

УДК 004.89

doi: 10.15622/rcai.2025.020

## ЕЩЕ ТРИ ВОПРОСА<sup>1</sup> (НА ПОНИМАНИЕ), АДРЕСОВАННЫЕ «ТОВАРИЩАМ ПО ПАРТИИ»

В.К. Финн (*v.k.finn@yandex.ru*)

М.А. Михеенкова (*m.mikheyenkova@yandex.ru*)

М.И. Забейайло (*m.zabzhailo@yandex.ru*)

Федеральный исследовательский центр  
«Информатика и управление» РАН, Москва

Обсуждаются представления об интеллектуальности систем искусственного интеллекта (ИИ), методах оценки качества формируемых ими результатов, а также некоторые проблемы организации экспертизы проектов и решений в области ИИ.

**Ключевые слова:** искусственный интеллект, исследования и разработки, экспертиза, оценки качества формируемых системами ИИ решений, подготовка кадров.

Сегодня словосочетание Искусственный Интеллект (ИИ) прочно вошло и закрепилось в самых различных составляющих нашей обыденной жизни. И дело не только в том, что на тему ИИ, как принято говорить, не высказывает тех или иных суждений, по-видимому, лишь только абсолютно ленивый. Бесспорный факт, что за последние примерно четверть века область внимания, а также все более смелых ожиданий, связанных с применениями активно развивающихся ИИ-технологий и решений, фантастически расширилась. Если 80 лет назад эта проблематика увлекала преимущественно лишь пионеров-исследователей да и, пожалуй, военных (для которых во все времена характерен интерес к перспективным технологиям и решениям), то сегодня уже могут потребоваться существенные усилия, чтобы указать такую область окружающей нас жизни, в которой на обсуждались бы перспективы применения тех или иных ИИ-решений.

Сопутствующими эффектами в такой ситуации стали *hype*<sup>2</sup> вокруг проблематики ИИ вместе с целым набором мифов и легенд, циркулирую-

---

<sup>1</sup> См. также [Забейайло, 2023].

<sup>2</sup> Бесконтрольный ажиотаж (англ.).

щих в обществе (в том числе – даже в профессиональной среде исследователей). Перспективы дальнейшего развития в этом направлении текущей ситуации не могут не вызывать у экспертов вполне обоснованных опасений. Один из очевидных факторов риска здесь – это рост «пузыря» завышенных ожиданий общества в части практических результатов применения ИИ-систем и решений, дополняемый перспективами последующего глубокого разочарования «широких масс интересантов», уставших дожидаться обещанных им фантастических прорывов и достижений.

Настораживающие профессионалов элементы реальности окружающей нас действительности – это заполнившие инфосферу всевозможные обсуждения фантастических преимуществ интеллектуализации всех сфер нашей жизни: использования интеллектуальных пылесосов или холодильников, интеллектуальных<sup>3</sup> домов и интеллектуальных производственных технологий, необходимости прямо завтра перейти на использование интеллектуальных систем управления, возникающие повсеместно кафедры и отделы интеллектуальных систем, а также многое другое. (Справедливости ради следует отметить, что в ответственных ситуациях – в частности, при использовании ИИ-технологий и решений в критически значимых приложениях – обычно используется более корректная терминология: например, упоминаются *системы с элементами ИИ* и т.п.).

К сожалению, используемая на текущий момент (причем даже в профессиональном сообществе ИИ-специалистов) терминология далеко не всегда предлагает однозначное (и операционально<sup>4</sup> корректное) толкование задействованных понятийных конструкций. Следствием этого оказываются не только возникновение и поразительная живучесть целого ряда мифов о некоторых аспектах проблематики ИИ<sup>5</sup>, но и ряд проблем, связанных с завышенными и, вообще говоря, мало оправданными ожиданиями от применения ИИ-систем в практически значимых приложениях.

Не менее запутанной оказалась ситуация с оценкой достаточности оснований для принятия результатов, формируемых системами ИИ. Ограниченная применимость традиционных средств обоснования результатов работы ИИ-систем – использования математического инструментария

---

<sup>3</sup> Справедливости ради следует заметить, что в актуально используемой лексике некоторых национальных языков части подобных смысловых неоднозначностей (а иногда – и очевидных коллизий) удастся избежать. Так, например, английский термин *smart* позволяет «смягчить» толкование «интеллектуальности» поведения некоторых технических устройств (называемых *smart devices*) в соотнесении с представлениями об интеллектуальности поведения человека в тех или иных требующих «использования интеллекта» ситуациях.

<sup>4</sup> Позволяющее аргументированно отделить удовлетворяющие ему объекты от неудовлетворяющих

<sup>5</sup> См., в частности, [Финн и др., 2023] и др.

дедуктивных доказательств или же средств стандартного статистического анализа данных при решении прикладных задач в открытых, постоянно пополняемых новой информацией предметных областях – стала драйвером развития новых подходов, методов и математических моделей. Актуализировалась проблематика обеспечения доверия и доверенности систем ИИ. Использование инструментов доказательства, понимаемого как приведение формируемого ИИ-системой заключения к неоспариваемости относительно заданного контекста (например, конкретного множества накапливаемых фактов эмпирического исследования и др.) расширяется за счет его интеграции со средствами формирования содержательной интерпретации получаемых результатов, а также их неформального объяснения и аргументации. Стандартом *de facto* стало использование так называемых контрафактуальных схем обоснования, учитывающих как сходства, так и различия в описаниях анализируемых эффектов. Бесспорный триумф использования технологий искусственных нейронных сетей (ИНС) и так называемых больших языковых моделей (БЯМ) при решении важных прикладных задач вывел в разряд критически значимых потребность в надежном и неоспариваемом обосновании формируемых ими в каждом конкретном случае результатов и заключений (потребность все-таки иметь возможности формировать приемлемый ответ на классический вопрос «*Что есть истина?*»).

Очевидным (причем теперь уже не только для разработчиков, но и для более широкого круга Лиц, Принимающих Решения (ЛПР), промышленных заказчиков, ответственных руководителей органов государственного управления, представителей структур-регуляторов и др.) барьером на пути разработки и развития ИИ-систем для различных промышленных приложений оказалась проблема подготовки квалифицированного заказчика – как собственно ЛПР, способного и готового принять на себя ответственность за принимаемые им решения и их последствия, так и обеспечивающих его компетентной поддержкой экспертов (в том числе – способностью обеспечить со стороны заказчика корректную постановку решаемой задачи для исполнителя, а также проконтролировать корректность представляемых исполнителем результатов решения). *Как, чему именно, где и какими силами* готовить таких специалистов – вот еще один комплекс вопросов, без эффективного решения которых отвечать на технологические вызовы развития ИИ-систем и решений буде крайне затруднительно.

Таким образом, представляется актуальным и полезным уточнить <sup>6</sup> понимание экспертами профессионального ИИ-сообщества (см. ранее, [Финн и др., 2023], [Забейайло, 2023] и др.) еще по трем позициям:

(i) Как понимать термин *интеллектуальность* систем ИИ (СИИ)?

---

<sup>6</sup> Например, через публикации, обсуждения на ИИ-конференциях и семинарах, ...

- (ii) Какими средствами следует воспользоваться для *оценки приемлемости* (надежности, обоснованности, ...) заключений и рекомендаций, формируемых конкретной СИИ?
- (iii) Что полезно предпринять для повышения уровня ИИ-подготовки ЛПР и экспертов, определяющих направления развития и приоритеты финансирования ИИ-исследований и разработок в нашей стране?

Вариант ответа на каждый из этих трех вопросов и будет предложен ниже.

### **1 Так в чем же, собственно, интеллект? (О критериях интеллектуальности систем ИИ)**

К классу систем ИИ (СИИ) представляется целесообразным отнести все такие компьютерные системы, которые позволяют решать те или иные задачи, относимые к ИИ как области исследований и разработок. При этом достаточно часто вместе с термином СИИ также используется термин *интеллектуальные системы* (ИС). Некоторые авторы рассматривают эти два термина – СИИ и ИС – как синонимы, что на наш взгляд представляет собою глубокое заблуждение. Если рассмотреть понимание этих двух терминов в самом общем виде (так сказать с высоты птичьего полета), то в случае СИИ мы имеем дело с эмуляцией (своего рода «протезированием») в том числе и сугубо вычислительными средствами отдельных функций интеллекта человека (естественного интеллекта – ЕИ), используемых им при решении «трудных» (как принято считать. требующих «интеллекта») задач. В таком случае, в чем же *интеллектуальность* СИИ, позволяющих аргументированно отнести их к классу ИС?

Представляется вполне естественным считать, что интеллектуальность СИИ определяется возможностями *воспроизведения (имитации)* в них (компьютерными средствами – !) тех или иных *функций (способностей)* ЕИ. При таком подходе представления об ИИ – в т.ч., определение ИИ как исследования и разработки, которые направлены на имитацию и усиление функций ЕИ компьютерными средствами, – открывает возможности для целенаправленного и контролируемого<sup>7</sup> построения систем ИИ. Однако, при этом потребуются уточнить наши представления об ЕИ (в том числе – перечислить или же привести примеры способностей ЕИ, характеризующих собственно понятие *интеллект*). Не менее значимо и разделение таких способностей на *допускающие* и *не допускающие* (по крайней мере на текущем уровне развития) моделирование компьютерными средствами.

---

<sup>7</sup> Позволяющего обходить различные мифы и угрозы типа, например, «восстания интеллектуальных пылесосов» или же всеобщей безработицы, которую сулит широкое внедрение ИИ-систем и решений.

Двигаясь в обозначенном направлении – уточняя наши представления о характеристиках ЕИ, воспользуемся хорошо известным подходом<sup>8</sup> феноменологического характера. Еще в 20-е годы прошлого века психологи, изучая структуру ЕИ, предложили определение «интеллект – это то, что измеряется тестами интеллекта» (см., например, работу Э.Боринга [Boring, 1923], увидевшую свет в 1923 году, и др.). Процедурный каркас феноменологического подхода Э. Боринга – это психологические тесты и обработка их результатов математическими методами статистики (приоритет – метод главных компонент [Pearson, 1901], использованный для выявления факторов наибольшего влияния) позволил сформировать «эко-систему» из 10 так называемых «широких» способностей (таких, например, как «*работа с памятью* – запоминание и «извлечение» информации из памяти, «*способности к восприятию* – визуальная и слуховая обработка информации, «*чтение и письмо*, «*способности оперировать количественными величинами*, «*работа со знаниями*, а также способность выстраивать гибкие «*последовательности логичных «шагов» мысли* и др.), дополненных сопутствующими им «узкими» способностями, каждая из которых детализирует соответствующие «широкие» способности. Современная теория структуры человеческого интеллекта, предложенная Кэттеллом, Хорном и Кэрроллом [Schneider, et al 2018], стала естественным развитием и обобщением исторически сложившегося (см. [Boring, 1923] и др.) подхода. При этом следует особо отметить, что центральная роль в теории Боринга-Кэттелла-Хорна-Кэрролла отводится умению человека *рассуждать и формировать понятия*, решая *новые задачи в незнакомых ситуациях*.

Возвращаясь к данному выше определению ИИ как имитации и усиления компьютерными средствами системообразующих способностей ЕИ, обратим внимание на разделение формализуемых и не формализуемых на текущий момент свойств и способностей ЕИ. В работах [Финн, 2023] и др. представлен перечень из 13 способностей ЕИ, эффективно переносимых сегодня на компьютер. Вот краткий вариант их перечня:

- (1) обнаружение существенного в данных;
- (2) порождение последовательности «цель – план – действие»;
- (3) подбор посылок, релевантных цели рассуждения;
- (4) способность к рассуждению: вывод следствия из посылок;
- (5) синтез и взаимодействие познавательных процедур (например, индукции, аналогии и абдукции с последующим применением дедукции);

---

<sup>8</sup> См. предложенное Д.С.Миллем [Милль, 2007] определение понятия *theoretical economy* как область исследований, которыми занимаются *theoretical economists*, а также более поздние «реинкарнации» этого методологического приема (физика как то, чем занимаются физики, и т.д.).

- (6) рефлексия – оценка знаний и действий (как рациональная и аргументированная реакция на состояние знаний и результаты действий)
- (7) способность к объяснению – ответ на вопрос «Почему?»;
- (8) аргументация при принятии решений;
- (9) познавательное любопытство и способность к распознаванию;
- (10) способность к обучению и использование памяти;
- (11) способность к интеграции знаний для образования концепций и теорий;
- (12) способность к уточнению неясных идей – преобразованию их в понятия;
- (13) способность к изменению системы знаний при получении новых знаний и изменений познавательных ситуаций.

Параллельно следует принять во внимание, что (по крайней мере, на текущий момент) за рамками области компьютерной формализации способностей ЕИ оказываются не только ряд таких критически значимых характеристик интеллекта как, например, интуиция и воображение, но и совокупность функций мозга, характеризующих высшую нервную деятельность человека.

Следуя далее уже намеченным путем и принимая данное выше определение ИИ (как области исследований и разработок, ориентированных на имитацию и усиление познавательных функций человека компьютерными средствами), можно предложить операциональный критерий интеллектуальности систем ИИ:

*отнесение СИИ к классу ИС обеспечивает реализация в оцениваемой СИИ рассматриваемых – см. выше и [Финн, 2023] – способностей ЕИ (в том числе – приближенная имитация тех или иных из перечисленных ЕИ-способностей компьютерными средствами).*

Упрощенная (и как следствие – «радикализованная») версия Критерия: *СИИ можно считать интеллектуальной, если она способна реализовать тот или иной вариант рассуждений (т.е. ИС – это «рассуждающая» СИИ).*

Разумеется, авторы допускают существование и других – отличных от предложенного ими выше – определений ИИ и, как следствие, других вариантов Критерия интеллектуальности СИИ. При этом предполагается, что каждая такая версия Критерия (прежде чем стать «инструментом» поддержки принятия соответствующих оценочных решений) должна пройти публичное критическое обсуждение в профессиональном сообществе специалистов, занятых практическими исследованиями и разработками в области ИИ.

## **2. «Чем же сердце, наконец, успокоится?» (Об оценке достаточности оснований для принятия результатов, формируемых системой ИИ)**

Необходимость применять СИИ в анализе открытых – постоянно пополняемых новыми данными (т.е. так называемый эффект *Open*, характерный для *Big Data*) – предметных областей повлекла существенное расширение соответствующего «инструментария» – подходов, математических моделей, методов и алгоритмов оценки достаточности оснований для принятия результатов, формируемых СИИ в той или иной конкретной ситуации. Осознано отсутствие возможностей использовать в условиях эффекта *Open* целый ряд традиционных средства компьютерного анализа данных. Так, для применения доказательств дедуктивного типа открытый характер анализируемых данных порождает в общем случае непреодолимые препятствия на пути формирования аксиоматического описания предметной области, а уж о доказательстве утверждений типа теорем о полноте в таких ситуациях вообще нет оснований говорить. Другой пример – использование «инструментов» статистического анализа данных натывается в условиях эффекта *Open* на отсутствие возможностей формировать генеральные совокупности, репрезентативные выборки из которых могли бы стать надежным основанием для обучения СИИ.

Как следствие, внимание исследователей и разработчиков развернулось в сторону подходов, ориентированных на содержательную интерпретацию, неформальное объяснение и аргументацию результатов, формируемых системами ИИ. Отсечение не интерпретируемых результатов стало эффективным инструментом борьбы с артефактами переобучения. Объяснение (как ответ на вопрос «Почему»?) позволяет использовать изначально скрытые в анализируемых данных эмпирические зависимости причинно-следственного типа как базу для обоснования приемлемости формируемых системой ИИ выводов и заключений. Все большую популярность завоевывают решения на основе так называемого контрафактуального подхода [Pearl, 1999], [Pearl, 2005], [Höfler, 2005 и др.], позволяющие проверять порождаемые рекомендации СИИ на предмет их фальсифицируемости. Активно развивается проблематика – подходы, модели и алгоритмы – оценки доверия (в значении *believe*) и доверенности (в значении *trust*) результатов, формируемых той или иной системой ИИ. В основе этих решений – те или иных механизмы сравнения сходства и различия в описаниях анализируемых (проверяемых на фальсифицируемость) свойств и эффектов.

К сожалению, по-прежнему открытым остается в общем случае вопрос о достаточности оснований для принятия результатов, порождаемых с использованием ИНС или Больших Языковых Моделей (БЯМ). В случае

БЯМ отсутствие надежных средств формирования ответа на классический вопрос «Что есть истина?», адресуемый к предлагаемым ими результатам, породил в последние несколько лет целую индустрию эвристических сервисов, которые призваны помочь пользователям в использовании БЯМ (см., например, *prompt engineering*, *retrieval augmented generation* и др.).

Все более широко востребованными при оценке достаточности оснований для принятия результатов, генерируемых системами ИИ, становятся аргументационные подходы. Среди направлений, тесно связанных с успехами в создании объяснимых (приемлемых) решений (а, следовательно, заслуживающих доверия систем), заметное место заняли формальные теории аргументации [Sugas et al., 2021]. Аргументация как таковая служит двум целям: *обоснование* и *убеждение*. В первом случае предъявление посылок вывода фиксирует состоятельность позиции. В этом смысле дедуктивное доказательство, *гарантирующее* истинность заключения, может также рассматриваться как частный случай аргументационной схемы. Однако, как уже говорилось, в открытом мире истинность заключения может ставиться под сомнение. Здесь можно рассчитывать лишь на ту или иную степень его вероятности или правдоподобия и, соответственно, приемлемости. Значительный вклад в развитие и применение аргументации в ИИ и информатике вносят так называемые «системы абстрактной аргументации» [Dung, 1995]. Ключевым понятием в этой структуре является атака аргументов – абстрактное формальное отношение, представляющее возможность опровержения одного аргумента другим. В дальнейших разработках системы абстрактной аргументации были существенно расширены: системы биполярной аргументации используют отношения атаки и поддержки, количественной – с учётом «веса» аргументов [Baroni et al., 2019], обобщённой аргументации [Gabbay, 2016] – любое количество отношений. Системы нечёткой аргументации позволяют представить относительную силу взаимосвязи между аргументами и степень их принятия [Janssen et al., 2008]. Генеративные модели с использованием байесовской аргументации (в рамках байесовского подхода к рассуждениям) оказываются действенными для достижения устойчивости к состязательным атакам [Cerutti, 2022]. Тесная связь логики и аргументации, известная ещё со времён Аристотеля, поддерживается последними исследованиями в области представления знаний и рассуждения в ИИ (см., например, [Багин, 2019], [Besnard et al., 2020], [Dastani et al., 2020]). В [Финн, 2021a] предложены оригинальные отечественные варианты четырёхзначных логик с аргументационной семантикой и неассоциативными связками, в [Финн, 2021b] развита схема многоуровневой аргументации.

Аргументация как когнитивный феномен является предметом изучения различных дисциплин – философии, лингвистики, юриспруденции, политологии, когнитивных наук и т.д. Это расширяет горизонты фор-



мальных подходов к принятию решений с помощью аргументации, обоснованию на основе аргументов с учётом базовых (внешних) знаний, формированию иерархических систем аргументации (например, на основе доверия источникам или неких количественных оценок), аргументационного диалога *«pro»* и *«contra»* и т.д. На базе аргументации создаются рекомендательные системы, классификаторы, планировщики. В последнее время наблюдается также усиление таких систем моделями ML для формирования аргументации в рамках обучения на положительных и отрицательных примерах или разнонаправленных вознаграждений/штрафов в обучении с подкреплением.

Таким образом, выбор средств, которыми следует воспользоваться для *оценки приемлемости* (надежности, обоснованности, ...) заключений и рекомендаций, формируемых конкретной СИИ, – еще один актуальный вопрос для обсуждения в профессиональном ИИ-сообществе.

### 3. «А судьи – кто?» (Об экспертизе и кадрах)

Опыт последних примерно 5 лет в части организации и проведения в нашей стране крупных конкурсов по проблематике ИИ высветил критически значимую проблему. Речь о несоответствии уровня квалификации (в обсуждаемой здесь области) участвовавших в проведении этих конкурсов ответственных руководителей и привлекаемых ими экспертов целям и задачам развития данного направления науки и технологий в РФ. Собственно, и ранее, например, в период существования РФФИ, нередко решения о финансировании или же, наоборот, не финансировании тех или иных групп исследователей принимались не на основе беспристрастной оценки качества и перспектив предлагаемых ими проектов, а по принципу «близости» к соответствующим ЛПР или же с учетом «конкуренции» определенных исследовательских сообществ (а также их лидеров). При этом, как это было, например, в случае конкурсов РФФИ второй половины 2010-х годов, «квантование» выделяемых в таком образом организованных процедурах в виде грантовой поддержки сумм в диапазоне 3-5 миллионов рублей не оказывало существенного влияния на состояние ИИ-исследований и разработок в целом по стране. Однако, при выходе на проекты с общим финансированием порядка 1 млрд. рублей (см., в частности, Конкурс на организацию так называемой первой волны Центров ИИ, запущенный постановлением Правительства РФ летом 2021 года) влияние уровня экспертизы на принимаемые решения приобрело критически значимый характер. Если, например, взглянуть на имеющую место ситуацию, рассуждая по аналогии, то любому непредвзятому наблюдателю было бы странно наблюдать за экспертизой и принятием решения в ситуации, когда ведущим в экспертизе плана проведения нейрохирургической операции на головном мозге человека был бы избран, например, хирург-онколог, который является бесспорным специалистом в области

хирургического лечения рака простаты. Мотивация ЛПР в крупных конкурсах обсуждаемого типа, сводимая к фигуре рассуждений вида *«в ведущих промышленно-развитых странах мира ИИ-методы и технологии класса NNN активно развиваются, следовательно, и мы в нашей стране должны будем пойти именно этим путем»*<sup>9</sup>, разумеется, не лишена здравого смысла. Однако, она не учитывает ряд критически значимых для нашей страны факторов – наличия своих научных и инженерных школ с длительной успешной историей развития и собственными достижениями, асимметрии технической вооруженности отечественных исследователей в сравнении с основными зарубежными конкурентами<sup>10</sup> и др.

В чем причины такой ситуации? Одним из приоритетных факторов, на наш взгляд, является проблема недостаточного уровня подготовки, а более широко – формирования квалифицированного заказчика и консультирующих его экспертов. Где взять квалифицированных экспертов? Как, кому, по каким направлениям и где готовить кадры для ИИ как области исследований и разработки? Таким образом, хорошо известный вопрос «А судьи – кто?» легко трансформируется в рамках национальных усилий по развитию ИИ-решений и систем в достаточно жесткий вопрос об экспертизе и кадрах.

Опыт некоторых крупных российских ИТ-проектов последних пяти лет (см. например, формирование Центров НТИ, Лидирующих Исследовательских Центров по заданным направлениям и тематикам, 6 центров ИИ так называемых первой «волны», а также следовавшие за эти конкурсы «второй» и «третьей» волн и др.) продемонстрировал, что:

- реальной альтернативой формально открытым конкурсам в условиях небесспорной экспертизы (а также ориентации ЛПР такого конкурса в первую очередь на «статус» организаций победителей, а уж потом на содержание предлагаемых работающими в них исследовательскими группами проектных предложений) может быть, например, прямое включение в государственное задание «достойным» организациям соответствующих работ (обеспеченных соответствующим финансированием). При таком – альтернативном конкурсному – подходе можно избежать непродуктивного расхода сил и времени соискателей на подготовку объемных конкурсных заявок<sup>11</sup>, а также дать возможности получателям соответствующего государственного заказа привлекать (разумеется, под свою ответственность) полезных им высоко квалифицированных соисполнителей по конкретным направлениям работ;

---

<sup>9</sup> Грубо говоря, схема принятия решения выглядит примерно следующим образом: *«Слышали про ИНС. Это – наше 'все'. Будем развивать это направление»*

<sup>10</sup> См., например, уровень обеспеченности вычислительными ресурсами отечественных разработчиков ИНС-решений в сравнении с конкурентами в США и КНР.

<sup>11</sup> Часто используемых лишь при аргументации отказа соискателю

• критически значимым для последующего выполнения проектных работ и последующих процедур сдачи-приемки получаемых результатов оказывается организационно-контентный процесс, позволяющий ответственно сформировать задачи для потенциальных исполнителей. Очевидные критичные аспекты здесь – это, в частности:

- кто именно (какие требования к квалификации, в каких именно конкретных направлениях, ...) и
  - как именно сформирует действительно необходимые государствену цели и задачи (деньги ведь – бюджетные, и отчетность за их использование предусматривает, в том числе, и достаточно жесткие меры государственного контроля и ответственности);
  - кто именно (квалификация, в каких направлениях науки и технологий) и по каким критериям будет проводить экспертизу конкурсных заявок?
- вопрос о том, как (по каким критериям) принимать решение о выборе победителей (дать «своим»; выбрать необходимые для индустрии решения, отобрать те работы, которые способны обеспечить возврат инвестиций, ...) на национальной уровне проектов, оказывается критически значимым;
- «тонкой» настройки и аккуратной организации требует мониторинг исполнения проектов: важно располагать «инструментами» управления, позволяющими не потерять вложенные инвестиции, и при этом иметь эффективные возможности обходить возникающие по ходу проекта трудности и проблемы;
- в каждом конкретном проекте следует принимать особое – учитывающее специфику данного проекта – решение о том, как организовать процедуры сдачи-приемки (в т.ч. – научную, техническую и финансовую экспертизу), чтобы обеспечить успешное завершение проекта.

Итак, отдельного обсуждения в профессиональном сообществе ИИ-исследователей и разработчиков заслуживают, на наш взгляд, три проблемы:

- (1) подготовки квалифицированного заказчика (включая формирование среды, которая способна обеспечить заказчика-ЛПР профессиональной экспертизой в требующемся ему круге вопросов или предметной области);
- (2) формирования профессиональной экспертизы (чтобы иметь неоспариваемый ответ на вопрос «А судьи – кто?»);
- (3) систематической подготовки квалифицированных исполнителей (чтобы иметь неоспариваемое представление о том, кто именно и чему именно учит, ориентируясь, как минимум на 5 профилей подготовки специалистов по ИИ: «теоретики», разработчики программных систем ИИ, квалифицированные пользователи ИИ-систем, специали-

сты по управлению проектами в области разработки и внедрения ИИ-систем и решений, а также – специальные курсы повышения квалификации в области ИИ для специалистов различного профиля – от государственных чиновников до преподавателей учебных заведений, обучающихся проблематике ИИ).

## Заключение

Сегодня ИИ как область исследований и разработок по-прежнему находится в стадии становления, характеризуемой еще только формируемой собственной проблемно-ориентированной терминологией, а также специфической системой понятий и методов исследования. Именно по этой причине представляется важным определиться в нашей профессиональной среде ИИ-специалистов с актуальными именно для нас приоритетами понимания перспектив и выбора путей дальнейшего развития. Было бы полезно сформировать общее видение того, на чем именно сфокусироваться, чтобы не отстать «навсегда», наслаждаясь своими былыми достижениями и, не жалея сил и времени, поддерживая готовность «к лихой кавалерийской атаке на танки противника<sup>12</sup>».

## Список литературы

- [Вагин, 2019] Вагин В.Н. Элементы теории аргументации и её роль в интеллектуальном анализе данных // В кн.: Вагин В.Н. Знания и убеждения в интеллектуальном анализе данных. – М.: Физматлит, 2019. – 536 с.
- [Забейайло, 2023] Забейайло М.И. Три вопроса (на понимание), адресованные «товарищам по партии» // КИИ-2022 (21–23 декабря 2022 г.). Труды конф. / под ред. В.В. Борисова, Б.А. Кобринского. – М.: РАИИ, 2022. – Т. 1. – С. 280-289.
- [Милль, 2007] Милль Д.С. Об определении предмета политической экономии и о методе исследования, свойственном ей // В книге: Основы полит. экономии с нек. приложениями к социальной философии. – М.: Эксмо, 2007. – С. 985-1023.
- [ПК] Польские кавалерийские бригады в 1939 году. – <https://chestnut-ah.livejournal.com/842458.html>.
- [Финн, 2021a] Финн В.К. Стандартные и нестандартные логики аргументации // В.К. Финн. Искусственный интеллект: методология, применение, философия. – М.: ЛЕНАНД, 2021. – 468 с. – С. 337-363.

---

<sup>12</sup> Поучительным примером здесь может быть, в частности, ситуация с различными видами кавалерии в качестве основной ударной силы польской армии второй половины 30-х годов. По данным [ПК] в сентябре 1939 года относительная доля кавалерийских частей в польской армии была в 5 раз выше, чем в германской: примерно 10,5% к 2,1%. Такая разница была следствием как особенностей актуальной на тот момент военной доктрины Польши, так и старых национальных кавалерийских традиций.

- [Финн, 20216] Финн В.К. Двенадцать тезисов об аргументационных системах // Финн В.К. Интеллект, информационное общество, гуманитарное знание и образование. – М.: ЛЕНАНД, 2021. – 464 с. – С. 191-221.
- [Финн, 2023] Финн В.К. Искусственный интеллект. Методология, применения, философия. – 2-е изд. испр. и доп. – М.: ЛЕНАНД, 2023. – 464 с.
- [Финн и др., 2023] Забейайло М.И., Михеенкова М.А., Финн В.К. О некоторых актуальных мифах современного искусственного интеллекта // Труды XXI Нац. по ИИ с междунар., Смоленск 16-20 октября 2023 г. – Т. 1. – С.190-200.
- [Baroni et al., 2019] Baroni P., Rago A., and Toni F. From fine grained properties to broad principles for gradual argumentation: A principled spectrum // *Int. J. Approx. Reason.* – 2019. – 105. – P. 252-286.
- [Besnard et al., 2020] Besnard P., Cayrol C., Lagasquie-Schiex M.-C. Logical theories and abstract argumentation: A survey of existing works // *Argument & Computation.* – 2020. – Vol. 11(1-2). – P. 41-102.
- [Boring, 1923] Edwin G. Boring. Intelligence as the Tests Test It // *New Republic.* – 1923. – 36. – P. 35-37.
- [Cerutti, 2022] Cerutti F. Supporting Trustworthy Artificial Intelligence via Bayesian Argumentation // *Lecture Notes in Computer Science.* – 2022. – Vol. 13196. – P. 377-388. – [https://doi.org/10.1007/978-3-031-08421-8\\_26](https://doi.org/10.1007/978-3-031-08421-8_26).
- [Čyras et al., 2021] Čyras K., Rago A., Albini E., Baroni P., Toni F. Argumentative XAI: A Survey // *Proc. of the Thirt. Int. Joint Conf. on AI (IJCAI-21).* – P. 4392-4399.
- [Dastani et al., 2020] Dastani M., Dong H., van der Torre L. (eds). Logic and Argumentation // *Lecture Notes in Artificial Intelligence.* – 2020. – Vol. 12061.
- [Dung, 1995] Dung P.M. On the Acceptability of Arguments and its Fundamental Role in Non-Monotonic Reasoning, Logic Programming, and  $n$ -person Games // *Artificial Intelligence.* – 1995. – Vol. 77. – P. 321-357.
- [Gabbay, 2016] Gabbay D.M. Logical foundations for bipolar and tripolar argumentation networks: preliminary results // *J. Log. Comput.* – 2016. – 26(1). – P. 247-292.
- [Höfler, 2005] Höfler M. Causal inference based on counterfactuals // *BMC Med. Res. Methodol.* – 2005. – 5, 28. – <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-5-28>.
- [Janssen et al., 2008] Janssen J., De Cock M., Vermeir D. (2008) Fuzzy Argumentation Frameworks // *Proceedings of the 12<sup>th</sup> Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems.* – P. 513-520.
- [Pearl, 1999] Pearl J. Causality: models, reasoning, and inference. – N.-Y.: Cambridge Univ. Press, 2000. – 384 p.
- [Pearl, 2005] Pearl J. Probabilities of causation: three counterfactual interpretations and their identifications // *Synthese.* – 1999. – 121. – P. 93-149.
- [Pearson, 1901] Pearson K. On lines and planes of closest fit to systems of points in space // *Philos. Mag.* – 1901. – 2(11). – P. 559-572.
- [Schneider et al 2018] Schneider W.J. and McGrew K.S. The Cattell-Horn-Carroll Theory of Cognitive Abilities // in *Contempor. Intellect. Assessm.: Theories, Tests, and Issues.* – 4th ed., ed. D.P. Flanagan & E.M. McDonough (N-Y: Guilford, 2018). – P. 73-163.