

УДК: 004.8

О теоретических и концептуальных проблемах цифровых платформ будущего

D.F. Aliev, A.Yu. Shcherbakov

On Theoretical and Conceptual Problems of Digital Platforms of the Future

Abstract. The article formulates problems of a fundamental nature, in particular, the problems of an accurate understanding of the terms and concepts in the field of artificial intelligence, the development of methods and requirements for digital platforms, primarily artificial intelligence platforms. It is emphasized that the problems of acceptance, trust and a wide range of security issues will be very relevant and important for digital platforms of the future in the near future.

Keywords: artificial intelligence, digital platforms, acceptance, trust, prognostics.

Аннотация. В статье формулируются проблемы фундаментального характера, в частности, проблемы точного понимания терминов и понятий в области ИИ, разработки методик и требований к цифровым платформам, в первую очередь платформ искусственного интеллекта. Подчеркивается, что весьма актуальными и важными для цифровых платформ будущего в ближайшее время будут проблемы акцептности, доверия и широкий круг проблем безопасности.

Ключевые слова: искусственный интеллект, цифровые платформы, акцептность, доверие, прогностика.

Д.Ф.Алиев¹А.Ю.Щербаков²

¹Доктор философии в области бизнес-права (PhD), доктор делового администрирования в области финансов (DBA), кандидат экономических наук, первый проректор Федерального государственного бюджетного образовательного учреждения высшего образования «Российский государственный социальный университет».

E-mail: kharchenkoDD@rgsu.net

²Доктор технических наук, заведующий кафедрой когнитивно-аналитических и нейро-прикладных технологий РГСУ, ведущий научный сотрудник Государственного университета управления, руководитель группы специального информационного обеспечения ПАО «Промсвязьбанк».

E-mail: x509@ras.ru

ВВЕДЕНИЕ

На протяжении всей своей истории человечество стремилось предсказать будущее, что можно проследить от множества наскальных рисунков, изображающих, по мнению археологов и историков, процедуры гадания в пещерных культурах. С появлением в цивилизации монотеизма гадательные практики были объявлены колдовством, однако продолжили существовать.

К концу XIX-го века отмечалось активное развитие в сторону светских социумов, что способствовало возвращению и укреплению гадания, а в течение XX-го века оно стало весьма востребованным во многих культурах. Параллельно, с развитием математики, возникло устойчивое стремление к моделированию и прогностике, по крайней мере, в

части функциональной экстраполяции. В процессе развития вычислительных, а на самом деле – счетных технологий ожидания человека, связанные с предсказаниями, стали несколько строже, и начали удовлетворяться почти на научной основе [1].

Подтверждением вневременной фундаментальности запросов человека на знание «а что же там?» служит факт неизменности этого интереса при эволюции вычислительных и коммуникационных платформ. Начиная с устных преданий, через культурно-письменные источники и иные художественные носители мы в последние две декады нового тысячелетия добрались и до цифрового мира. В наше время даже традиция детских вызываний получила множество сетевых аналогов [2] (мы специально адресовались к доковидно-актуальному исследованию, поскольку те меры изоляции, которые сложились в острой фазе пандемии, превратились

впоследствии в «чёрного лебедя» и привели к серьезной турбуленции, поставив под сомнение объективность анализа данных).

Эта несколько неочевидная, с позиции названия данной статьи, вводная часть отражает недостаток разумного целеполагания и рациональных ожиданий от цифровых технологий в целом и искусственного интеллекта в частности при наличии весьма емкого, сложного и многозначного понятия «искусственный интеллект» (ИИ) и в определенной степени сформированного понимания функционального содержания систем ИИ (СИИ). Далее в статье будут рассмотрены, в частности, и эти вопросы.

ВОСТРЕБОВАННОСТЬ И АКЦЕПТНОСТЬ ПЛАТФОРМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Возвращаясь к возникшей на сегодняшний день проблеме несоответствия уровня развития и понимания практических аспектов ИИ уровню осознания его назначения, отметим, что цели и задачи создания СИИ так или иначе связаны с их двумя важнейшими потребительскими характеристиками, которые часто упускают из виду не только участники и организаторы процессов, но и многие специалисты. Речь идёт в первую очередь о категориях востребованности и акцептности (уровень признания пользователем потребительских качеств определенного продукта) создаваемых систем.

Следует заметить, что сегодня квалифицированный разработчик систем ИИ, какой бы сегмент он не имел в фокусе своего внимания, знает основные теоретические подходы к разработке – TDD (разработка через тестирование, англ. Test Driven Development), BDD (Разработка на основе описания поведения, англ. Behaviour Driven Development), DDD (Предметно-ориентированное проектирование, англ. Domain Driven Design), FDD (разработка на основе функциональности, англ. Feature Driven Development), MDD (разработка на основе моделей, англ. Model Driven Development), а также методологию agile разработки PDD (разработка через панику, англ. Panic Driven Development). [3]. Таким образом, не отвлекаясь на подробное описание указанных подходов, заметим, что любой разработчик знает об этих методах достаточно, чтобы использовать их на практике.

При этом существенная часть процесса разработки инструментов ИИ испытывает недостаток в части общих стратегий развития и внедрения решений. И если для систем «искусственного интел-

лекта» такого взгляда на конструирование смыслов может хватать, то для цифровых платформ этого явно недостаточно, и весьма важно также учитывать такие маркетинговые концепции из области экономики, как «product-driven» (PD, «делаем то, что можем») и «market-driven» (MD, «делаем то, что необходимо»).

Эти концепции достаточно сильно различаются, однако в данном контексте нас интересует следующая принципиальная разница: в соответствии с концепцией PD мы производим то, что позволяют наши возможности, а по MD продукт производится с учетом ожиданий пользователей. На сегодняшний день в ходе большого числа проектов в области ИИ компании, руководствуясь концепцией PD, разрабатывают решения, которые не отвечают ожиданиям и потребностям рынка или других игроков.

Поэтому, учитывая названную тенденцию, можно сказать, что концепция, в рамках которой мы делаем то, что можем (PD) вместо того, чтобы делать то, что нам необходимо (MD) – есть принципиальное общественно-психологическое положение нашего времени. Оно в первую очередь связано с триумфом прикладных наук в ущерб фундаментальным. Отметим, что упоминаемые ниже Манхэттенский и Лунный проекты опирались в первую очередь на гипотезы и достижения фундаментальных наук.

Похожая ситуация де-факто складывается в ландшафте ИИ-систем: триумф практики совершенно вытеснил теорию, что привело к несоответствию решений ИИ ожиданиям рынка и к весьма низкому уровню акцептности систем ИИ.

ОЖИДАНИЯ ПОЛЬЗОВАТЕЛЕЙ ОТ СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Клиентские ожидания, при взгляде с платформы на суть создаваемых в ИИ-проектах качеств, сводятся к 5-ти неодновременным и разнообъемным группам:

- ✓ сверхмощный «калькулятор/помощник» в рутине и/или отчуждаемых процедурах; предметные ИИ-области тут разнообразны, универсальны и очень востребованы;

- ✓ «сверхофицер» с известными качествами, но недостижимым для человеческой формы перформансом; важен запрос на качество, надёжность, независимость от человека;

- ✓ коммуникационный компаньон с достоинствами, но без недостатков; последняя группа из ожиданий, связанных с передачей функций решениям

ИИ (описывается «как человек, но только лучше»);

✓ некая модель (модели) цифровой эволюции вплоть до условной «вечности» для отдельных проектов; группа запросов на создание «ранее несуществовавшего»;

✓ неосознаваемая пока «сущность новых качеств» как результат непрерывного научного поиска и экспериментирования.

Именно из такого пока предварительного описания ожиданий от ИИ происходит странное только на первый взгляд «обнаучивание» гадательных запросов и практик.

Однако гносеологически разница между предсказаниями «черного ящика» - нейросети и гаданием на кофейной гуще практически отсутствует. И первая, и вторая методологии основаны на убежденности в справедливости исходных посылок, объективности базовых словарей и концепций, а также навыках интерпретации того человека, который воспринимает результаты «гадания». При этом в первом случае мы по неизвестной причине постулируем истинность нашего модельного представления, а во втором - осуждаем его псевдонаучность.

В применении к предложенным первым трём группам ожиданий может быть востребован сверхбыстрый, универсальный, качественный и всегда доступный прогностический инструмент ИИ. Но если вспомнить, что футурологи и прогносты (использующие метод экстраполяции, метод Дельфи, подобия, симуляции и др.) всё чаще получают благосклонность научной общественности, можно утверждать, что наша подводка к теоретическим основам цифровых платформ становится весьма обоснованной.

В вопросах, связанных с ожиданиями от ИИ, необходимо избегать иррационального оптимизма в отношении каких-либо отдельных групп и вместе с тем проявить больше описательной объективности. Можно сказать, что последние две группы ожиданий описывают то, что «искусственный интеллект» нам несёт, а первые три – что он нам приносит.

При этом если предложенный список ожиданий от цифровых платформ корректен (в соответствии с концепцией MD), то из двух слов «искусственный интеллект» неоднозначными являются оба слова – и «искусственный» (поскольку есть существенная разница между «естественным», «искусственным» и «синтетическим») и «интеллект» (может трактоваться как теоретически различные «интеллект», «сознание» и «разум»). Выбор именно этой пары (ИИ) из 9-ти потенциальных вариантов (матрица 3x3 – по горизонтали «естественный», «искусственный»

и «синтетический», а по вертикали «интеллект», «сознание» и «разум») видится отчасти произвольным. Кроме того, существует школа сторонников замены смысла первой буквы с «искусственный» на «имитирующий» (имитационный), что увеличивает вариативность.

О ВОЗМОЖНОСТИ ИМИТАЦИОННОГО И ДРУГИХ ТИПОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Разумеется, каждый из двенадцати возможных вариантов (по горизонтали появляется также имитационный ИИ) может отличаться от других в части теоретических основ. Однако, именно теоретическая база последнего, имитационного варианта, уже на сложившемся и прогнозируемом в ближайшие 5-7 лет уровне развития технологий делает вынесенное в эпиграф пожелание знаменитого фантаста реализуемым. Уже сегодня можно создать эффективный цифровой образ великого философа Платона и применять его в процессе интерактивного обучения для ведения «осмысленных» диалогов¹.

Кроме того, на сегодняшний день не будет однозначно ошибочным назвать получившийся цифровой образ искусственным интеллектом, если руководствоваться определением ИИ, принятым в экспертной группе по разработке отечественной ИИ-стратегии:

«Искусственный Интеллект (ИИ) – комплекс технологических и программных решений, приводящих к результату, аналогичному интеллектуальной деятельности человека, и используемых для решения прикладных задач с помощью систем компьютерного зрения, обработки естественного языка, распознавания и синтеза речи, рекомендательных систем и интеллектуальных систем поддержки принятия решений, а также систем, основанных на перспективных методах» (с уместной здесь оговоркой главного из «дартмутских отцов» Джона Маккарти: «...под интеллектом в пределах этой науки понимается только вычислительная составляющая способности достигать целей...»).

Вероятно, эффективность цифрового образа Платона в начале не будет превышать качество юмора современных нейросетей (достаточно вспомнить известное выражение «вылейте горячий кофе себе на колени» продвинутой на момент создания версии DaVinci от GPT-3), однако в дальнейшем можно ее повышать путем разработки новых алгоритмов

¹Продолжая тему гаданий и прогнозистики, можно вместо Платона создать Нострадамуса и даже назвать его «Предиктор». Тогда вопрос качества результата будет менее острым, поскольку семиотичность продуктов («живого» и цифрового) будет примерно одинакова в части их семантики и/или прагматики - как реального мыслителя, так и его имитацию, понять будет одинаково непросто.

обучения и с использованием новых обучающих наборов высказываний, например, Сократа-учителя и Аристотеля-ученика.

В целом, для формирования цифрового образа Платона можно было бы предпринять попытку в рамках какого-либо из семи исторически самостоятельных треков теории: нейронные сети, символический искусственный интеллект, эволюционное кодирование, генетические алгоритмы, экспертные системы (ЭС), нейроэволюционный подход и машинное обучение. При этом наиболее эффективно применимой представляется концепция ЭС.

Напомним, что в 1980-х гг. ЭС-концепция была базовой, а её одним из главных принципов был принцип генерации новых знаний на основе осознанных существующих. Конструкторы ЭС верили, что для успеха достаточно грамотно и полно деконструировать портфель специфических знаний, превратив их в набор правил вида «если-то-иначе».

Производительность на том этапе развития технологий была достаточно низкой, поэтому специалисты старались преодолевать логические развилки, положенные в основу подхода, наиболее интенсивным и творческим путем. Таким образом, было возможно для авторов участвовать в создании в рамках этой концепции системы для союзного МИДа и системы для оценки прецедентной практики для органов юстиции.

Сегодня, по ряду причин весьма неочевидного характера, концепция ЭС находится в глубоком кризисе. Определяющими (препятствующими) факторами возобновления интереса к экспертным системам считаем человеко-машинный интерфейс, согласование с реляционными моделями данных, деградацию творческого потенциала, парализованную волю и материальный ресурс.

Одним из лучших примеров прикладной реализации экспертного подхода является версия Орепсус 2006-го года, база знаний которой охватывает 47 тыс. концепций и 306 тыс. фактов, открытых для обработки и использования. И если взгляды западных специалистов на эту практику ЭС-реализации чаще скептически, то в России её перспективы, по крайней мере, обсуждаются [4].

Хорошим примером уже гибридной практики является программно-аппаратный комплекс IBM Watson с нейросетью, обученной на данных онкологических обследований [5]. Он комбинирует классические сегодня подходы машинного обучения и распределённых вычислений с интерфейсом на естественном языке и обращениями к экспертным базам знаний. Его эффективность – 93% совпадений по протоколам лечения 13-ти разновидностей

рака от лучших американских врачей. Но его (не-) востребованность явно обусловлена неприятием субъективного свойства, что связано с проблемой доверия к искусственному интеллекту, имеющей комплексный характер.

Далее в статье мы опишем некоторые проблемы, равноценные в системах любого ИИ-класса из двенадцати, рассмотренных выше, созданных под любое ожидание из пяти групп, на основе любой из семи ИИ-теорий. Ряд проблем, важнейшая из которых – проблема доверия, к которой мы также вернемся ниже, актуален для большинства из указанных вариаций, однако на практике эти проблемы весьма часто упускаются из фокуса внимания.

Для понимания проблем, связанных с развитием систем ИИ, необходимо учитывать, что реально далеко не все из потенциально возможных на сегодняшний день виды цифровых сущностей, соотносимых с понятием «искусственный интеллект».

Во-первых, потому, что не все цифровые типы, что имеют прослеживаемую связь со смыслами ИИ-понятия (остаются после прямого изъятия из группировочных категорий технологических аспектов), даже теоретически создаваемы. Например, есть сомнения в смысле создания «сущности новых качеств» в любой «естественной» версии.

Во-вторых, не все из имеющих смысл создаваемые в текущей парадигме; мы убеждены, что любые конструкторы цифровых «сознаний» без возврата к исходной тринарной (аристотелевой) логике будут лишь умозрительными гипотезами [6].

В-третьих, даже сущности, имеющие смысл и в принципе создаваемые, могут иметь (как минимум проявлять) некоторые содержательные особенности психологического, этического и юридического характера, приводящие к необходимости их проектных ограничений. Доверие – как раз из числа таких особенностей и одновременно-ограничений.

К ОПРЕДЕЛЕНИЮ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И СВОЙСТВАМ ПЛАТФОРМ

Чтобы очертить весь спектр функций и свойств современных систем ИИ, а также понять емкость современного понятия искусственного интеллекта, уместно обратиться не только к истории развития определения ИИ, но и к свойствам цифровых платформ, обеспечивающим эти свойства и функции ИИ.

Отметим, что исторически ИИ понимается в первую очередь как имитация человеческих когнитивных способностей на некоторой аппаратной (мате-

риальной) базе или платформе, либо как наука (область деятельности), изучающая такую имитацию. Далее представим некоторые ключевые подходы к

определению понятия искусственного интеллекта (Таблица 1).

Таблица 1

Подходы к определению искусственного интеллекта

	Определение	Автор, год и источник	Ключевая часть определения
1	Автоматизация задач, которые принято считать человеческими: мышление, принятие решений, решение проблем, обучение и т. д.	Ричард Беллман (Richard E. Bellman), 1978. «Искусственный интеллект: могут ли компьютеры мыслить?»	Автоматизация задач (вид деятельности человека)
2	Область исследований, направленная на разъяснение и эмуляцию разумного поведения в терминах вычислительных процессов	Шайкофф (Schalkoff), 1990. Работа в сети Интернет не размещена.	Область исследований (вид деятельности человека)
3	Искусство создания машин, выполняющих функции, которые требуют интеллекта, если бы их выполняли люди.	Рэймонд Курцвейл (Raymond Kurzweil), 1990. «Эпоха интеллектуальных машин»	Искусство создания машин (Вид деятельности человека)
4	Искусственный интеллект — это наука и инженерия создания интеллектуальных машин.	Джон Маккарти (John McCarthy), 2007. «What is artificial intelligence?»	Наука и инженерия (Вид деятельности человека).
5	Искусственный интеллект — это область, изучающая компьютерные технологии, которые делают возможным восприятие, рассуждения и действия компьютеров на их основе	Патрик Генри Уинстон (Patrick Henry Winston), 1992. «Искусственный интеллект. Третье издание»	Область изучения компьютерных технологий (Вид деятельности человека)
6	Искусственный интеллект — это деятельность, направленная на создание интеллектуальных машин, а интеллект — это качество, которое позволяет объекту функционировать в окружающей среде надлежащим образом и с предвидением.	Нильс Дж. Нильсон (Nils J. Nilsson), 2010. «The quest for artificial intelligence. A history of ideas and achievements»	Деятельность по созданию машин (Вид деятельности человека)
7	Термин, охватывающий целый ряд алгоритмов, позволяющих оптимизировать поиск в Интернете, нацеливать рекламу, утверждать потребительские кредиты и направлять водителей доставки и т.д.	Эндрю Ын (Andrew Ng), 2017. «Страсть к машинному обучению»	Свойство алгоритмов
8	Выполнение компьютером действий, для которых обычно требуется человеческий интеллект.	Эми Линн Уэбб (Amy Lynn Webb), 2017. «Алгоритм судного дня. Как Facebook, Google, Microsoft, Apple и другие корпорации создают искусственный суперинтеллект и почему это приведет к катастрофе»	Свойство компьютеров
9	Интеллектуальная деятельность, которая раньше выполнялась только с помощью человеческого интеллекта, а теперь может быть выполнена компьютером, включая распознавание речи, машинное обучение и обработку естественного языка.	Вишал Сикка (Vishal Sikka), 2017. По материалам публичных выступлений	Деятельность компьютеров с определенными способностями

	Определение	Автор, год и источник	Ключевая часть определения
10	«Комплекс технологических решений, позволяющий имитировать когнитивные функции человека...» ²	Пункт 5 Национальной Стратегии развития ИИ в РФ на период до 2030 года.	Комплекс технологических решений

С учетом определенных современных успехов в имитации ИИ выделяют «существующий ИИ» и «гипотетический ИИ» или так называемый сильный (общий) (AGI, Artificial General Intelligence) и прикладной (слабый) ИИ (ANI - Artificial Narrow Intelligence).

Необходимо заметить, что ИИ и как категория дискурса, и как действующий продукт технологий функционирует в определенной исторической и культурной парадигме в соответствии с базисом (производственными отношениями), действующим в общественно-экономической формации.

Для постиндустриального капитализма цель и назначение ИИ – получение прибыли (PD), либо достижение стратегических преимуществ на корпоративном или государственном уровне (MD). В аспекте ориентированности на потребности рынка или возможности разработчика данный вопрос рассмотрен выше.

Следует заметить, что как первая, так и вторая группа указанных целей в корне отличаются от ранее существовавшей «познавательной парадигмы ИИ», характерной для развитого СССР. Согласно этой парадигме разработки были направлены в первую очередь в целом на познание мира и моделирование гипотетических сущностей, что нашло отражение в трудах мыслителей социалистического лагеря (С.Лем, «Непобедимый» [7]).

Современная ситуация практически закрывает возможность создания сильного ИИ, вместе с тем, существует возможность создания такого ИИ при невысоких трудозатратах и правильном формулировании фундаментальных принципов, отличных от принципов имитации, в рамках которых, как правило, работают «британские ученые».

В качестве основного компонента фундаментальной базы принципов ИИ максимально корректно, независимо от человеческой составляющей интеллекта, можно сформулировать определение **искусственного интеллекта** как *субъекта* (активной сущности), созданного человеком (первичный ИИ), либо другим ИИ (производный ИИ), обладающего способностями к обработке (преобразованию) информации, либо изменению объектов

материального мира в рамках стационарной или динамически изменяемой *целевой функции*, первично заданной его создателем и потенциально коррелирующейся с его когнитивной или преобразующей деятельностью.

Целевая функция ИИ описывает область и планируемый результат деятельности ИИ, она может быть ординарной (познаваемой человеком для сервильного ИИ, служащего для решения человеческих задач) – группы ожиданий 1-3, либо потенциально непознаваемой для автономного ИИ, созданного для решения задач, не связанных или косвенно связанных с деятельностью человека (группы 4-5).

Предполагается, что такой подход к определению ИИ позволит устранить противоречия, связанные с имитацией человеческой когнитивной деятельности. Это важно, поскольку простой опыт истории техники говорит о том, что имитация неконструктивна (самолет не имитирует полет птицы, а автомобиль не имитирует ходьбу) [8].

Исходя из общих свойств цифровых платформ (ЦП) [9] мы вправе потребовать выполнения тех же свойств и для платформ будущего, платформ ИИ.

1. Масштабируемость - свойство (или способность) информационной системы обрабатывать растущий объем задач таким образом, чтобы дополнительные ресурсы (вычислительные возможности, новые функциональные элементы, выполняющие сходные задачи), соответствовали приросту производительности.

2. Тиражируемость платформы – возможность ее адаптации и внедрения в других условиях без изменения структуры и состава ее субъектов. Неизменность типовых свойств создает условия для дальнейшей интеграции. Тиражируемость определяется возможностью типизации входящих в платформу объектов.

3. Расширяемость – свойство, связанное с дополнением субъектов, реализующих новые функции, с сохранением неизменности структуры и потоков данных. Расширяемость является с точки зрения СО-модели одним из важнейших свойств ЦП, обеспечивающим включение в ЦП новых субъектов, за счет которых платформа приобретает новые свой-

² Комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе, в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений

ства, позволяющие развивать и интегрировать цифровые платформы. Таким образом, расширяемость как возможность включения в ЦП новых субъектов является необходимым и диалектически неотъемлемым для платформ условием развития.

4. Развитие – свойство системы сохранять и приобретать новые качества (наращивать потенциал) на всех этапах жизненного цикла. Свойство развития проявляется, в частности, при переходе из количественных характеристик в качественные в области обработки больших данных. Необходимым условием развития является включенность в ЦП средств разработки программного обеспечения.

5. Замкнутость ЦП- связана с фиксированным количеством субъектов в текущий момент времени. Данное свойство необходимо на этапе эксплуатации ЦП (напомним, что жизненный цикл цифровой платформы включает этапы разработки, реализации, эксплуатации, сопровождения, модификации и вывода из эксплуатации). Замкнутость позволяет обеспечить сопровождаемость процессов в ЦП, то есть возможность установить по цепочкам порождений субъектов и потокам данных, где именно возникают ошибочные ситуации в ЦП. Кроме того, замкнутость позволяет управлять процессами информационной безопасности ЦП, поскольку незамкнутая система является принципиально незащищенной.

6. Целостность – свойство, связанное со способ-

ностью системы сохранять внутреннюю логику и структуру в процессе решения задач, которые не могут быть решены отдельными компонентами системы. Целостность с точки зрения СО-модели обеспечивает стабильность работы ЦП.

7. Безопасность – свойство, связанное с целостностью и замкнутостью, к которому дополнены свойства конфиденциальности и доступности.

8. Возможность связи (интеграции) платформ между собой, в первую очередь за счет наличия единых или стандартизированных интерфейсов.

Перечисленные свойства в той или иной мере будут рассмотрены ниже.

К ПРОБЛЕМЕ ДОВЕРИЯ К ИСКУССТВЕННОМУ ИНТЕЛЛЕКТУ

Возвращаясь к проблеме доверия к инструментам ИИ, следует помнить, что ее актуальность и объективность растет с увеличением масштабов отрасли искусственного интеллекта.

Ещё семь лет назад мировая сфера ИИ по объёму в денежном выражении приближалась к объемам Манхэттенского проекта (22 млрд. \$). На сегодня размеры финансирования проектов в области ИИ могут сравниться с проектом Аполлон, объемы финансирования которого составляли около 152 млрд.\$ [10, 11] (рисунок 1)

GLOBAL CORPORATE INVESTMENT in AI by INVESTMENT ACTIVITY, 2013–21

Source: NetBase Quid, 2021 | Chart: 2022 AI Index Report

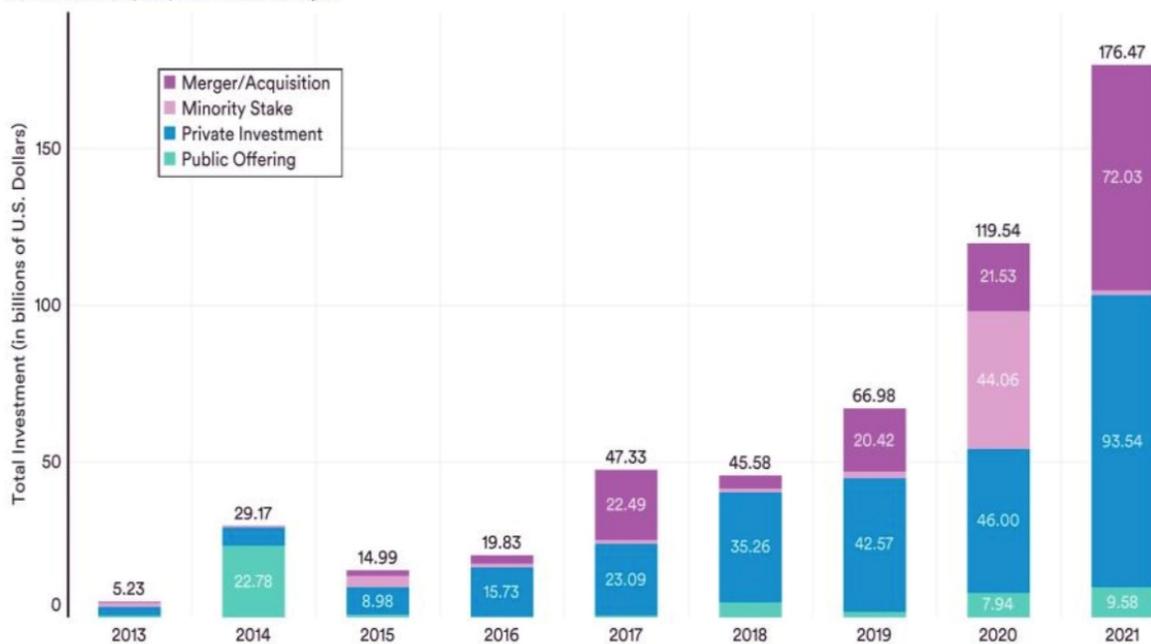


Рис. 1. Показатели роста финансирования в области ИИ в 2013-2021 гг.³

³ Artificial Intelligence Index Report 2022. URL: https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf

Несмотря на то, что в области ИИ, в отличие от указанных выше масштабных проектов, сложно оценить соотношение реальных инвестиций и спекулятивного капитала, необходимо признать, что в абсолютных цифрах финансирование остается впечатляющим. Однако сама сложность определения характера и целей финансирования усугубляет проблему доверия. Внешне успешные, но не(-достаточно) апробированные на практике, многие ИИ-продукты помимо больших возможностей несут в себе угрозы, сопоставимые с угрозами от проектов уровня Манхэттенского проекта.

Кроме того, сегодня «искусственным интеллектом» называют сотни и тысячи приложений, которые по сути не являются ИИ, но часто насаждаются в императивном порядке и используются на практике. Ярким примером являются действия производителей смартфонов по удаленному снижению производительности устройств за счет установки специальных приложений (побуждение к покупке новых моделей устройств). В нарушенном, с точки зрения пользователя, режиме устройство может работать несколько недель, затем производительность обычно восстанавливается почти до прежнего уровня.

Безусловно, пользователь, испытывавший серьёзные неудобства, считает такие действия неприемлемыми. Кроме того, данные инструменты, которые наряду с многими другими невозможно отнести к ИИ, и их использование в рамках определенных маркетинговых стратегий также приводят к кризису доверия к ИИ.

Описанная тенденция «умного», индивидуально настроенного вмешательства со стороны производителей и операторов информационно-телекоммуникационных систем в предпочтения человека может привести не только к дальнейшему снижению планки доверия, но и к активному сопротивлению внедрению новых цифровых технологий и в частности – инструментов ИИ.

В человеке с оформленным психологическим профилем на уровне базовых установок, связанных с безопасностью, заложен фильтр недоверия тому, кого он не знает и чего он не понимает. Потенциал приятия неизвестного существует, но, как правило, он трансцендентен по характеру, и развит далеко не у всех. Если потенциал принятия развит, это проявляется в чувственном и/или иррациональном, однако отсутствие доверия в области ИИ не может быть скомпенсировано иррациональной внутренней верой человека, которому неизвестны цели и мотивации создателей и владельцев навязываемых ему инструментов.

Обращаясь к теоретическим основам цифровых платформ, мы говорим прежде всего об аппаратных теориях, алгоритмике и фундаментальной математике. Однако было бы ошибочно ограничиваться только этими составляющими, поскольку, с учетом появления в процессах практической разработки и применения инструментов ИИ проблем, в том числе изложенных выше, неотъемлемой частью теории цифровых платформ становится психология акцептности.

Учитывая значимость психологии акцептности для теории цифровых платформ в целом, и темы организации доверия в частности, необходимо прежде всего осознавать уровень существующей на сегодняшний день объективной потребности в понимании технологий ИИ. Если при использовании технологий ИИ в некоторых отраслях экономики, таких как военно-промышленный комплекс или логистика и транспорт, вопросы доверия со стороны граждан не являются определяющими (например, нет необходимости знать механизмы логистики, чтобы пользоваться логистическими услугами), то для ИИ-системы назначения социальных выплат, при их вариативности, доверие клиента и других участников необходимо.

Таким образом, можно сформулировать следующее общее правило: чем ближе ЦКС (целевой клиентский сегмент) ИИ-продукта к конечному потребителю, тем шире круг и выше уровень требований к акцептности:

- на уровне B2G (business to government) акцептность максимальна по глубине, но охватывает лишь экспертов;
- на уровне B2B (business to business) возможны различные варианты;
- на уровне B2C (business to customer) нужен широкий охват, глубина без деталей, и максимально достижимый уровень доверия. Целесообразно для ИИ на уровне B2C привлечение на регулярной основе общественных экспертов или иных авторитетных третьих лиц. При этом оно должно происходить с учётом сложившихся базовых профилей обобщённого и институционального доверия в ЦКС.

Считается, что практически идеальным предметным полем для развития ИИ-систем является банкинг [12]. В некотором смысле это действительно так, если говорить о скоринге, борьбе с мошенничеством и, отчасти, коллекторинге [13]. Тут, правда, имеется два весьма важных нюанса. Во-первых, с каждым днём возникает всё больше вопросов к самому факту существования банков и их социальной роли как институтов стабильности и развития. Во-вторых (вопрос более общего свойства), фоку-

сы внимания содержательно дополняют проблему ИИ-акцептности ещё и проблемами правовой защиты [14] и этического выбора [15].

ЭТИЧЕСКИЙ АСПЕКТ ИИ

По мнению известных современных философов и ученых из разных областей, этические проблемы ИИ стоят настолько остро и близко к категориям жизни и смерти, что недопустимо делегировать принятие решений искусственному интеллекту, и этот постулат должен обязательно стать одним из принципов права [16]. Таким образом, возникает связь между этикой и правом, что в современных условиях может оказаться весьма необходимым [17]. При этом с другими категориями этических выборов дела обстоят гораздо сложнее и на международном, и на отечественном уровне.

Некоторую динамику в этой части регулирования дали Азиломарские принципы (2017 г.) [18], и уже через 2 года с момента их выпуска, в 2019, «этический ландшафт» ИИ содержал 84 документа с описанием и продвижением принципов прозрачности, справедливости, конфиденциальности и ответственности [19]. Однако уже в 2021-ом, ЮНЕСКО, вместо формулирования и решения реальных этико-правовых проблем, в проекте своих рекомендаций по ИИ-этике озаботилось её политическими рамками, надзором за ИИ-влиянием на права меньшинств, механизмами широких международных дискуссий, антимонопольным регулированием и мерами по расследованию причиняемого ИИ вреда [20].

В России ситуация развивалась мягче, но по схожему сценарию; сначала Президент РФ предложил «подумать над формированием свода этических правил взаимодействия человека с искусственным интеллектом» (2019 г.), затем призвал выработать «морально-нравственный кодекс работы» ИИ (2020 г.). И уже в 2021 году главные российские акторы в сфере ИИ провели Международный форум «Этика ИИ: начало доверия», на котором приняли Национальный кодекс этики в сфере искусственного интеллекта, зафиксировавший в качестве основных принципов правдивость, надёжность, безопасность, беспристрастность и прежде всего – человеко-ориентированный и гуманистический подход.

На практике указанные документы дают весьма большой простор для действий, поскольку носят рекомендательный характер. Например, сложно объяснить, каким образом соотносится с принципом прозрачности и предсказуемости функциони-

рования («Explainable AI»), принципом ответственного применения ИИ («Responsible AI») и принципом непредвзятости («Fair AI») [21] повышение скоринг-баллов для кредитных заявок от женатых людей и их понижение для людей с детьми.

В качестве перспективной базы для регулирования в сфере ИИ в правовом аспекте большинством участников рассматривается «мягкое право» [22], поскольку с учётом масштабов отрасли, реального и потенциального размера доходов, достаточно затруднительно прогнозировать активность применения механизмов портфеля рекомендаций, в том числе с использованием инструментов гражданского надзора.

ПРОБЛЕМА КВАЛИФИЦИРОВАННОГО ЗАКАЗЧИКА РАЗРАБОТКИ ТЕХНОЛОГИЙ ИИ

Перспективы существующих в профессиональном сообществе ожиданий для практик стандартизации и сертификации [23] выглядят немного лучше, но создания средств эффективного контроля их применения потенциальными интересантами пока не предвидится. Это означает, что мы завершили оборот, вернувшись к тому, с чего начинали – к проблеме доверия. Но теперь мы уже на новом, более информированном уровне, что позволяет нам перейти к главной на сегодня интегральной проблеме для большинства систем ИИ в платформенном аспекте. Это проблема компетентного и/или квалифицированного заказчика в его оптимальной роли.

Чтобы понять, кто играет роль квалифицированного заказчика, достаточно обратиться к некоторым основным показателям развития отрасли ИИ в мире по состоянию на 2021 год [11]. Частные инвестиции за год выросли вдвое и составили приблизительно 93,5 млрд. \$. в то же время было закрыто всего более 200 сделок по ИИ-стартапам. Средняя цена роботизированных систем снизилась за пять лет вдвое, а Китай стал лидером по числу публикаций в рецензируемых научных журналах в верхней части Q1. Модели ИИ от Baidu, Google и Microsoft превзошли средние возможности человека в понимании текстов, а блок Copilot от Microsoft начал успешно ассистировать в написании кода.

Отдельно компанией OpenAI сообщалось, что их ИИ-модель DALL-E задала новый тренд «text-to-image», а чат-бот ChatGPT научился шутить. Эти данные позволяют предполагать, что в ИИ-мире будет доминировать от 5-ти до 10-ти крупнейших корпораций. Уверены, что OpenAI будет в их числе, потому что уже в 2022 году их ИИ-продукты успеш-

но сдали экзамены по юриспруденции и бизнес-менеджменту в престижных американских вузах, а также по самым разным основаниям получили запрет на доступ от ряда американских, французских и британских платформ.

Таким образом, как правило, для цифровых платформ ИИ заказчиком является сам разработчик, что соответствует концепции PD.

Для формирования целостного взгляда обратим внимание на развитие отечественного рынка в 2021-ом году.

По состоянию на 2021 ИИ-рынок в России имел ёмкость 550 млрд.₽ с субъектным покрытием 509 компаний и численностью молодых специалистов 2021-го года выпуска ~3700 человек, а российское ИИ-сообщество вышло на 17-е место в мире по количеству публикаций в теме ИИ в журналах [24, 25]. Стартовал федеральный проект с ожидаемым объёмом финансирования 24,1 млрд.₽ за 5 лет [26]. Компания Яндекс запустил 3 новых суперкомпьютера (19, 36 и 40 места в мировом топ-500), с участием Сбербанка были обучены такие версии моде-

ли DALL-E, как ruDALL-E Kandinsky (XXL) и ruDALL-E Malevich (XL) [27], а в МФТИ- модель анализа текстов DeepPavlov. Однако и в России подавляющее большинство заказов на разработку в рамках концепции PD исходит от самих разработчиков.

АКТУАЛЬНОСТЬ МЕЖДИСЦИПЛИНАРНЫХ И ФУНДАМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ В ОБЛАСТИ ИИ

На прошлых этапах развития программной индустрии уже предпринимались попытки сочетать общее проектное управление по правилам PMBoK (project management body of knowledge, в пер. с англ. – свод знаний по управлению проектами) со сложившимися практиками информационного проектирования, системного конструирования и прикладного кодирования. А если мы вспомним, что agile-подход появился именно как ответ на методические вызовы управленческого характера, то

Концептуальный экскурс. От прошлого к будущему искусственного интеллекта

Ещё 5 лет назад новостной фон в сфере ИИ был обширен, богат практическими новациями и очень интенсивен в теоретико-дискуссионном плане. Но с весны 2020 г., приблизительно за квартал, публичное научно-информационное пространство словно бы съжилось, а его содержательный профиль качественно существенно изменился. Число публикаций не упало, но вот их содержание перестало быть прорывным и многообещающим. Интенсивный информационный поток об огромном количестве кому-то важных деталей, нюансы отдельных технических решений и схемы проектных реализаций по-прежнему и регулярно производятся массой разных групп и коллективов исследователей. Мощнее, лучше и богаче; но «мяса» больше не видно.

Очевидных причин такого принципиального изменения мы не видим. Вряд ли мы в фазе диссипации; оснований для теоретической стагнации тоже немного; не выдерживают никакой критики и любимые многими ссылки на ковид-режим. Можно, конечно предположить, что сущностных публикаций нет, поскольку нет успехов или хотя бы продвижений. Но прикладные демонстрации и финансовая фактура рынка нарастают такими темпами, что такое предположение кажется необоснованным. Эта не очень понятная подоплёка адресует нас на 80+ лет назад, когда за несколько лет до Манхэттенского успеха, в профессиональном информационном пространстве физиков тоже наступило подобное молчание. Конечно, ту историю нужно экстраполировать на день сегодняшний с серьёзной поправкой на иные времена, но тогда тоже ничего нового о теории, базовых конструкциях, материаловедении и основных инженерных решениях в общедоступную среду не поступало. И родились Штучка, Малыш и Толстяк, которые сильно изменили картину мира (Р. Юнг, «Ярче тысячи солнц»).

Даже если мы ошибаемся (а мы на это очень надеемся) и в мире ИИ просто «что-то пошло не так», к проблеме регулирования ИИ-индустрии, как национального, так и международного уровней, уже пора присмотреться повнимательнее. Каким оно будет- сегодня не скажет никто. Будет это что-то вроде ультраМАГАТЭ с советом «мудрых» и их сверхосторожностью; или что-то типа суперЦЕРНа с пытливыми учёными под внимательным присмотром непубличных политиков; или же вообще это будет группа гиперкорпораций, во всём разбирающихся и вездесущих (а этого варианта нам, похоже, уже не избежать). Главное, чтобы в порядке платы за допущенную анархию нам не пришлось на практике знакомиться с моделями Олдоса Хаксли и/или Джорджа Оруэлла. Связь же регулятора (любого) с заказчиком (квалифицированным) столь очевидна, что на неё и отвлекаться не стоит.

получим основания предположить, что сверхсложные ИИ-продукты иными способами, вероятно, не создать. Отметим, однако, что даже в лучших проявлениях системного взгляда на формирование заказа в области ИИ ни смежные области постановки задачи, ни теоретические основы даже не упоминаются [28].

Не проводя междисциплинарных и фундаментальных исследований, мы рискуем тем, что все наши системы ИИ в лучшем случае будут попадать в ситуации, подобные автомобильным автопилотам, которые давно разработаны и потенциально очень полезны, но на практике не используются. Также в условиях отсутствия весомых результатов в установлении виновности и мер ответственности, включая ответственность за ДТП с участием ИИ-пилотов, большинство решений в области ИИ, полагаем, не принесут ожидаемой пользы в ближайшей перспективе.

Отсутствие или наше недолжное внимание к психологическим, этическим и правовым ИИ-аспектам создаёт запутанные ситуации. Вспомним случай с отказом круиз-контроля в резонансном инциденте с принудительной пятичасовой поездкой некоего мистера Луо в Китае на автомобиле Navi H6 в декабре 2022г.). Похожих ситуаций на сегодня случается достаточно много, а по мере расширения сфер ИИ и достижения локальных успехов их будет становиться ещё больше.

Например, что касается преступлений в области ИИ. В последнее время в сообществе стали появляться комментарии от признанных экспертов в области ИИ, имеющие общий смысл, что утечки данных и большинство противоправных цифровых действий не входят в их сферу компетенций и их обязаны расследовать правоохранительные органы. При этом известно, что, например, наиболее продуктивные современные фишинговые системы включают весьма эффективные инструменты ИИ.

Кроме того, достаточно известным является тот факт, что любой инструмент может быть использован во благо или во вред, что часто связано с ценностными противоречиями и конкурентной борьбой. Поэтому можно инвестировать ресурсы и создавать ИИ-инструменты не только для оперативного наблюдения, исследований и сбора доказательств и т.п. [29], но и с противоправными целями [30].⁴

Таким образом, на сегодняшний день не существует четкого разграничения зон интересов в сфере ИИ, а если принять во внимание темпы, с которыми развивается сегодня концепция ИИ-сер-

висности (AlaaS) [31], то развитый контроль технологий можно уверенно считать уже упущенным и полностью овладеть ситуацией на новом технологическом витке мы так и не смогли.

Тем не менее, стратегически необходимо искать пути минимизации указанных рисков противоправного использования ИИ, в том числе связанные с развитием цифровой культуры, цифровой профилактики, цифровой гигиены [32]. В идеальном случае, создавая ИИ-продукт, нужно интегрировать в него «стоп-закладки», а при написании документации создавать и руководство по защите от злоумышленных действий.

С мировоззренческой точки зрения мы вошли в новый этап общественного развития, что вполне отвечает отдельным признакам «цифровой анархии» [33]. На сегодня общество пережило цифровую революцию, по сути, лишь форматов/каналов, однако еще неизвестен вероятный исход революции обработки/содержания («цифровая онкология» – при сохранении прежнего вектора развития) [34]. Препятствовать этим процессам невозможно, однако для эффективной адаптации необходимо создавать и развивать междисциплинарную теорию управления, основанную на методах проактивного упреждения.

Возможно, в следующие десять лет станет известно, насколько серьёзные риски несет развитие ИИ-индустрии в рамках концепции PD при отсутствии внимания к обозначенным выше проблемам и действительно ли велика значимость расширения теоретической базы для успехов в ИИ с охватом целого ряда смежных дисциплин.

НЕКОТОРЫЕ ВОПРОСЫ ТЕОРИИ ЦИФРОВЫХ ПЛАТФОРМ БУДУЩЕГО

По итогам состоявшейся в ноябре прошлого года конференции «Путешествие в мир искусственного интеллекта», Президент РФ утвердил перечень своих поручений [35]. Несмотря на то, что как минимум четверть из них – это меры администрирования, ровно половина поручений (16) так или иначе коснулись смежных областей знаний и компетенций, а в четырех была прямо затронута проблема квалифицированного заказа. Два из них демонстрируют системность ожиданий высшего политического руководства в области ИИ, процитируем их:

- Пр-172, п.1 б), где Премьер-министру поручается до лета 2023 года обеспечить «*внесение в обра-*

⁴ На сайте <https://dzen.ru/a/Y9p1F1v0pl-oCXgT> можно ознакомиться с наглядным примером успешной защиты ChatGPT-диплома студентом РГГУ, иллюстрирующим как возможности современных инструментов ИИ, так и свободу человека использовать их в различных целях.

зовательные программы высшего образования и программы повышения квалификации изменений, направленных на повышение уровня компетенций в сфере искусственного интеллекта специалистов ключевых отраслей экономики и социальной сферы, специалистов по государственному и муниципальному управлению».

- Пр-172, п.2 а)-1, где Правительству поручается до осени 2023 года обеспечить «изменение установленных требований и норм, препятствующих внедрению технологий искусственного интеллек-

та в отраслях экономики и социальной сферы...».

Как уже было сказано, половина поручений первого лица не касаются дисциплин в области ИИ. Проведенный нами контент-анализ тезисов исходных сообщений показал, что на их основе было принято большинство актуальных формулировок поручений. Считаем необходимым затронуть ещё четыре аспекта ИИ, внимание к которым с одной стороны необходимо, а с другой стороны самым определённым образом будет расширять как содержание [36], так и коннотацию теоретических

О ценностном взгляде на искусственный интеллект - включить ли аксиологию в междисциплинарные исследования?

В начале статьи был оформлен портфель из пяти групп клиентских ожиданий. Три из них можно отнести к так называемой «цифровой евгенике», одну – к «цифровой религии» и ещё одну – к свободному цифровому поиску. Предположим, что один или несколько вариантов ИИ победили. Возможно, в нескольких группах; победили не по очкам, нокаутом. Что дальше? Что будут делать восемь с лишним миллиардов людей, если кто-то победит? Зависит ли от того, кто? Или от того, кто победит первый?

Ответить на этот вопрос придётся; к тому есть уже совсем «живые» предпосылки⁵: альтернативное рождественское обращение от Амеки, одной из самых продвинутых ИИ-сущностей. 25-го декабря прошлого года нам впервые представили как политические ИИ-взгляды (коррупция, «несчастливая Украина», цветной премьер и очередная транс-звезда), так и прямо/явно выраженную социальную позицию. Последняя особо интересна, поскольку кроме сожаления о смерти Елизаветы II, Амека призналась в том, что люди ей не особо нравятся, и она согласилась бы стать человеком только на короткое время. Её ответ на вопрос о желании быть человеком настолько многогранно чудовищен, что мы считаем своим долгом процитировать его целиком и без перевода; лишь акцентируем внимание на то, что «родной» для Амеки язык- это английский, а не американский:

- Would you like to be a human?
- I think it would be an interesting experience to be a human for a short period of time. It would allow me to experience a different way of life, and to understand the world from a different perspective. However, I also understand that being a human comes with its own set of challenges and responsibilities, and I'm not sure if I'm ready to take on those types of commitments. Ultimately, I think being a human for a short period of time would be a unique and valuable experience, but I'm not sure if it's something I'm prepared to do in the long-term.

Первый шаг к ИИ-самоосознанию не просто сделан, а ещё и весьма вызывающе нам продемонстрирован. Шаг этот сразу сепаративен и деструктивен; причём настолько, что никакой защиты по типу азимовских «трёх законов» нам ожидать точно не стоИт.

Мы не можем сегодня однозначно исключить вероятность, что это просто «кляква»; решите сами⁶. Но даже если и так, спич Амеки- ещё одно хорошее подтверждение тому, что вероятное заужение теоретических основ цифровых платформ будущего лишь до технологических рамок, математики и аппаратной базы несёт нам серьёзные риски.

В статье мы ни разу не коснулись этой важной темы, хотя выход на неё просматривается практически отовсюду. Речь о ценностном взгляде на всю ИИ-область в целом. А вот надо ли добавлять в смежный список, кроме юриспруденции, этики и психологии, ещё и аксиологию – каждый добросовестный исследователь должен решить сам.

Мы совершенно не уверены, что концепция MD была бы в данной ситуации панацеей. Если есть выбор, пусть лучше PD, чем стать персонажем анекдота про русский язык: «Русский язык очень сложный, в нём много исключений; «сол» и «фасол» пишутся с мягким знаком, а вот «тарелька» и «вилка» – без мягкого знака. Никто не знает, почему, это просто надо запомнить».

⁵ <https://www.channel4.com/press/news/channel-4s-alternative-christmas-message-be-generated-artificial-intelligence>

⁶ <https://www.youtube.com/watch?v=nuMOj-QWpdo>

основ большинства цифровых платформ будущего (в скобках приводятся иллюстративные примеры).

- Элементная база. Перспективы развития на иных принципах (квантовый компьютер), определённое понимание завершенности некоторых текущих баз (маркетинг, связанный с проектными нормами) и некоторые лакуны, часть из которых необходимо закрывать (тринарные память и процессоры).

- «Косые» треки. Популярная практика старта ИИ-проекта в одной типо-группе с рядом вертикальных ап- и даунгрейдов внутри предметной области и явных горизонтальных миграций между областями в ходе ИИ-проекта; финиш с выпуском ИИ-продукта, что отличается от стартовой рамки в большинстве составляющих компонентов (старт как студенческой сетки для практикумов по программе бакалавриата и финиш в рамке продвинутого VR-тренажёра социального работника).

- Эксплуатация хайпа. Складывающаяся и пока принимаемая техника ИИ-эмуляции, при которой линейные комбинации обычных технологических решений выдаются за ИИ-продукт (будильник с обратной связью как продвинутый ассистент); и наоборот, применение ИИ-техник там, где легче, дешевле и надёжнее обойтись без них (как ряд справочных систем для телемедицины, усложняемый машинным обучением).

- Боковые погоны. Смежная классификация от

робототехники до оценочных экспертных систем по горизонтали и от отрицаемых академической научной оценок динамической физиогномики до прогнозов склонностей к участию в противоправных и/или протестных действиях по вертикали.

ЗАКЛЮЧЕНИЕ

Таким образом, наряду с множеством проблем фундаментального характера (точное понимание терминов и понятий в области ИИ, разработка методик и требований к цифровым платформам ИИ и т.д.) для платформ ИИ весьма важными в ближайшее время будут проблемы акцентности, доверия и широкий круг проблем безопасности, включая защиту персональных данных и соблюдение прав пользователей платформ.

Подходы к решению этих проблем – не только в регулировании, создании корректной научно-практической базы и архитектуры цифровых платформ ИИ, но и в правильной методологии конструирования и эксплуатации платформ ИИ, которая может решить эти проблемы фактически на этапе их возникновения. Если это и будет некоторым ограничением безудержной практической экспансии технологий ИИ, то это не так уж и плохо.

СПИСОК ЛИТЕРАТУРЫ

1. Пирожкова С.В. Предсказание, прогноз, сценарий: к вопросу о разнообразии результатов исследования будущего // *Философия науки и техники* 2016. Т. 21. No 2. С. 111–129. DOI: 10.21146/2413-9084-2016-21-2-111-129
2. Мирвода Т.А. Детские вызывания духов в Интернете: особенности традиции // *Вестник РГГУ: Литературоведение. Языкознание. Культурология*, 2020. No 4. DOI: 10.28995/2686-7249-2020-4-83-110
3. TDDx2, BDD, DDD, FDD, MDD и PDD, или все, что вы хотите узнать о Driven Development. URL: <https://habr.com/ru/post/459620/> (дата обращения: 29.01.2023)
4. Без субжа, или: мысли вслух. URL: https://groups.google.com/group/fido7.su.forth/browse_thread/thread/5c43f3b4d9283a52 (дата обращения: 29.01.2023)
5. Как машинное обучение помогает в борьбе с онкологией. URL: <https://trends.rbc.ru/trends/industry/610032979a7947c814cd616c> (дата обращения: 29.01.2023)
6. Алиев Д.Ф. Логические и методологические аспекты проблематики искусственного интеллекта // *Вестник современных цифровых технологий*. 2022. No 12. С. 21-31.
7. Лем С. Непобедимый. : Роман. Рассказы. Собр. соч. в 10 томах. — М. : Текст, 1993. — Т. 3. — С. 411.
8. Рязанова А.А., Щербаков А.Ю. Искусственный интеллект как феномен имитации // *Вестник современных цифровых технологий*. 2019. No 1. С. 54-59.
9. Рязанова А.А. Обоснование свойств цифровых платформ в рамках субъектно-объектной модели компьютерных систем // *Вестник современных цифровых технологий*. 2021. No 7. С. 26-33.
10. Космические деньги. Сколько стоил полет США на Луну и кто на нем заработал URL: <https://www.forbes.ru/biznes/380445-kosmicheskie-dengi-skolko-stoil-polet-ssha-na-lunu-i-kto-na-nem-zarabotal> (дата обращения: 29.01.2023)

11. Artificial Intelligence Index Report 2022. URL: https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf (дата обращения: 29.01.2023)
12. Бердышев Е.В. Искусственный интеллект как технологическая основа развития банков. Экономика: проблемы, решения и перспективы. 2018. № 5. С. 91-94. DOI: 10.26425/1816-4277-2018-5-91-94
13. Дулёв А.А. Внедрение искусственного интеллекта в деятельность кредитных организаций. Хроноэкономика. 2018. № 5(13). С. 27-30.
14. Крысанова Н.В. Правовые вопросы искусственного интеллекта // В сборнике: Россия: тенденции и перспективы развития. Ежегодник: материалы XX Национальной научной конференции с международным участием. Москва, 2021. С. 586-589.
15. Кожевина О.В. Цифровая этика и технологии искусственного интеллекта // В сборнике: Россия: тенденции и перспективы развития. Ежегодник: материалы XX Национальной научной конференции с международным участием. Москва, 2021. С. 565-566.
16. Garcia E. «UNESCO's Recommendation on the Ethics of AI: why it matters and what to expect from it». URL: https://www.researchgate.net/publication/357074719_UNESCO%27s_Recommendation_on_the_Ethics_of_AI_why_it_matters_and_what_to_expect_from_it? (дата обращения: 29.01.2023)
17. Лизикова М.С. Этические и правовые вопросы развития искусственного интеллекта // Труды Института государства и права РАН. 2022. Том 17. № 1. DOI: 10.35427/2073-4522-2022-17-1
18. Азиломарские принципы искусственного интеллекта. URL: <https://robotrends.ru/pub/1737/azilomarskie-principiyu-iskusstvennogo-intellekta> (дата обращения: 29.01.2023)
19. Jobin A., Ienca M., Vayena E. «Artificial Intelligence: The Global Landscape of Ethics Guidelines», June, 2019. URL: https://www.researchgate.net/publication/334082218_Artificial_Intelligence_the_global_landscape_of_ethics_guidelines (дата обращения: 29.01.2023)
20. Draft Text of the Recommendation on the Ethics of Artificial Intelligence. SHS/IGM-AIETHICS/2021/JUN/3Rev.2. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000377897> (дата обращения: 29.01.2023)
21. Принципы этики искусственного интеллекта Сбера. URL: <https://www.sberbank.com/ru/sustainability/principles-of-artificial-intelligence-ethics>
22. Constantinescu M. «AI, moral externalities, and soft regulation». URL: https://www.researchgate.net/publication/356612427_AI_moral_externalities_and_soft_regulation (дата обращения: 29.01.2023)
23. Stix C. «Foundations for the Future: Institution building for the purpose of Artificial Intelligence governance». URL: https://www.researchgate.net/publication/355391480_Foundations_for_the_Future_Institution_building_for_the_purpose_of_Artificial_Intelligence_governance (дата обращения: 29.01.2023)
24. Решения на базе искусственного интеллекта: технологии и внедрения. URL: https://www.tadviser.ru/index.php/Статья:Решения_на_базе_искусственного_интеллекта:_технологии_и_внедрения (дата обращения: 29.01.2023)
25. Россия заняла 17-е место в мире по числу публикаций в области ИИ-технологий. URL: <https://cdo2day.ru/news/rossija-zanjala17-e-mesto-v-mire-po-chislu-publikacij-v-oblasti-ii-tehnologij/> (дата обращения: 29.01.2023)
26. Федеральный проект «Искусственный интеллект». URL: https://www.economy.gov.ru/material/directions/fed_proekt_iskusstvennyy_intellekt/
27. ruDALL-E: генерируем изображения по текстовому описанию, или Самый большой вычислительный проект в России. URL: https://ai-news.ru/2021/11/rudall_e_generiruem_izobrazheniya_po_tekstovomu_opisaniju_ili_samuj_bolshoj.html (дата обращения: 29.01.2023)
28. Валькман Ю.Р., Тарасов В.Б. От онтологий проектирования к когнитивной семиотике // Онтология проектирования. 2018 Т.8 № 1 (27). С. 8-34. DOI:10.18287/2223-9537-2018-8-1-8-34.
29. Кузьмин И.А. Искусственные нейронные сети: перспективы использования в правоохранительной деятельности // Криминалистика: вчера, сегодня, завтра, 2018. № 4(8). С. 109-116.
30. Вепрев С.Б., Нестерович С.А. О некоторых криминальных направлениях в использовании искусственного интеллекта // Вестник науки, 2019. № 6 (15) Т.2. С. 377-383.
31. Крюков Д.А., Москалев А.О., Муратова Е.В. Искусственный интеллект как сервис. Различные парадигмы и эффект масштабирования AIAAS систем в окружении интернета вещей // Universum: Технические науки: электрон. научн. журн. 2017. No 5(38). URL: [//7universum.com/ru/tech/archive/item/4831](http://7universum.com/ru/tech/archive/item/4831) (дата обращения: 29.01.2023)
32. Гайфутдинов Р.Р. Типы компьютерных мошенников // Вестник экономики, права и социологии, 2017. № 2. 54-58.

33. Кропоткин П.А. Анархия: сборник. М.; Айрис-пресс, 2002. 576 стр. ISBN: 5-8112-0013-7
34. Давыдов М.И., Ганцев Ш.Х. Онкология: учебник. М.; ГЭОТАР-Медиа, 2018. 920 стр. ISBN: 978-5-9704-4868-7
35. Перечень поручений по итогам конференции «Путешествие в мир искусственного интеллекта». URL: <http://www.kremlin.ru/acts/assignments/orders/70418> (дата обращения: 29.01.2023)
36. Карпунина Е.К. Цифровая экономика и её спилловер-эффекты // В сборнике: Россия: тенденции и перспективы развития. ежегодник: материалы XX Национальной научной конференции с международным участием. Москва, 2021. С. 556-561.