



# Мораль для робота

## ЕВАНГЕЛЬСКИЙ ВЗГЛЯД НА ЭТИЧЕСКИЕ ПРОБЛЕМЫ ТЕХНОЛОГИЧЕСКИХ ИННОВАЦИЙ

Феоктист, епископ Переславский и Угличский

Искусственный интеллект — одно из самых спорных явлений, связанных с развитием цифровизации в современном мире. С одной стороны, все мы ежедневно сталкиваемся с его проявлениями — смартфоны и голосовые помощники, обработка персональных данных, диагностика заболеваний и многое другое. С другой стороны, есть этические вопросы разработки и использования интеллектуальных компьютерных систем, которые до конца не решены и тревожат православное сообщество. Существуют ли перспективы создания машинного разума в обозримом будущем, какие цифровые технологии допустимы с точки зрения Евангелия, в чем действительно могут помочь умные роботы и можно ли доверять их выводам относительно человека, читателям «Журнала Московской Патриархии» рассказал руководитель рабочей группы Межсоборного присутствия епископ Переславский и Угличский Феоктист.

### Занимательная философия

Последние десятилетия искусственным интеллектом интересуются люди самых разных специальностей — от философов и программистов до политиков, религиозных деятелей и домохозяйек. Наверное, не получится найти современного человека, который бы не имел своего мнения относительно технологий искусственного разума.

Огромное влияние на отношение большинства людей к этому достижению научно-технического прогресса оказывают произведения

писателей-фантастов и снятые по мотивам их книг фильмы и сериалы. То, что знают обычные люди об искусственном интеллекте, как правило, основывается лишь на художественных произведениях, в ряде случаев подкрепленных громкими заявлениями IT-гигантов<sup>1</sup>, опубликованными в СМИ.

Если же начать пристально оценивать то, что происходит в области разработки таких технологий, то мы довольно быстро убедимся в отсутствии реальной перспективы создания полноценного искусственного интеллекта в обозримом будущем. По этой причине размышления о том,

как нам с ним взаимодействовать, что это такое в философском и богословском смысле, и все связанные с этим фобии ничем не отличаются, скажем, от размышлений относительно безопасности воздушных перевозок людьми, жившими тысячу лет назад. Это интересно и увлекательно, но не более того.

Но есть реальные вопросы и проблемы современных технологий, на которых стоит сосредоточить свое внимание. Конечно, такого рода вопросов и проблем — множество, они затрагивают самые разные сферы нашей жизни.

Один из них состоит в определении границы между допустимым и недопустимым использованием технологий искусственного интеллекта с точки зрения Евангелия.

### **Алгоритмы, способные моделировать интеллектуальную деятельность человека**

Отрадно, что вопросы этического измерения применения современных технологий привлекают к себе все большее количество ученых. Для примера можно привести масштабный и качественный доклад «Этика и “цифра”: этические проблемы цифровых технологий», который был сделан в минувшем году Российской академией народного хозяйства и государственной службы при Президенте РФ<sup>2</sup>. Этой же академией в текущем году был подготовлен еще один доклад на схожую тему: «Этика и “цифра”: от проблем к решениям»<sup>3</sup>. Обе эти работы представляют собой серьезные научные труды, которые дают возможность погрузиться в проблему и детально разбирают вопрос этических вызовов, стоящих сегодня перед человечеством в связи с развитием цифровых технологий.

Авторы докладов предлагают несколько важных определений этики искусственного интеллекта, а также самого искусственного интеллекта. Эти определения очень важны, потому как они закладывают фундамент дальнейших рассуждений. Так, предлагается считать, что «искусственный интеллект — это наука и технология, включающая набор средств, позволяющих компьютеру на основании накопленных знаний давать ответы на вопросы и делать на базе этого экспертные выводы, то есть получать знания, ко-

торые в него не закладывались разработчиками. Область “искусственного интеллекта” является междисциплинарной и входит в комплекс компьютерных наук, а создаваемые на ее основе технологии относятся к информационным. В общем случае под искусственным интеллектом принято понимать алгоритмы и программные системы, способные моделировать интеллектуальную деятельность человека»<sup>4</sup>.

Этические вопросы — это вопросы прикладного характера, поэтому исследование этики цифровых технологий исключает какие-либо футурологические построения; следовательно, далее речь пойдет исключительно о так называемом слабом, или прикладном, искусственном интеллекте, под которым понимаются «прикладные системы на базе искусственного интеллекта, направленные на решение какой-либо конкретной задачи; фактически это все решения, которые окружают нас сейчас. Помощь такого искусственного интеллекта делает жизнь более комфортной, а работу более производительной. Подобные системы постоянно совершенствуются, уже сейчас многие конкретные виды работ системы с искусственным интеллектом делают лучше людей»<sup>5</sup>.

Слабый искусственный интеллект, как известно, противопоставляется сильному, или универсальному, который можно определить как «гипотетический искусственный интеллект,





способный решать любые интеллектуальные задачи и самостоятельно осуществлять целеполагание, — то есть сравнимый с человеческим по своим свойствам, а во многих смыслах даже превосходящий его»<sup>6</sup>.

Исследователи сходятся во мнении, что на настоящем этапе развития науки не обнаружены какие-либо теоретические препятствия для возникновения сильного искусственного интеллекта, «однако вряд ли стоит ожидать его появления в ближайшее десятилетие из-за необходимости решить большое количество сложных научных и практических задач реализации»<sup>7</sup>. Также стоит отметить, что многие исследователи настроены значительно менее оптимистично и утверждают, что появление сильного искусственного интеллекта — это вопрос не нескольких десятилетий, а нескольких веков.

### Проблема создания машинной этики

Теперь стоит дать определение этики искусственного интеллекта, которая должна лежать в основе как разработки, так и использования технологий искусственного интеллекта. Здесь важно напомнить, что задачу систематизации этических норм поставил в 2019 году Президент РФ В. В. Путин: «Сейчас во всем мире разворачивается дискуссия о социальных аспектах и последствиях использования искусственного интеллекта, это очень важная тема. Предлагаю профессиональному сообществу, компаниям

подумать над формированием свода этических правил взаимодействия человека с искусственным интеллектом»<sup>8</sup>.

Такая работа сейчас действительно ведется. Но вот в чем дело: необходимо отличать этику взаимодействия человека с искусственным интеллектом от этики самого искусственного интеллекта. Это две совершенно разные области: этику технологий искусственного интеллекта «от этики других областей отличает проблема этического поведения интеллектуальных систем в ситуации, когда ее решение касается людей. Принципиально важно, что система искусственного интеллекта способна самостоятельно принимать решения, касающиеся человека, и анализировать данные в таких объемах и с такой скоростью, как человек делать не в состоянии (следовательно, человек не может проверить верность решений). Соответственно, основная проблема — определение того, насколько решения, принимаемые интеллектуальной автономной системой, соответствуют этическим нормам, то есть насколько она этична». Таким образом, перед нами стоит проблема создания так называемой машинной этики. Необходимо снова сделать оговорку: речь идет не об этике «думающих» машин, а об этике программ, которые оперируют большим количеством данных и на основании этих данных предлагают человеку некие выводы. Другими словами, «основной вопрос — это этичность поведения самой системы, что вызывает целый ряд проблем. Предстоит





понять, в чем заключается этичность решения»<sup>9</sup>. Авторы доклада «Этика и “цифра”: этические проблемы цифровых технологий» признают, что «на этот вопрос очень сложно получить однозначный ответ от моральных философов. Однако без привлечения профессиональных специалистов по этике разработчики этической компоненты искусственного так и останутся в плену своих бытовых представлений о морали. Следовательно, принципиально важно установить связи между этическими концепциями, понятиями и сущностями и техническими элементами, в которых они могут быть воплощены, то есть решить вопросы конструктивных определений и онтологий»<sup>10</sup>.

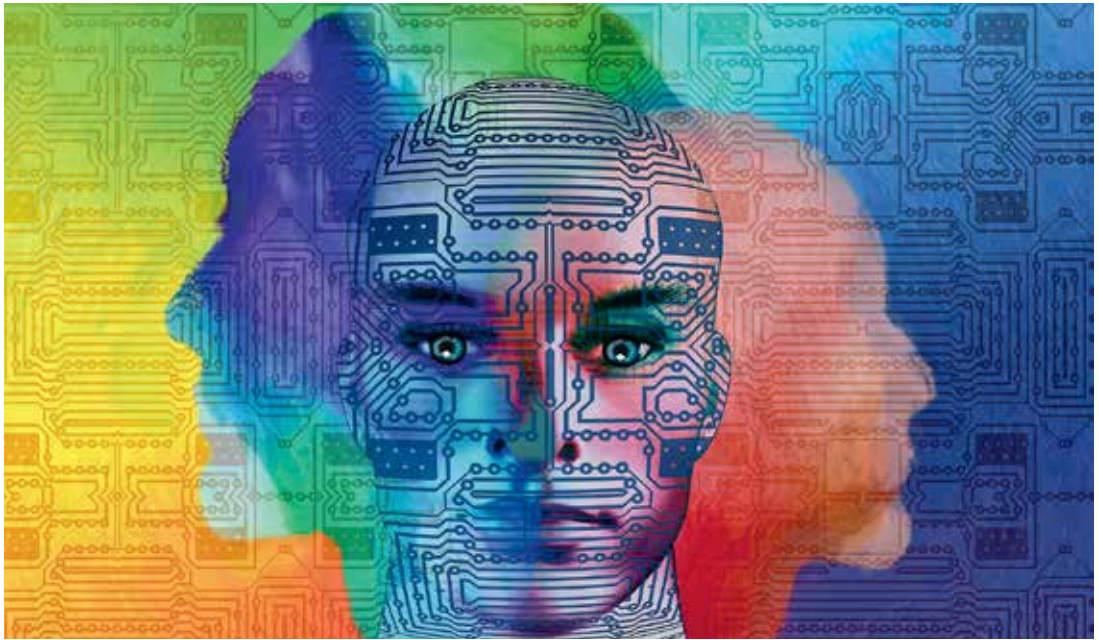
Существуют попытки создания математического аппарата для решения указанной проблемы, в рамках этой задачи особо «актуально рассмотрение и использование различных неклассических логик (например, многозначных), механизма многокритериальной классификации, вероятностных подходов и так далее»<sup>11</sup>. Но математический аппарат — это лишь инструмент, который может быть использован для реализации бесчисленного множества алгоритмов, тогда как суть вопроса состоит в том, что именно должно лежать в основе этики искусственного интеллекта.

## Божественная логика Христовых заповедей

Все мы прекрасно понимаем, что этические системы могут быть самыми разными, в том числе и такими, которыми оперируют запрещенные на территории нашей страны экстремистские группировки и секты. Но этика этих сообществ — ярко негативный пример, можно привести более нейтральные системы: этично то, что считает таковым большинство, или этично то, что определило как таковое государство. А может быть, этично то, что полагают таковым философы, занимающиеся разработкой проблем морали? Какую этику выбрать в качестве этики искусственного интеллекта? С христианской точки зрения этично и непререкаемо лишь то, что определил как таковое Господь наш Иисус Христос. Но здесь возникает серьезная проблема: этика искусственного интеллекта нуждается в формализации для того, чтобы она была программно реализована в цифровых системах, а Спаситель неоднократно и самыми разными способами подчеркивал недопустимость формализации морали. Современные Ему попытки создания этической системы Он неоднократно и очень жестко критиковал и обличал, говорил, что такая система уводит человека, ей подчиненного,

*Этика искусственного интеллекта нуждается в формализации, а Спаситель неоднократно подчеркивал недопустимость формализации морали.*





в противоположную от Бога сторону. В качестве наиболее яркого примера такого рода критики можно привести 23-ю главу Евангелия от Матфея, в которой собраны обличения Христом книжников и фарисеев. На основании Ветхого Завета они успешно выстроили этическую систему с использованием ряда формальных логических принципов, с помощью которых можно было вывести на основании Пятикнижия Моисея этически верное с точки зрения иудаизма отношение к абсолютно любому явлению жизни, а также к человеку или какому-либо материальному объекту. По большому счету любая другая этика, кроме этики христианской, подобна этике евангельских книжников и фарисеев, так как в ее основе всегда лежат некоторые законы, которые повинуются формальной логике. Примеры такого рода законов можно привести еще раз: правильно то, что считает таковым большинство; правильно то, что считает таковым государство. Это непротиворечивые законы, на их основе относительно просто выстроить логику принятия этических решений системами искусственного интеллекта.

Но не так обстоит дело с христианской этикой. Несмотря на то что Евангелие содержит в себе заповеди, которые дал Христос, эти заповеди невозможно свести к единой непротиворечивой системе, потому как даже на уровне формулировок они противоречат друг другу. Но это — на уровне формальной логики, то есть всех тех логик, которые способен выстроить,

постичь и формализовать человек. Христовы заповеди повинуются совершенно иным закономерностям, их логика принципиально не может быть системно выстроена, потому что она имеет не человеческое, а Божественное происхождение, соответственно, она не подчиняется ни человеческому восприятию, ни тем более языку. Христос выражал эту логику одним лишь словом: любовь. Полная же формулировка звучит так: *Возлюби Господа Бога твоего всем сердцем твоим, и всею душою твоею, и всею крепостию твоею, и всем разумением твоим, и ближнего твоего, как самого себя* (Лк. 10, 27). В каждом конкретном случае вопрос о том, какое именно решение и действие будет проявлением любви, зависит от бесчисленного числа факторов, которые принципиально невозможно свести в единую систему.

### **Подспорье для решения многих задач, но не всех**

На основании всего вышеизложенного можно сделать довольно банальный и предсказуемый вывод: любая запрограммированная этика порочна. Следовательно, любая система искусственного интеллекта, которая принимает решение относительно человека, — «плохая». А «хорошими» с христианской точки зрения можно считать лишь системы, которые являются вспомогательными, при этом должно быть соблюдено





условие прозрачности их алгоритма для конечного пользователя.

Конечно, такие умозаключения могут показаться имеющими в себе дух алармизма, пусть и сдержанного. Но есть и другая сторона вопроса: любая система искусственного интеллекта, которая способна делать вывод, касающийся человека, делает этот вывод на основании несравненно большего числа данных, чем используем мы. Вынося свой вердикт другому человеку, мы ориентируемся лишь на несколько деталей, далеко не все из которых можно считать объективными и заслуживающими внимания. К примеру, у нас может быть эстетическое неприятие тех или иных внешних особенностей человека, что мешает делать относительно него выводы, претендующие на объективность. Но тем не менее различные выводы мы все равно делаем, причем регулярно. Тогда как искусственный интеллект, встроенный, к примеру, в поисковые интернет-системы, ориентируется в своих выводах на запросы пользователей, а это уже совершенно иной уровень знания человека. Безусловно, он тоже не абсолютен, но при этом по объективным критериям он более достоверен, чем то впечатление, которое дают внешние характеристики человека. Но и здесь велика вероятность ошибки, хотя бы по той причине, что человек не сводится к сумме своих слов, мыслей и действий.

Исходя из всего вышеизложенного можно сделать заключение, что системы искусствен-

ного интеллекта могут быть прекрасным подспорьем в решении очень многих задач, но их выводы относительно человека — лишь еще один из множества факторов, которые используются, когда есть необходимость принять решение, касающееся другого индивидуума; само же принятие решения должно оставаться за человеком. При этом каждый из тех, кто вовлечен в процесс принятия такого рода решений, должен руководствоваться словами Христа Спасителя и всеми силами стараться видеть ближнего таким, каким его создал и видит Бог, то есть богоподобным. Так смотреть на человека может только другой человек при содействии Духа Святого, никакой машине, даже самой совершенной, это не подвластно.

#### ПРИМЕЧАНИЯ

<sup>1</sup> Довольно подробный разбор этих вопросов см., напр.: *Маркус Г., Дэвис Э. Искусственный интеллект: перезагрузка. Как создать машинный разум, которому действительно можно доверять*. М., 2020.

<sup>2</sup> *Этика и «цифра»: этические проблемы цифровых технологий*. М.: РАНХиГС, 2020. 207 с.

<sup>3</sup> *Этика и «цифра»: от проблем к решениям* / под ред. Е. Г. Потаповой, М. С. Шклярчук. М.: РАНХиГС, 2021. 184 с.

<sup>4</sup> *Этика и «цифра»: этические проблемы цифровых технологий*. С. 43.

<sup>5</sup> *Искусственный интеллект. К выбору стратегии*. М.: РАНХиГС, 2018 / Текст: электронный // URL: <https://www.ranepa.ru/images/News/2019-05/30-05-2019-AI-report-2019.pdf>.

<sup>6</sup> Там же.

<sup>7</sup> Там же.

<sup>8</sup> Путин: РФ способна стать одним из глобальных лидеров в сфере искусственного интеллекта // ТАСС: информационное агентство России: [сайт]. Текст: электронный. URL: <https://tass.ru/ekonomika/7096728>.

<sup>9</sup> *Этика и «цифра»: этические проблемы цифровых технологий*. С. 57.

<sup>10</sup> Там же.

<sup>11</sup> Там же. С. 60.