

УДК 165+004
DOI 10.15372/PS20230409
EDN REQDOM

А.С. Зайкова

НОЯБРЬ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Термин «зима искусственного интеллекта» был предложен по аналогии с термином «ядерная зима» для описания цепной реакции, вызванной пессимистическими прогнозами экспертов и приводящей к резкому сокращению исследований в сфере ИИ. Впоследствии термин «зима ИИ» под влиянием понятия «цикл ажиотажа» стал характеризовать инвестиционный цикл ИИ-исследований и только упрочился с появлением термина «весна ИИ» как синонима «бума ИИ». Последние несколько лет отдельные исследователи предрекают зиму ИИ, предупреждая о возможных проблемах, следующих за «пиком завышенных ожиданий». Однако более тщательный анализ показывает, что при предсказании зимы ИИ не учтено несколько факторов, ключевыми из которых являются институциональные и экономические, а также упущен из виду принципиально иной подход современного общества к ИИ-исследованиям. Это значит, что даже с наступлением «ноября ИИ» в обозримом будущем зима фундаментальных ИИ-исследований так и не наступит.

Ключевые слова: наука; искусственный интеллект; философия ИИ; зима ИИ; весна ИИ; нейросети; генеративные модели; отчет Лайтхилла; отчет ALPAC.

A.S. Zaykova

NOVEMBER OF ARTIFICIAL INTELLIGENCE

The term “artificial intelligence winter” was coined by analogy with the term “nuclear winter” to describe the chain reaction caused by experts’ pessimistic forecasts and resulting in a drastic decline in AI research. Later on, the term “AI winter,” influenced by the concept of “hype cycle,” began to describe the investment cycle of AI research and only became stronger with coining the term “AI spring” as a synonym for “AI boom.” Over the past few years, several researchers have been predicting an AI winter and warning of potential problems following the “peak of inflated expectations.” However, a more thorough analysis shows that predicting an AI winter does not take into account some factors, the key ones being institutional and economic, as well as ignores the fundamentally different approach of modern society to AI research. This means that even with the advent of “AI November,” the winter of basic AI research will not come in the near future.

Keywords: science; artificial intelligence; philosophy of AI; AI winter; AI spring; neural networks; generative models; Lighthill report; ALPAC report

В 1973 году парламент Великобритании попросил Дж. Лайтхилла оценить состояние исследований в области искусственного интеллекта. Изначально речь шла только об обзоре исследований в самой Великобритании, однако выводы Лайтхилла относились к состоянию дел в целом. Этот отчет, описывающий множество неоправданных надежд и завышенных ожиданий, в результате привел к сокращению финансирования работ в области искусственного интеллекта [1, р. 117; 7, р. 22], что Р. Шанк и М. Мински в 1984 г. ретроспективно назвали «зимой ИИ» – цепной реакцией, порожденной пессимизмом экспертов в отношении ИИ-исследований.

Последние несколько лет или даже десятков лет отдельные эксперты также указывают на близость зимы ИИ. Сначала это были осторожные высказывания о том, что глубокое обучение исчерпало весь свой потенциал, а после появления генеративных моделей компании OpenAI – предупреждения об их несовершенстве и завышенных ожиданиях от нейросетей. Однако вместо наступления зимы началась новая «весна ИИ» – «бум ИИ», порожденный разработками компании OpenAI и значительно повлиявший на экономическую систему в целом и на научные исследования в частности.

В этой статье мы проанализируем причины, которые традиционно считаются предвестниками зимы ИИ, и покажем, что несмотря на отдельные признаки скорой зимы ИИ, для фундаментальных научных исследований в области ИИ в обозримом будущем она так и не наступит.

Отчеты Лайтхилла и ALPAC: первые зимы ИИ

Согласно позиции Р. Шанка и М. Мински, зима ИИ представляет собой цепную реакцию, которую вызывает пессимизм ИИ-экспертов. Мнения экспертов влияют на прессу, множась этот пессимизм, который приводит к значительному сокращению финансирования и, как следствие, к прекращению дорогостоящих или долгих исследований. В 1973 г. ряд экспертов уже высказывали осторожные прогнозы о невозможности достичь поставленных перед данной областью целей в обозримые сроки. Именно эти осторожные предположения и повлияли на заказ Дж. Лайтхиллу отчета о состоянии дел в сфере разработок ИИ [3], который оказался смелым, громким и в полной мере радикальным.

В своем отчете для построения более объективной картины Лайтхилл разделил все исследования в области ИИ на категорию А, категорию С и категорию В – в зависимости от цели разработок.

Категория А, или *Advanced Automation*, – исследования, направленные на осуществление возможности замены людей машинами при выполнении конкретных задач: научных, промышленных или военных. Это, в частности, исследования в области распознавания образов, символов и речи, а также синтезирование речи и машинного перевода. К этой категории относятся также исследования, затрагивающие и любые механические действия с необходимостью интеллектуального сопровождения, такие как исследования во враждебной среде или упаковка контейнеров. К этой же категории Лайтхилл отнес и исключительно интеллектуальную деятельность, такую как хранение и поиск информации, обобщение материала и доказательство теорем, решение проблем и принятие решений.

Категория С – компьютерное исследование центральной нервной системы. Эти исследования были посвящены использованию компьютеров для построения и оценки теорий сознания: нейрофизиологических, психологических, феноменальных. К этой области Лайтхилл относил также исследование отдельных функций ЦНС и попытку познания самого человеческого интеллекта.

Наконец, категорию В (*Bridge*) Лайтхилл представил как промежуточную между категориями А и С. Согласно его позиции, основной целью исследований в этой области было создание роботов, способных к имитации некоторых конкретных человеческих функций, таких как зрение, распознавание машинного текста или принятие решений.

Лайтхилл высказался очень резко и однозначно, демонстрируя проблемы во всех перечисленных областях, однако некоторые конкретные области исследований категории А он выделил особо: «...Самые печально известные разочарования оказались в области машинного перевода, где огромные суммы были потрачены с очень небольшим полезным результатом, как показал тщательный анализ, проведенный в 1966 г. Национальной академией наук США. Этот вывод не поколеблен никакими последующими событиями» [3]. Действительно, отчет ALPAC (*Automatic Language Processing Advisory Committee* – Консультативный комитет по автоматической обработке языков), в 1966 г. продемонстрировал, что текущие достижения в области кибернетики, психологии и лингвистики не позво-

лили за десять лет приблизиться к идеалу машинного перевода, что привело не только к снижению финансирования исследования в сфере машинного перевода, но и к появлению некоторого «стыда» и закрытости у тех исследователей, которые продолжали им заниматься [6]. Некоторые авторы также используют термин «зима ИИ» для обозначения последовавшего за отчетом ALPAC затишья в сфере машинного перевода. Впрочем, такое применение этого теормина является исключительно редким, поскольку упомянутый отчет касался только одной из сфер ИИ-исследований, в отличие от отчета Лайтхилла.

Но вернемся к последнему. Лайтхилл использует пример с машинным переводом, для того чтобы продемонстрировать упадок всей области А. И хотя он указывает на проблемы в областях В и С, эти области кажутся сложнее для освоения и для извлечения экономической прибыли, в результате чего неоправданных ожиданий в них было значительно меньше, и, как следствие, значительно меньшим было и сокращение финансирования соответствующих исследований [3].

Если же мы будем рассматривать отчет ALPAC в отношении машинного перевода и отчет Лайтхилла в отношении А-категории исследований в сфере ИИ, то стоит заметить, что оба этих отчета спровоцировали переход от осторожных, но пессимистических «ноябрьских» прогнозов экспертов, предлагаемых инвесторам и обществу почти в частном порядке, к суровым зимам, лишившим исследователей финансирования, а общество – результатов исследований.

Вторая зима ИИ

Как мы уже упоминали выше, термин «зима ИИ» был введен в 1984 г. Р. Шанком и М. Мински. Исследователи обращали внимание, что экспертные системы почти полностью порождены методами программирования двадцатилетней давности. Новейшие разработанные программы были не умнее предыдущих, но они могли быстрее принимать глупые решения. Шанк и Мински использовали аллюзию на мрачный сценарий ядерной зимы, предвещая печальное будущее для исследователей ИИ: «...Оптимизм покидает сначала исследовательское сообщество, потом и общественное мнение, ведущие деятели в сфере ИИ подвергаются насмешкам. Финансирование исследований в области искусственного интеллекта прекраща-

ется, а ветераны искусства обработки списков оказываются в холоде и темноте, работая таксистами» [1, р. 203]. Шанк и Мински предрекали крах текущей системы финансирования, и он действительно случился совсем скоро. Следующую после зимы ИИ 1973–1980 гг. тяжелую пору исследователи связывают с упадком экспертных систем и датируют периодом 1987–1993 гг., хотя иногда упоминают другие даты: 1984–1993 гг. (начиная отсчет с конференции AAAI – Американской ассоциации искусственного интеллекта, где Шанк и Мински выступили со своим предупреждением) или 1987–2000-е гг. (связывая начало периода с упадком экспертных систем и его завершение – с появлением «глубокого обучения»).

Тем не менее если исходить исключительно из объема финансирования, то период 1987–1993 гг. больше всего подходит для определения зимы ИИ. Однако если зима машинного перевода была связана с разочарованием конкретно в области машинного перевода и долгим отсутствием успехов, а первая зима ИИ – с общим несоответствием ожиданий и сложностью коммерческого использования разработок, то к 1987 г. экспертные системы в виде специализированного оборудования ИИ на основе LISP, несмотря на общее устаревание технологии, применялись повсеместно и приносили коммерческую прибыль. Тем не менее появление конкурентов LISP-машин в виде настольных компьютеров от Apple и IBM привело к краху всего рынка LISP-машин. Именно с этим крахом и связывают падение финансирования, однако, на наш взгляд, наиболее критичным стало сокращение инициативы стратегических вычислений. В 1983 г. DAPRA (Управление перспективных исследовательских проектов Министерства обороны США) направило большой объем инвестиций на исследования для достижения ряда практических целей и потенциального создания полноценного ИИ. К 1985 г. на базе 60 учреждений реализовывалось 92 проекта с общим инвестированием в размере более чем 100 млн долл. с начала действия программы. Однако длительность программы оказалась разочарывающей: уже в 1987 г. Дж. Шварц, возглавивший IPTO (Управление технологий обработки информации), назвал экспертные системы «умным программированием» и урезал финансирование ИИ, тем самым снова запустив очередную зиму в сфере стратегических исследований [4].

Жизненный цикл разработки ИИ

Можно назвать целый ряд причин, лежащих в основе «похолоданий» в любой сфере исследований: ажиотаж в СМИ, институциональные и экономические факторы, недостаток фундаментальных исследований, неспособность существующих технологий к адаптации, а также технические недостатки оборудования. Но если последние два аспекта присутствуют в любых разработках довольно часто (если не сказать всегда), то отдельные институциональные и экономические факторы, в том числе приводящие к недостатку фундаментальных исследований, можно выявить отдельно. Особенную же роль в жизненном цикле любой технологии играет первый фактор – ажиотаж в СМИ.

Для любой технологии существует так называемый «цикл ажиотажа»: технологический прорыв влечет за собой бурную реакцию в СМИ, которая подпитывает саму себя, что в результате конструирует «пик завышенных ожиданий», за ним следует «долина разочарования», которая в одних случаях может привести к краху всего направления исследований, а в других сменяется процессом повсеместного внедрения технологии в производство. Шанк и Мински не думали об этом цикле. Они использовали термин «зима ИИ» по аналогии с популярным тогда термином «ядерная зима», таким образом пытаясь указать на необратимую цепную реакцию, вызванную пессимизмом экспертов в сфере ИИ [1]. Однако термин «зима ИИ» как аналогия термина «ядерная зима» смешался с понятием «цикл ажиотажа» и в результате прижился как обозначение этапа жизненного цикла финансирования разработок. Появление термина «весна ИИ» как синонима термина «бум ИИ» [8] только упрочило смысл понятия зимы в исследованиях искусственного интеллекта как чего-то временного и естественного для жизненного цикла любых разработок.

Действительно, различные этапы исследований укладываются в этот цикл. Все начинается с «бума», или «весны»: появляются новые исследования, запускающие бурную реакцию ажиотажа. Для «лета» характерны институциональные изменения сферы исследований и рынка труда, а также серьезные экономические перестроения, как это было в 1983–1986 гг. в фундаментальных исследованиях и на рынке LISP-машин и как это происходит сейчас. «Осень», в свою очередь, предполагает потенциальное разочарование экспер-

тов: от осторожных «сентябрьских» прогнозов до повсеместного «ноябрьского» уныния из-за расхождения ожидаемых и фактических результатов. И наконец, для того чтобы перейти от ноября ИИ к зиме ИИ, необходим триггер – отчет или решение эксперта, который и приводит к фактическому снижению финансирования.

Стоит отметить, что такой триггер для перехода «ноября» в «зиму» необходим не только для стратегических исследований и концептуальных научных разработок, но и при создании и разработке более или менее конкретных технологий, особенно в категориях А и В по терминологии Лайтхилла: от создания виртуальных медицинских помощников до всего направления коннекционизма (1969 г., книга «Перцептроны» [5]) или машинного перевода (1966 г., отчет ALPAC [6]).

Зима ИИ близко?

Хотя в результате появления генеративных ИИ-систем авторства OpenAI можно говорить о буме ИИ, или весне ИИ, отдельные руководители, разработчики и инвесторы, как и в последние несколько десятков лет, указывают на близость зимы ИИ. Предыдущий всплеск осторожных, но удручающих прогнозов возник после того, как «глубокое обучение» стало привычным и развитие нейронных сетей (в частности, чат-ботов) замедлилось. Но за последние несколько лет произошли значительные изменения всей нашей жизни, связанные с появлением генеративной языковой модели ChatGPT 3.0 и нейронных сетей DALL-E и Midjourney. Нейросети могут создавать изображения, писать стихи, переводить общие и специальные тексты, искать информацию и делать выжимку из больших объемов литературы, т.е. все то, что Лайтхиллу еще пятьдесят лет назад казалось невозможным. Изменился образ жизни множества технических специалистов и представителей творческих профессий, а покупка, обучение и разработка подходящих нейросетей приобрели значительную коммерческую ценность, в которой когда-то сомневался Лайтхилл. Изменился и рынок труда: только в США за 2022 г. было открыто почти 800 тыс. вакансий, связанных с ИИ [2]. В целом, мы можем говорить о полном изменении институциональной структуры исследований. Ранее целый ряд фундаментальных исследований мог зависеть от одного инвестора, который, в свою очередь, при принятии решений о сокращении или увеличе-

нии финансирования мог руководствоваться мнением одной экспертной компании или даже только одного эксперта.

Сейчас эксперты продолжают делать осторожные либо пессимистические прогнозы. Исследователи и разработчики предостерегают от излишних надежд и предсказывают новую зиму, а пользователи с большим удовольствием находят ошибки и несовершенства нейросетей и генеративных языковых моделей. Так, нейросети и другие интеллектуальные системы, решая поставленные задачи, не отвечают за точность предоставленных фактов, дат, упоминаемых исторических событий; они приводят несуществующие источники и не держат контекст беседы; они не способны понимать многозначность слов и предложений; они значительно искажают образы и не схватывают структуры, в результате чего у футболистов появляются третьи и четвертые ноги, у художников – дополнительные пальцы, а текст приобретает особый вид, значительно отличающийся от любого текста, написанного человеком. Но все эти проблемы не являются нерешаемыми, а кроме того, они не всегда могут быть критичными в конкретной профессиональной сфере.

Вместо этого ожидания зимы искусственного интеллекта объясняется, во-первых, неоправданно высоким повсеместным интересом к теме высоких технологий и, во-вторых, желанием разработчиков получить дополнительное финансирование разработок в сфере ИИ в сочетании с опасением не оправдать ожидания массовой аудитории или инвесторов. Но текущая ситуация принципиально отличается от ситуации 1973 или 1986 г., поскольку большинство исследований имеют практическую и коммерческую ценность, а представители крупных IT-компаний, заявляя о кризисе и предполагаемых снижении инвестиций и сокращении исследований, одновременно финансирование увеличивают. Компании, использующие высокие технологии, неожиданно оказываются в инвестиционной гонке, и в этом случае сокращение исследований – последнее, чего разработчикам следует ожидать. Наконец, в 1964, 1973 и 1982 гг. переход от ноября ИИ к зиме ИИ совершался в результате некоторого отчета эксперта перед инвесторами либо утверждения эксперта, обладающего высокой репутацией или наделенного возможностью контролировать большую часть финансирования в сфере ИИ-разработок (Дж. Лайтхилл, ALPAC, Дж. Шварц, М. Мински и С. Паперт), однако сейчас способность единолично распоряжаться таким объемом исследований или подобное доверие к одному эксперту и даже команде экспертов кажутся невозможными.

Тем не менее, ряд экспертов вспоминают про возможную зиму. Более того, в марте 2023 г. было опубликовано открытое письмо, подписанное более чем тысячей исследователей, инженеров и предпринимателей, среди которых довольно известные личности, включая глав корпораций, с просьбой прекратить обучение нейросетей. В этом письме, помимо озвученных опасений о непредсказуемости продуктов разработок ИИ, предлагалось «избирать» технологических лидеров, которые в исключительном порядке могли бы заниматься инновационными разработками. Неизвестно, действительно ли главы корпораций готовы пойти на это и хотят ли они стать теми самыми «избранными» лидерами, но такое положение дел снова сделает зиму ИИ возможной, поскольку в этих случаях вместо конкуренции в сфере фундаментальной науки нас ожидает строго определенное развитие – или стагнация – под контролем «избранных» корпораций. Впрочем, возможно, это было лишь попыткой притормозить появление новых команд разработчиков и замедлить рост инвестиций в сторонние компании, не вовлеченные в написание «открытого письма». Тем не менее значимого эффекта это письмо не дало, а возможно, лишь добавило ажиотажа, но не в отношении всех исследований ИИ в целом, а только в отношении исследований в сфере генеративных языковых моделей.

Различные интеллектуальные и псевдоинтеллектуальные системы прочно вошли в нашу жизнь, и их развитие продолжается. Может быть, именно страх зимы ИИ побуждает отдельных разработчиков к все новым попыткам исследований, в том числе в тех сферах, которые ранее – ошибочно – были признаны неперспективными, как это было с коннекционизмом и машинным переводом. Однако необходимость разработки нейросетей для экономического успеха самых разных компаний, текущее состояние институциональной структуры и обилие экспертов привели к ситуации, когда потенциальные отчеты экспертов об упадке в ИИ-исследованиях не смогут повлиять на все исследовательское поле ИИ в целом и повлечь значимое сокращение его финансирования.

В таких условиях «зима» всего направления ИИ просто невозможна, хотя, разумеется, отдельные сферы исследований в области ИИ могут испытывать краткосрочные снижения финансирования.

Поэтому мы можем ожидать «зиму чистых генеративных моделей», «зиму компьютерного зрения» или еще одну «зиму машинного перевода», однако фундаментальные и концептуальные исследования будут продолжаться, не только развивая существующие направления ИИ-исследований, но и создавая новые, переходя от «лета» и «осени» ИИ, может, даже от «ноября» ИИ сразу к новой «весне». Однако такая смена времен лишь демонстрирует регулярность цикла инвестиционных влияний, а мы по-прежнему, как и пятьдесят лет назад, далеки от создания «общего» или «сильного» ИИ, точно как это отметил Дж. Лайтхилла: «Подводя итог, можно сказать, что это свидетельство и все остальные данные о работе ИИ в категории В за последние двадцать пять лет, полученные автором, в некоторой степени обнадеживают в отношении программ, написанных для работы в узкоспециализированных предметных областях, когда программирование занимает очень много времени, учитываются результаты человеческого опыта и человеческого интеллекта в соответствующей области. Но совершенно обескураживают программы общего назначения, стремящиеся имитировать аспекты решения проблем деятельности ЦНС человека в довольно широкой области. Такая универсальная программа, желанная долгосрочная цель деятельности ИИ, кажется отдаленной как никогда» [3].

Литература

1. *Crevier D.* AI: The Tumultuous Search for Artificial Intelligence. N.Y.: Basic Books, 1993.
2. *Goswami R.* Here's where the A.I. jobs are. CNBC. Apr 5, 2023. URL: <https://www.cnbc.com/2023/04/05/ai-jobs-see-the-state-by-state-data-from-a-stanford-study.html> (дата обращения: 01.10.2023).
3. *Lighthill J.* Artificial Intelligence: A General Survey // Artificial Intelligence: A paper symposium. UK: Science Research Council, 1973.
4. *McCorduck P.* Machines Who Think. 2nd ed. Natick, Mass.: A.K. Peters, Ltd., 2004.
5. *Minsky M., Papert S.* Perceptrons: An Introduction to Computational Geometry. 2nd ed. Cambridge, Mass.: The MIT Press, 1972.
6. *Nirenburg S., Somers H.L., Wilks Y.* ALPAC: The (in)famous report // Readings in Machine Translation. Cambridge, Mass.: The MIT Press, 2003. P. 131–135.
7. *Pause* Giant AI Experiments: An Open Letter (22 March 2023). URL: <https://futureoflife.org/open-letter/pause-giant-ai-experiments> (дата обращения: 01.10.2023).
8. *Russell S., Norvig P.* Artificial Intelligence: A Modern Approach. Prentice Hall, 2010.

References

1. *Crevier, D.* (1993). *AI: The Tumultuous Search for Artificial Intelligence*. New York, Basic Books.
2. *Goswami, R.* (2023). Here's where the A.I. jobs are. CNBC. Apr 5, 2023. Available at: <https://www.cnbc.com/2023/04/05/ai-jobs-see-the-state-by-state-data-from-a-stanford-study.html> (date of access: 01.10.2023).
3. *Lighthill, J.* (1973). *Artificial Intelligence: A General Survey*. In: *Artificial Intelligence: A paper symposium*. UK, Science Research Council.
4. *McCorduck, P.* (2004). *Machines Who Think*, 2nd ed. Natick, Mass., A.K. Peters, Ltd.
5. *Minsky, M. & S. Papert.* (1972). *Perceptrons: An Introduction to Computational Geometry*, 2nd ed. Cambridge, Mass., The MIT Press.
6. *Nirenburg, S., H.L. Somers & Y. Wilks.* (2003). ALPAC: The (in)famous report. In: *Readings in Machine Translation*. Cambridge, Mass., The MIT Press, 131–135.
7. *Pause Giant AI Experiments: An Open Letter* (22 March 2023). Available at: <https://futureoflife.org/open-letter/pause-giant-ai-experiments> (date of access: 01.10.2023).
8. *Russell, S. & P. Norvig.* (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

Информация об авторе

Зайкова Алина Сергеевна – Институт философии и права СО РАН (630090, Россия, г. Новосибирск, ул. Николаева, 8)/
zaykova.a.s@gmail.com

Information about the author

Zaykova, Alina Sergeevna – Institute of Philosophy and Law, Siberian Branch of the Russian Academy of Sciences (8, Nikolaev st., Novosibirsk, 630090, Russia).

Дата поступления 25.10.2023