально востребована в периоды культурных и познавательных кризисов. Мы живем именно в такое время. Сегодня время философии.