

УДК 1:001/159.964/004.338/519.782

DOI: 10.26456/vtphilos/2022.4.022

ЭКЗИСТЕНЦИАЛЬНЫЕ АСПЕКТЫ СИЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ПРОБЛЕМА СВОБОДЫ

В.Э. Войцехович¹, И.Р. Скиба²

¹ ФГБОУ ВО «Тверской государственной университет», г. Тверь

² Институт философии Национальной академии наук Беларуси,
г. Минск, Беларусь

Целью данной статьи является раскрытие систем искусственного интеллекта в экзистенциальном контексте и освещение влияния экзистенциальных категорий на процесс формирования интеллектуальных технологий.

Ключевые слова: искусственный интеллект, экзистенциальный контекст, свобода, интеллектуальные технологии.

Восстановить конкретное посредством сложения или организации элементов, которые от него абстрагированы, удастся не больше, чем в системе Спинозы получить субстанцию из бесконечного сложения ее модусов
Жан-Поль Сартр «Бытие и ничто» [10]

В контексте тематики интеллектуальных технологий в целом и искусственного интеллекта в частности экзистенциальный вопрос практически никогда не поднимается. Сфера технологий и экзистенциальный уровень бытия кажутся не имеющими вовсе никаких пересечений. Единственный довольно притянутый пример соприкосновения данных областей в информационном пространстве – цитата Илона Маска 2015 г. об опасности чересчур интенсивного развития систем искусственного интеллекта и ярко-коннотативное интерпретирование Маском искусственного интеллекта как «главной экзистенциальной угрозы» для человечества. Примерно в том же ключе высказывался и основатель Microsoft Билл Гейтс: «Я в том лагере, который обеспокоен перспективой развития супер-интеллекта. Сначала машины будут выполнять большую часть работы за нас, однако не будут обладать супер-интеллектом. Это хорошо, если мы правильно будем этим управлять. Через несколько десятилетий искусственный интеллект станет достаточно развитым, чтобы стать причиной для беспокойства» [6]. С подобными положениями были также согласны ныне покойный Стивен Хокинг, Джеймс Баррат (автор сборника конструктивных опасений [1]), Ник Бостром (в книге «Искусственный интеллект. Этапы. Угрозы. Стратегии» [3]) и многие иные исследователи, философы и общественные деятели.

Учитывая тот факт, что среди сугубо прагматически ориентированных корпораций и, соответственно, в глазах их руководства системы искусственного интеллекта полностью отождествляются с теоретическим объектом слабого искусственного интеллекта, а демаркация «сильный искусственный интеллект – слабый искусственный интеллект» в области материализации систем вовсе не проводится, то следует предположить, что в случае реализации проекта именно *сильного* искусственного интеллекта (даже в зачаточном его состоянии) реакция тех слоев общества, которые поверхностно разбираются в данной тематике, представлялась бы весьма неоднозначной.

И причины этого мы видим, во-первых, в неестественном генезисе систем искусственного интеллекта, т. е. в отсутствии опоры на феномен самоорганизации и контрпродуктивном редуцировании самоорганизации до уровня «машинного обучения»; и, во-вторых, в экзистенциальной проблеме свободы, которая теснейшим образом связана как с проблемами разработки систем сильного искусственного интеллекта, так и с восприятием его со стороны общественного сознания.

Для понимания детерминант настороженного отношения к стремительному развитию систем искусственного интеллекта, а также для того, чтобы способствовать прояснению ситуации относительно возможностей возникновения в будущем машины, действующей вопреки человеческой воле, необходимо определиться с самим понятием свободы в целом, а затем – и применительно к искусственному интеллекту.

О понятии свободы рассуждали со времен зарождения философии и говорят по сей день. В истории развития понятия свободы понятие творческой свободы постепенно вытесняет понятие свободы от препятствий (принуждения, казуальности, судьбы, Ананке). В древнегреческой философии (у Сократа и Платона) речь идет, прежде всего, о свободе в судьбе, затем о свободе от политического деспотизма (у Аристотеля и Эпикура) и о бедствиях человеческого существования (у Эпикура, стоиков, в неоплатонизме). Отдельно стоит школа киников (Диоген Киник), в которой свободой считалось «освобождение от всех необходимостей и природный, естественный образ жизни». В Средние века подразумевалась свобода от греха и проклятия церкви, причем возникал разлад между нравственно необходимой свободой человека и всемогуществом Бога.

Демокрит считал свободным «того, кто ни на что не надеется и ничего не боится». Датский философ Сёрен Къеркегор, основатель экзистенциального направления в философии, писал, что «страх есть головокружение свободы» [7, с. 160]. Бернард Шоу проводил прямую параллель между свободой и ответственностью в формулировке «Свобода означает ответственность. Вот почему большинство людей боится ее» [15]. Владимир Иванович Вернадский писал о связи свободы применительно к процессу научного поиска: «Для научного развития необходимо признание полной свободы личности, личного духа, ибо только при этом условии может одно

научное мировоззрение сменяться другим, создаваемым свободной, независимой работой личности» [4, с. 129]. Герберт Спенсер интерпретировал свободу сквозь призму социологии: «Никто не может быть совершенно свободен, пока все не свободны» [11, с. 175].

Однако описания свободы остаются метафорами, интерпретациями и не более. Определений (как родо-видовых отношений) нет. Только с эпохи Возрождения начинаются неполные определения.

В эпоху Ренессанса и последующий период под свободой понимали беспрепятственное всестороннее развитие человеческой личности. Со времен Просвещения возникает понятие свободы, заимствованное у либерализма и философии естественного права (Гоббс, Гроций, Пуфендорф; провозглашенный в 1689 г. в Англии «Билль о правах»), сдерживаемое все углубляющимся научным взглядом, признающим господство всемогущей естественной причинности и закономерности. В немецкой философии, начиная от Мейстера Экхарта, включая Лейбница, Канта, Гёте и Шиллера, Гегеля, а также немецкий идеализм до Шопенгауэра и Ницше, ставится вопрос о свободе как вопрос о постулате нравственно-творческого соответствия сущности и актуально достигнутого уровня ее развития. Так, Гегель подвергает критике представление о свободе как о произволе: «Когда мы слышим, что свобода состоит вообще в возможности делать все, чего хотят, то мы можем признать такие представления полным отсутствием культуры мысли» [5, с. 163]. По Марксу, коммунизм – общество свободных личностей. Согласно марксизму, человек мыслит и поступает в зависимости от побуждений и среды, причем основную роль в его среде играют экономические отношения и классовая борьба [8, с. 387]. Спиноза определяет свободу как любовь к Богу и любовь Бога к человеку: «Из этого мы ясно понимаем, в чем состоит наше спасение, или блаженство, или свобода – а именно в постоянной и вечной любви к Богу или в любви Бога к человеку» [12, с. 15]. Сходство Спинозы и Гегеля состоит в том, что оба мыслителя утверждают: свобода – в любви к Богу, т. е. эволюции человека в направлении к Абсолюту, в частности, в движении к высшим духовным ценностям и их реализации.

Тем не менее, несмотря на разнообразные определения свободы, негласно подразумевается положительный ответ на вопрос о том, свободен ли человек вообще – в плане некоего уровня самостоятельности в контексте поведения и принятия решений. По крайней мере подобный вывод напрашивается с точки зрения здравого смысла при обращении внимания на само наличие системы уголовной и административной ответственности. А также это делает возможным существование системы культуры, государственности и цивилизованного социума в противовес хаосу и анархии. Если вспомнить работу Зигмунда Фрейда «Недовольство культурой» [13] и провести некую аналогию между абсолютизированным либидо и вездесущей свободой, то можно заметить, что роль свободы человека в детерминировании

возможности существования социокультурной реальности весьма амбивалентна: с одной стороны, постулат о свободе личности позволяет структурировать общество и формировать законодательство, с другой стороны, и то, и другое реализуется в ущерб личной свободе каждого отдельно взятого члена этого общества. Подразумевая именно это, Анри Бергсон и выделял два типа моральной организации социума: открытую (довольно утопичную) и закрытую (актуально функционирующую) [2].

Гипотетическая свобода выбора является, за редкими исключениями, общепризнанной (пример исключения: классический психоанализ с детерминизмом бессознательного) и порождает, в частности, страх человечества перед искусственным интеллектом. Феномен страха перед полноценно функционирующим (даже слабым) искусственным интеллектом обеспечивается психическим механизмом проективной идентификации, описанным в отдельных вариантах психоанализа (Мелани Кляйн, Уилфред Байон). Это происходит при проецировании некоторых фрагментов внутреннего мира (психореальности) вовне (за грань субъективной телесности) с последующей идентификацией (Я-тождественностью) со спроецированным, но с сохранением некоторой степени осознания того, что внешний объект есть Другой (Ж. Лакан) и чуждый.

Соответственно, при таком процессе объект проективной идентификации отождествляется с базальным прообразом Плохого объекта и, как следствие, возникает пресекаторная тревога (страх преследования), т. к. объект воспринимается как угрожающий и желающий причинить вред. Разумеется, описанный выше процесс практически полностью бессознателен и не доступен для рефлексии «невооруженного» сознания, но феномен свободы играет в нем крайне значимую роль.

Здесь можно предположить, что проецирование и последующее инкорпорирование (помещение внутрь) в систему искусственного интеллекта естественной человеческой свободы (это психологический, а не технический процесс) и вызывает страх перед ним. Тем не менее на данный момент слабый искусственный интеллект лишен свободы выбора, т. к. подчинен алгоритму.

Возможно ли появление свободы у искусственного интеллекта в будущем? Если дальнейшие разработки будут вестись в направлении организации самостоятельности, независимости от человека искусственного интеллекта, его собственной эволюции, основанной на поисках, хаосе, мутациях, отборе, тогда за много поколений возможно такое усложнение интеллектуальной системы, что в ней откроется «свобода» (по Гегелю – как актуализация потенций), а также сознание, творчество, разум, т. е. появления сильного искусственного интеллекта.

Современные разработки идут именно в этом направлении: машинное обучение, глубокое обучение, нейронные сети основаны на все более «широких» алгоритмах, становятся все более сложными, их поведение все

более непонятно и непредсказуемо. В отдаленной перспективе это приближит искусственный интеллект к самостоятельному существованию.

Предполагаемое наличие у системы сильного искусственного интеллекта феномена свободы вызывает еще одну разновидность страха – страх неизвестности (неизведанного), описанный Г.Ф. Лавкрафтом. И при подобном подходе опасения выглядят действительно логичными.

На данном этапе целесообразно заметить, что мы избрали данный контекст развития текущего дискурса по причине наличия латентной взаимосвязи между специфическими особенностями обеспокоенности общества возможностью создания сильного искусственного интеллекта и ключевыми аспектами процесса его разработки на данном этапе, которые представляют собой пока слабые методы, не приближающие человечество к формированию систем сильного искусственного интеллекта, а реализующие «простое разбрызгивание одной области бытия на другую» (Сартр) [10]. «Бытие есть то, что оно есть и не есть то, что оно есть» – цитата из книги «Бытие и ничто» Ж.П. Сартра [10, с. 13]. Сартр, по всей видимости, хочет донести до читателя свою мысль о темпорализации бытия (интенциональности в будущее) и ее противопоставленности некой актуальной (здесь-и-сейчас) рефлексии: т. е. бытие как будто все время отстает само от себя и потому оно недостижимо для себя же самого. Здесь бытие можно представить в виде некоего художественного образа, некоего субъекта – тем не менее образ весьма наглядный. Действие всегда интенционально, также как и сознание. Но действие и осознание действия не параллельны, а последовательны. Сартр отождествлял решение действовать и само действие как некий единый акт. Решение, исходящее из свободного выбора, – базальный структурно-функциональный компонент свободного бытия-для-себя (сознания).

На основе этого мы выводим наше представление о свободе. Признавая сознание сущностью, а его существование – осуществлением процесса осознания, а также зная, что сущность предшествует существованию и постулируя синергетическую и хаотическую (а не механистическую) предопределенность решения случайными факторами информационного поля, мы выводим акт бессознательной свободы в виде гипотезы: «Человек не принимает решений, а лишь ознакомливается (путем осознания) с результатами уже принятых решений». Один из ведущих научных лидеров в области нейронауки и психолингвистики, а также теории сознания, доктор биологических наук, профессор Санкт-Петербургского государственного университета Т.В. Черниговская приводит верифицируемые научные доказательства того, как проходит процесс принятия решений: за некоторый промежуток времени в структуре центральной нервной системы осуществляется определенная нейро-гуморальная реакция, которая затем, с задержкой около двадцати секунд, передается импульсно, и человек осуществляет то или иное действие, считая, что это он сам сознательно принял данное решение в момент начала реализации осознаваемого поведенческого паттерна

[14]. И еще одна гипотеза, логически вытекающая из предыдущей и взаимосвязанная с ней: «Свободным является только бессознательное человека, а на сознательном уровне функционирует формальный детерминизм» [11]. И корни данной аксиомы уходят глубоко в историю восточной философии. В индуистской традиции, а затем в дзэн-буддизме красной нитью проходит ключевая мысль о необходимости упразднения феномена сознания для получения экзистенциальных и духовных дивидендов, таких как, к примеру, сатори (так называемое «состояние одной мысли»), нирвана (полное освобождение от беспокойства, свойственного «сансаре») и т. д.

Бессознательная свобода, по сути, это упразднение свободы в привычном смысле данного понятия и признание зависимости каждого конкретного решения от многочисленных неосознаваемых факторов и их совокупных воздействий: бессознательное выступает как актор принятия решения, *остальной* человек – как следствие, действие этого процесса. С философской точки зрения немаловажно, что у каждого решения априори ограничено число (в каждом случае совершенно точное и определенное) вариантов выбора конкретной интенции, а при наличии ограниченного количества вариантов речь может идти лишь об относительной свободе, а никак не об абсолютной. На предпочтение одного варианта другому влияет как актуальная ситуация, так и весь прошлый опыт в целом, а также темпорализация, которая, со своей стороны, сформирована тоже прошлым опытом, т. е. принятие решения со стороны системы обусловлено структурой (примеры: коннектом человека, искусственная нейронная сеть в процессе машинного обучения). Отсюда следующее предположение: «Структурированная система не способна принимать даже относительно свободных решений, не говоря уже об абсолютно свободных».

С позиций вышесказанного, логично будет вывести, что свобода, именно сознательная свобода, в качестве способности делать сознательный выбор и осознавать его причины весьма условно имеет право называться свободой, т. к. она бессознательна (бессознательная свобода), т. е. в некотором смысле вовсе не зависит от человеческого сознания. Абсолютная же свобода скорее абстрактна, чем реально имеет место быть. Отдаленными ее синонимами можно считать понятие нирваны в индуистской философии, понятие «сатори» в дзэн-буддизме, а также «чистое сознание» в феноменологии Эдмунда Гуссерля. Чтобы очертить некую конкретику в контексте различия между сознанием и бессознательным, мы предположим две из многочисленных демаркационных линий: сознание – структурировано, бессознательное – «бесструктурно» (имеет менее четкую структуру); сознание структурируется во времени, через время и в процессе осознания времени, а бессознательное – независимо от времени и с ним не контактирует. Вышесказанное вполне соответствует концепции классического психоанализа, в котором подразумевается, что бессознательное, в сущности, представляет собой сгусток желаний, которые жаждут немедленного исполнения вне зависимости от внешних условий и обстоятельств и лишь благодаря сначала

внешнему контролю, а затем уже и внутреннему – сознательному, появляется «принцип отсроченного удовлетворения», благодаря которому человеку удастся сдерживать желания и управлять ими. В целом же, в качестве рабочей гипотезы остановимся на понятии «бессознательной свободы» как единственно возможной. Под бессознательной свободой мы будем подразумевать *свободное принятие решений субъектом, с абсолютной невозможностью со стороны самого субъекта контролировать процесс принятия решений и каким-либо образом на него влиять.*

На данном этапе, с учетом вышесказанного, рассмотрим системы искусственного интеллекта. Засилье моделирования в области интеллектуальных разработок обусловлено, разумеется, прагматической необходимостью и повседневными нуждами – инвестиции в технологии должны быть рентабельны, но не стоит забывать, что в подобных случаях речь идет только лишь о системах слабого искусственного интеллекта – лого-машинах. Лого-машины, которыми только и представлены системы искусственного интеллекта на данном этапе развития интеллектуальных технологий, не обладают ни мышлением, ни интеллектом, ни, тем более, сознанием, наличие которого есть единственный критерий реального бытия системы сильного искусственного интеллекта – психо-машины. Таким образом, лого-машины априори не способны демонстрировать примеры даже относительной свободы, а все явления, которые могут быть похожи внешне на наличие чего-то внутреннего (интеллект, сознание, воля, желание и т. д.), являются внешней иллюзией. В тех же случаях, когда действия лого-машин являются не логичными и не прописанными в исходном коде – это абберрация, сбой, а не пример наличия чего-либо человекомерного. В лого-машине априори не может присутствовать ничего «внутреннего», «психического» в том смысле, в котором мы говорим о «внутреннем мире», «психическом» человека.

Причина этого – в генезисе. Повторим цитату Сартра, представленную в начале работы: «Восстановить конкретное посредством сложения или организации элементов, которые от него абстрагированы, удастся не больше, чем в системе Спинозы получить субстанцию из бесконечного сложения ее модусов» [10, с. 24]. Еще Платон в диалоге «Федон», цитируя фразы Сократа, подразумевал те же самые мысли, говоря, что человек не сложен из «костей, сухожилий и всего прочего» [9]. Рене Декарт в своем мысленном эксперименте также указывал на принципиальную невозможность «свести» человека в целом к телу.

В данных подходах мы не видим ничего эзотерического, эсхатологического или им подобного – мы не говорим о некоей «душе», монаде и т. д. Мы имеем в виду холистичный постулат о том, что целое больше всех частей вместе взятых, ибо обладает новыми – эмерджентными качествами, такими как, к примеру, сознание. И сформироваться подобная система способна только в процессе ауторегулирующегося саморазвития, самоорганизации, а никак не с помощью «сложения или организации элементов», т. е.

моделирования. В процессе многих миллионов лет эволюции психика прошла путь от прото-психических (простейших) качеств до осознания экзистенциальной категории бытия в контексте целостного мироздания и именно потому, что путь, полный синергетических скачков от бифуркации к аттрактору, был пройден, психика представляет собой то, что она есть сейчас – невероятно сложное образование, которое само не способно себя познать, ведь «оно всегда есть нечто большее, чем мгновение назад». Это естественный процесс эволюции. Он работает в контексте антропосоциогенеза; соответственно, именно генерирование (выращивание), а никак не моделирование следует применять при разработке систем сильного искусственного интеллекта в контексте техногенеза.

Генерирование отличается от моделирования тем, что при моделировании сначала создается некая максимально совершенная система, а затем она помещается в среду для обработки данных по формализованным схемам и выработки наиболее оптимальных способов реагирования на особенности среды, причем у системы, за редкими исключениями (пример исключения: эксперимент по просмотру нейросетью компании Google контента на площадке YouTube), имеется четко прописанная цель. Примерно так мы представляем антропный подход к формированию систем искусственного интеллекта – лого-машин. С другой стороны, при генерировании система не создается совершенной, а вынуждена сама себя создавать на основе максимально обширных и неформализованных схем и в тесном контакте с информационной средой, который первично происходит в момент появления у системы «способности» обрабатывать простейшие «информационные блоки». Примерно так мы представляем разработанный нами технотропный подход к формированию систем искусственного интеллекта – психо-машин.

И в данном контексте к генерированию напрямую относится понятие свободы. Ведь, если мы задаемся целью создать сильный искусственный интеллект – психо-машину, то мы обязаны обеспечить ее хотя бы относительной свободой, ибо это залог самой возможности возникновения феномена сознания – одного из ключевых критериев наличия человекомерности у машины. Отсюда следует, что в процессе разработки систем сильного искусственного интеллекта необходимо в первую очередь создавать «искусственное бессознательное». Далее, если человеческое структурированное сознание (сознательное) возникло в процессе естественной эволюции как некий неаддитивный, холистический феномен на основе бесструктурного бессознательного, предшествовавшего сознанию, то для обеспечения процесса естественного и саморегулируемого генезиса машинного сознания необходимо идти тем же естественным путем – путем генерирования систем, способных к самоорганизации, обладающих изначально одним лишь бессознательным.

Список литературы

1. Баррат Дж. Последнее изобретение человечества: искусственный интеллект и конец эры Homo sapiens / пер. с англ. Н. Лисова. М.: АНФ, 2015. 303 с.
2. Бергсон А. Два источника морали и религии / пер. с фр. М.: Канон, 1994. 382 с.
3. Бостром Н. Искусственный интеллект. Этапы. Угрозы. Стратегии / пер. с англ. С. Филина. М.: Манн, Иванов и Фербер, 2016. 404 с.
4. Вернадский В.И. Философские мысли натуралиста // ред. кол. А.Л. Яншин, С.Р. Микулинский, И.И. Мочалов; сост. М.С. Бастракова и др. М.: Наука, 1988. 520 с.
5. Гегель Г.Ф. Сочинения: в 2 т. // Философия природы / под ред. и с вступ. ст. А.А. Максимова. М.; Л.: Гос. изд-во, 1934. Т. 2. 683 с.
6. Гейтс Б. Предупреждение об опасности искусственного интеллекта. [Электронный ресурс]. URL: https://www.bbc.com/russian/rolling_news/2015/01/150129_rn_gates_ai_warning. (дата обращения: 12.06.2020).
7. Кьеркегор С. Страх и трепет. М.: Республика, 1993. 383 с.
8. Маркс К., Энгельс Ф. Соч.: в 55 т. / 2 изд. М.: Гос. изд-во полит. лит-ры, 1965. Т. 25, ч. 2. 552 с.
9. Платон. Федон // Сочинения: в 4 т. / общ. ред. А.Ф. Loseва, В.Ф. Асмуса, Л.Л. Тахо-Годи. СПб.: Изд-во С.-Петербур. ун-та: «Изд-во Олега Абышко», 2007. Т. 2. С. 11–96.
10. Сартр Ж.П. Бытие и ничто: Опыт феноменологической онтологии / пер. с фр., предисл., примеч. В.И. Колядко. М.: Республика, 2000. 639 с.
11. Спенсер Г. Автобиография // Образовательная библиотека; излож. А.Д. Коротаева. Санкт-Петербург, 1904. – № 6. – 178 с. (Серия 6)
12. Спиноза Б. Избранные произведения: в 2 т. М.: Госполитиздат, 1957. Т. 2. 727 с.
13. Фрейд З. Недовольство культурой. [Электронный ресурс]. URL: <https://www.litres.ru/zigmund-freyd/nedovolstvo-kulturoy/chitat-onlayn/> (дата обращения: 20.09.2022).
14. Черниговская Т.В. Язык и сознание: что делает нас людьми? [Электронный ресурс]. URL: <https://www.life-up>. (дата обращения: 02.02.2021).
15. Шоу Джордж Бернард: цитаты, афоризмы. [Электронный ресурс]. URL: <https://inpearls.ru/author/georgebernardshaw/page/15>. (дата обращения: 11.01.2022).

EXISTENTIAL ASPECTS OF STRONG ARTIFICIAL INTELLIGENCE. THE PROBLEM OF FREEDOM: MODELING OR GENERATION

V.E. Voytsekhovich¹, I.R. Skiba²

¹ Tver State University, Tver

² Institute of Philosophy of the National Academy of Sciences of Belarus, Minsk

The purpose of this article is the disclosure of artificial intelligence systems in an existential context and the coverage of the influence of existential categories on the formation of intellectual technologies.

Keywords: *artificial intelligence, existential context, freedom, intellectual technologies.*

Об авторах:

ВОЙЦЕХОВИЧ Вячеслав Эмерикович – профессор, доктор философских наук, профессор кафедры философии и теории культуры ФГБОУ ВО «Тверской государственный университет», г. Тверь. E-mail: Voytsekhovich.VE@tversu.ru

СКИБА Иван Рауфович – магистр педагогических наук, младший научный сотрудник Института философии НАН Беларуси (научный руководитель: кандидат философских наук А.В. Колесников), г. Минск, Беларусь. E-mail: gonzodzen@mail.ru

Authors information:

VOITSEKHOVICH Vyacheslav Emerikovich – Professor, PhD (Philosophy), Professor of the Department of Philosophy and Theory of Culture, Tver State University, Tver. E-mail: Voytsekhovich.VE@tversu.ru

SKIBA Ivan Raufovich – Master of Pedagogical Sciences, Junior Researcher at the Institute of Philosophy of the National Academy of Sciences of Belarus, (Scientific supervisor: Phd (Philosophical Sciences) Kolesnikov A.V.), Minsk, Belarus. E-mail: gonzodzen@mail.ru