

DOI: 10.26794/2587-5671-2024-28-3-206-217

УДК 33:004.8:336(045)

JEL G31

# Применение жизненного цикла модели для оценки инвестиций в искусственный интеллект на примере больших языковых моделей

Н.А. Никитин

Финансовый университет, Москва, Россия

## АННОТАЦИЯ

Объект исследования – жизненный цикл модели искусственного интеллекта (ИИ). **Цель** исследования состоит в разработке методологии жизненного цикла модели, описывающей экономическое содержание инвестиционного процесса в технологии искусственного интеллекта. В процессе исследования использовались как общенаучные **методы** анализа, синтеза, сравнения, абстракции, индукции и дедукции, так и проектные методологии жизненного цикла, взятые в качестве основы для разработки жизненного цикла модели с точки зрения создания стоимости. Анализ основывался на выявлении необходимых этапов разработки модели в терминах методологии CRISP-DM и определении особенностей каждого из них с точки зрения денежных потоков. Также были учтены модифицированные версии жизненного цикла модели, содержащие оценку рисков, в том числе модельного риска. В процессе исследования предложенная обобщенная методология жизненного цикла модели была уточнена для конкретной технологии ИИ – больших языковых моделей. В результате исследования автором предложена трехэтапная модель: описаны возможные опциональности между этапами и характеристика денежных потоков. Сделан **вывод** о том, что инвестиционный проект разработки ИИ содержит в себе несколько реальных опционов – на отказ, на сокращение, на расширение, на смену. Для больших языковых моделей сохраняется структура жизненного цикла и возможные опциональности. Особенность состоит в том, что в создании стоимости участвуют денежные потоки от разных направлений применения модели в бизнес-процессах. Результаты исследования имеют практическую значимость для среднего и крупного бизнеса, занимающегося самостоятельной разработкой ИИ моделей и/или применяющих их в своих бизнес-процессах. Предложенная концепция жизненного цикла модели также может использоваться для развития методологии оценки инвестиций в ИИ с использованием реальных опционов.

**Ключевые** слова: жизненный цикл модели; инвестиционная оценка; искусственный интеллект; денежные потоки; большие языковые модели

**Для цитирования:** Никитин Н.А. Применение жизненного цикла модели для оценки инвестиций в искусственный интеллект на примере больших языковых моделей. *Финансы: теория и практика*. 2023;28(3):206-217. DOI: 10.26794/2587-5671-2023-28-3-206-217

## ORIGINAL PAPER

# Application of a Model Life Cycle Concept to Investments in Artificial Intelligence Evaluation on the Example of Large Language Models

N.A. Nikitin

Financial University, Moscow, Russia

## ABSTRACT

The life cycle of an artificial intelligence model is the **object** of research. The **purpose** of the study is to develop a model life-cycle methodology that describes the economic content of the investment process in artificial intelligence technology. During the study, both general scientific **methods** such as analysis, synthesis, comparison, abstraction, induction and deduction were used, as well as project methodologies of the life-cycle, employed as the basis for the value creation life-cycle of the model. The analysis was based on identifying the necessary stages of model development in terms of the CRISP-DM methodology and determining the features of each of them in terms of cash flows. Modified versions of the model life-cycle containing risk assessment, including model risk, were also taken into account. In the process of research, the proposed generalized model life-cycle methodology was specified for a specific AI technology – large language models. As a result of the study, the author proposed a three-stage model. The possible optionality between the stages and the characteristics of cash flows are described. It was **concluded** that an investment project for the development of AI contains several real options – abandonment, reduction, expansion and replacement. For

large language models, the life cycle structure and possible optionalities are preserved. The peculiarity is that the value creation process involves cash flows from different areas of application of the model in business processes. The results of the study are of practical importance for medium and large businesses engaged in the independent development of AI models and/or applying them to their business processes. The proposed concept of the model life-cycle can also be used to develop a methodology for evaluating investments in AI using real options.

**Keywords:** model life-cycle; investment valuation; artificial intelligence; cash flows; large language models

**For citation:** Nikitin N.A. Application of a model life cycle concept to investments in artificial intelligence evaluation on the example of large language models. *Finance: Theory and Practice*. 2024;28(3):206-217. (In Russ.). DOI: 10.26794/2587-5671-2024-28-3-206-217

## ВВЕДЕНИЕ

В период с 2010-х гг. искусственный интеллект (далее — ИИ) переживает новый этап развития и тесной интеграции в общественные процессы, в том числе в экономические. Только за период с 2013 по 2021 г. объем корпоративных инвестиций в технологии ИИ, по данным Стэнфордского университета, вырос с 17 до 276 млрд долл. США [1]. Аналитические агентства оценивают вклад ИИ в ВВП мировой экономики на уровне 1,5% и прогнозируют, что ИИ станет новым драйвером для экономического роста в следующие 30 лет. Компании, инвестирующие в данный вид технологий, заинтересованы в наиболее оптимальном использовании ресурсов, что достигается за счет организации эффективного процесса разработки и внедрения. Этот процесс создания, внедрения и коммерциализации инноваций искусственного интеллекта называется жизненным циклом модели (далее — ЖЦМ). Понимание ключевых этапов жизненного цикла модели и их вклад в создание добавленной стоимости предприятия позволяет принимать корректные инвестиционные решения.

Одна из многообещающих технологий ИИ — большие языковые модели (БЯМ) — это самый современный метод работы с текстами и решением задач обработки естественного языка. Их ключевой особенностью является использование большого количества параметров для обучения (например, ChatGPT-4 использует более 100 трлн параметров), что позволяет генерировать осмысленный текст. Наиболее известными являются модели, разработанные американскими компаниями, такими как OpenAI (модели GPT-2, GPT-3, GPT-4), Google (модель BERT-large). Российские компании создали свои версии языковых моделей: например, в 2023 г. Яндекс внедрил в свои программные продукты YandexGPT, а Сбербанк разработал ГигаЧат. Объединенные с генеративными моделями<sup>1</sup> БЯМ решают значительно больший круг задач по

сравнению с моделями машинного обучения, которые направлены на решение одной конкретной задачи.

Несмотря на большее приближение к концепции общего искусственного интеллекта, предполагающего выполнение машиной широкого спектра задач на уровне человеческого интеллекта, за данным видом технологий сохраняется риск, присущий традиционным моделям машинного обучения, — модельный риск. Суть модельного риска состоит в снижении предикативных качеств модели. Показательным является исследование работы GPT-3.5 и GPT-4, показавшее, что за четыре месяца работы GPT-4 правильность<sup>2</sup> определения простых чисел снизилась с 97,6 до 2,4%, а доля корректных программных кодов, написанных моделью, сократилась с 52 до 10% [2].

Методология жизненного цикла может применяться при оценке инвестиций в современные технологии ИИ, но для этого требуется проанализировать существующие концепции и предложить такую модель, которая бы учитывала и проектные особенности разработки ИИ, и их экономическое содержание. Еще одной из задач исследования является апробация модели жизненного цикла посредством применения ее к самым современным технологиям ИИ — большим языковым моделям.

Научная новизна исследования заключается в разработке концепции жизненного цикла модели с точки зрения создания стоимости и ее применения для оценки инвестиций в технологии искусственного интеллекта. Теоретическая значимость обусловлена объединением проектного подхода к жизненному циклу модели с экономическим содержанием инвестиционного процесса и анализа встроенных опциональностей в нем. Возможность применения предложенной концепции жизненного цикла модели предприятиями при оценке инвестиций в ИИ определяет практическую значимость исследования.

<sup>1</sup> Прим. автора: генеративный ИИ, сокращенно от «генеративный искусственный интеллект», — это тип системы ИИ, которая может генерировать уникальный или оригинальный контент, такой как текст, аудио, видео или изображения по запросу.

<sup>2</sup> Прим. автора: правильность (ассурасу) показывает долю правильных ответов из всех вариантов исходов. Является наиболее примитивной при расчете, однако не показывает соотношение критически важных ошибок первого или второго рода применительно к конкретному типу задач.

## ЭВОЛЮЦИЯ КОНЦЕПЦИИ ЖИЗНЕННОГО ЦИКЛА МОДЕЛИ

В иностранной научной литературе концепция жизненного цикла модели возникла в начале 2000-х гг. в рамках проектного подхода, в которых процесс работы с данными организовывался как самостоятельный проект. Итогом такого проекта являлась интеллектуальная система (модель) по анализу данных. Данная методология в 2000 г. была сформулирована Колином Шерером и получила название Межотраслевой стандартный процесс интеллектуального анализа данных (CRoss-Industry Standard Process for Data Mining — CRISP-DM) [3]. Данная модель включает в себя шесть ключевых этапов разработки моделей анализа данных:

1. Бизнес-анализ. Данный этап подразумевает анализ ключевых целей бизнеса и способов их достижения с помощью моделей данных. В результате анализа бизнес-цели связываются с конкретными проблемами анализа данных.

2. Обработка данных. Второй этап подразумевает аналитическую работу с данными, необходимыми для достижения целей проекта. На данном этапе производится поиск взаимосвязей, проверка гипотез, подбор необходимых параметров и данных.

3. Подготовка данных. Представляет из себя техническую сторону процесса работы с данными и важнейшим этапом перед выбором модели, так как от качества данных и уровня их обработки будет зависеть финальный результат.

4. Моделирование. На данном этапе применяются различные техники обучения модели и параметры и на основании объективных статистических показателей калибруются до оптимальных решений. У разработчика есть возможность определять технику обучения и используемые параметры в зависимости от типа решаемой задачи. Дополнительные ограничения накладываются качеством и типом данных.

5. Оценка решения. По результатам разработки необходимо проконтролировать эффективность как статистическую, так и операционную. Результат работы модели всегда должен приносить улучшение бизнес-процесса, а также соответствовать всем ключевым требованиям и условиям его существования.

6. Внедрение. Финальный этап, подразумевающий введение модели в бизнес-процесс и его полное использование клиентом. Здесь выделяются такие этапы, как разработка плана внедрения, мониторинг и поддержание, финальный отчет по результатам проекта (рис. 1).

Значительно позже данная методология CRISP-DM дорабатывалась с целью упрощения взаимодейст-

вия участников процесса между собой. В результате появились модели Командного процесса обработки данных (Team Data Science Process — TDSP), которые выделяли четыре основные активности: бизнес-анализ, анализ данных, моделирование и внедрение таким образом, чтобы команды могли повторять их на любом этапе производственного цикла [4]. Аналогичные собственные модели были предложены компаниями Microsoft и IBM.

В 2021 г. исследователи на основании опросов руководителей финтех-компаний приходят к выводу, что в модели жизненного цикла необходимо детализировать оценку рисков и правила контроля за последствиями работы модели [5]. В 2022 г. была предложена концепция ЖЦМ, учитывающая работу как с операционными рисками, так и с социальными, этическими последствиями внедрения ИИ [6]. Концепция CDAC AI life cycle сохраняет проектный подход, разделяя его на три этапа — проектирование, разработка, внедрение, но при этом определяет обязательные шаги в каждом из этапов, учитывающие вышеперечисленные аспекты. Например, этап проектирования требует не только определения бизнес-проблемы, но и этические аспекты использования модели, на этапе разработки анализируется корректность интерпретации результатов работы модели, а внедрение требует анализ и оценку рисков.

## СТОИМОСТНОЙ ПОДХОД К ОПРЕДЕЛЕНИЮ ЖИЗНЕННОГО ЦИКЛА МОДЕЛИ

Российская исследовательская литература не выделяет жизненный цикл модели как самостоятельный объект инновационной и инвестиционной деятельности предприятия. Как правило, технологии ИИ рассматриваются в рамках жизненного цикла инноваций, который обобщенно передает основные этапы их создания:

1. Период зарождения инновационного продукта.
2. Период создания продукта.
3. Период вывода продукта на рынок.
4. Период формирования зрелости продукта [7].

Исследования в области применения ИИ предлагают направления его применения в различных отраслях. Например, с помощью ИИ решают проблемы повышения эффективности в розничной торговле, медицине, строительстве, транспорте и других отраслях [8]. Исследования, направленные на оценку эффективности инвестиций в сами технологии ИИ, все еще недостаточно представлены в отечественной научной литературе. Цикл создания инновационного продукта применим к технологиям искусственного интеллекта,

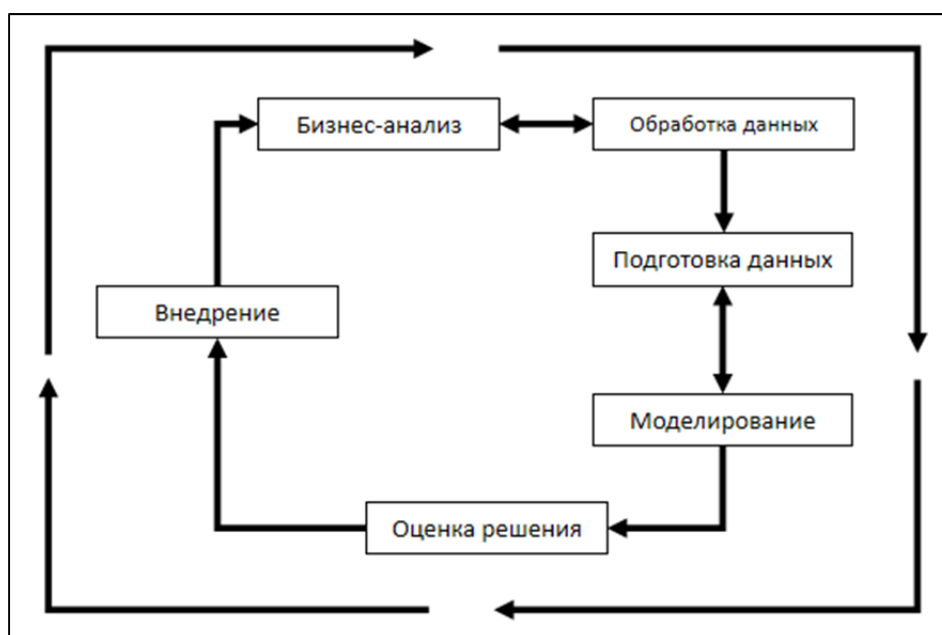


Рис. 1 / Fig. 1. Иллюстрация методологии межотраслевого стандартного процесса интеллектуального анализа данных (CRISP-DM) / Cross-Industry Standard Process for Data Mining (CRISP-DM) Methodology Illustration

Источник / Source: составлено на основании [4] / Compiled by the author based on [4].

но не учитывает его ключевую специфику — наличие модельного риска. Например, в отрасли финансовых технологий риски, связанные с развертыванием и стоимостью разработки ИИ, а также риски, связанные с качеством данных, алгоритмов, результатов работы, создают угрозу как крупным, так и средним участникам рынка [9]. Кроме того, технологии ИИ не всегда выступают в качестве продукта для внешнего клиента: часто разработка моделей ориентирована на внутренние потребности бизнеса с целью улучшить операционные процессы предприятия [10].

В зарубежной литературе данная проблема стала актуальной в связи с трансформацией бизнеса под воздействием ИИ. Например, исследователи Орстром и Рейм рассматривают разные модели создания стоимости предприятиями, разрабатывающими ИИ-модели, и выделяют три фазы инновационной бизнес-модели:

1. Определение предпосылок для создания стоимости в результате использования ИИ.
2. Определение технологий, необходимых для создания стоимости.
3. Разработка бизнес-модели предложения результатов работы ИИ [11].

Предложенные этапы включают в себя как бизнес-анализ, анализ технологий, потребности секторов экономики и потребителей, так и оценку потенциальных рисков. При этом рассматривается исключительно работа ИИ в качестве внешнего продукта. Данная

модель может описывать цепочку создания стоимости при коммерциализации технологий ИИ, но не всегда только их коммерциализация приносит эффект предприятию. Авторы замечают, что ИИ может влиять на доходы от направлений бизнеса, не связанных с ИИ напрямую, а также повышать операционную эффективность — снижать затраты предприятия.

В исследовании Моро-Висконти, посвященном оценке стоимости фирм, инвестирующих в ИИ, отмечается, что при их оценке необходимо использовать методику реальных опционов, так как успешные разработка и внедрение способны значительно улучшить показатели денежного потока фирмы. При этом необходимо отличать фирмы, у которых разработка ИИ основной вид деятельности от тех, которые используют ИИ-технологии для улучшения существующих процессов операционной деятельности. Анализ последних должен подразумевать определение инкрементальной стоимости как разницу между стоимостями «с применением ИИ в процессах» и «без применения ИИ в процессах» [12]. Описанные выше методологии жизненного цикла не позволяют оценить эффект опциональности, выгоды которой предприятие может оценивать и использовать на всех этапах разработки.

Для разработки стоимостной модели жизненного цикла в качестве каркаса будет использоваться модель CRISP-DM, так как она описывает необходимые этапы процесса работы с данными. Однако видим необходимость сократить их количество до трех: этапа



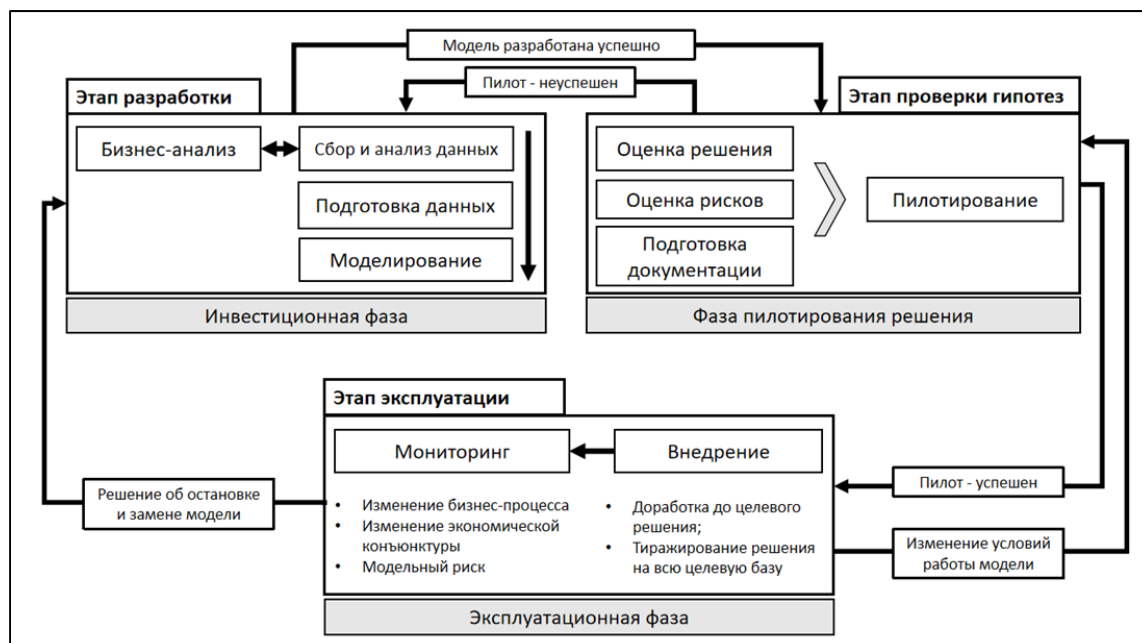


Рис. 2 / Fig. 2. Жизненный цикл модели ИИ с точки зрения создания стоимости / The Life Cycle of an AI Model in Terms of Value Creation

Источник / Source: составлено автором / Compiled by the author.

разработки, проверки гипотезы и этапа эксплуатации (рис. 2).

Этап разработки включает в себя бизнес-анализ, работу с данными и моделирование. На данном этапе предприятие осуществляет следующие действия:

1. Формирование первоначального замысла, определение бизнес-проблемы и связи между ней и задачей машинного обучения.
2. Оценка потребности в человеческих, технологических, финансовых ресурсах.
3. Разработка технического требования к модели искусственного интеллекта, в том числе к данным.
4. Разработка системы показателей успешности проекта, включающих в себя оценку статистической эффективности модели, операционные и финансовые метрики.
5. Непосредственная работа с данными, связанными с их приобретением, хранением, анализом и обработкой.
6. Обучение модели искусственного интеллекта, предварительная оценка качества ее работы.

После разработки модели / группы моделей и предварительной оценки их качества возникает опцион отказа: в случае, если модель не соответствует техническим требованиям процесса, то у предприятия есть возможность прекратить ее разработку. В таком случае компания избежит потенциальных убытков, которые могли бы возникнуть в случае эксплуатации модели. Кроме того, предварительный контроль качества модели позволит предотвратить ухудшение качества

процесса. Так как этап разработки подразумевает значительный объем инвестиций в данные, программное и аппаратное обеспечение, оплату труда специалистов, то в случае принятия решения о выходе из проекта компания сможет переориентировать часть ресурсов либо реализовать на рынке (например, продажа аппаратного обеспечения по остаточной стоимости).

Этап проверки гипотезы сам по себе является опционом на расширение. Он включает в себя следующие действия:

1. Оценка решения с точки зрения оптимальности и максимизации целевого показателя относительно других методов моделирования, в том числе без применения искусственного интеллекта.
2. Оценка рисков как для самой компании, так и для внешних участников экономических отношений, ведущих к возможному репутационным и финансовым издержкам.
3. Подготовка документации о валидации с детальным описанием работы модели, требуемыми данными, оценками рисков и качества модели.
4. Запуск модели на ограниченной части процесса (пилотная группа) с целью получения объективных фактических данных об эффективности процесса для принятия решения о расширении использования модели на весь процесс.

При проведении пилотирования ИИ решения формируют контрольную и тестовую группы: первая подразумевает использование решений без ИИ и ее результаты сравниваются с работой процесса

с применением ИИ. Для оценки эффективности используют статистические тесты для проверки значимости показателей, которые были выбраны в качестве индикаторов успеха [13]. Не всегда после улучшения статистической метрики улучшаются операционные или финансовые показатели. Например, в статье, описывающей применение ИИ в практике рентгенографии, делается вывод о том, что даже точные модели при применении в реальном процессе не всегда увеличивают качество диагностики, особенно во взаимодействии с сотрудником-человеком [14].

Данный этап может отсутствовать, если компания не создала условия для принятия решений на основе данных или отсутствует процесс валидации моделей машинного обучения сам по себе. Кроме того, пилотирование помогает зафиксировать взаимосвязь модельного риска и финансового эффекта.

С точки зрения создания стоимости этот этап увеличивает значимость управленческих решений, так как встроенный в него опцион позволяет малыми издержками нивелировать риск неопределенности инкрементальных денежных потоков инвестиций после внедрения модели в бизнес-процесс в целом. При оценке проектов с ИИ наличие этапа проверки гипотез определяет точность экономической оценки. Таким образом, инкрементальный денежный поток пилота от ИИ в момент  $t$  будет измеряться по формуле (1):

$$CF_t = CF_t^{ИГ} - CF_t^{КГ}, \quad (1)$$

где  $CF_t^{ИГ}$  — денежный поток от процесса с ИИ в целевой группе;  $CF_t^{КГ}$  — денежный поток от процесса без ИИ в контрольной группе.

На основании пилотных значений можно спрогнозировать денежные потоки после внедрения модели посредством экстраполяции денежных потоков за единицу драйвера пилота на объем драйвера в генеральной совокупности. Критерием успешности пилота может быть положительная чистая приведенная стоимость (NPV) [15]. В случае неуспешности пилота возникает либо опцион на смену, в результате которого проект возвращается в инвестиционную фазу, либо опцион отказа от инвестиций в данную модель и реализация ресурсов по остаточной стоимости.

Этап эксплуатации следует после положительного управленческого решения о тиражировании модели на весь бизнес-процесс. С момента внедрения процесс мониторинга моделей является обязательным условием контроля за операционными рисками. В силу существования модельного риска изменениями экономической среды или самого бизнес-процесса, в котором работает модель, необходима регулярная валидация ее эффективности. С точки зрения цепочки

создания стоимости данный этап представляет из себя доходную часть проекта. Одним из условий является необходимость учитывать риски. Например, в качестве решения можно использовать коэффициенты, снижающие объем денежных потоков на величину вероятности наступления этого вида риска. В формуле (2) представлена концепция учета денежных потоков на этапе эксплуатации с момента времени  $t + 1$  и для  $n$  периодов:

$$\sum CF_{\text{эксплуатации}} = CF_{t+1} * k_t + \dots + CF_{t+n} * k_{t+n}, \quad (2)$$

где  $CF_{t+n}$  — денежный поток тиража в период после окончания пилота в момент времени  $t$ ;  $k_t$  — это корректирующий коэффициент, учитывающий модельный риск.

Значение корректирующего коэффициента изменяется от 0 (модельный риск нивелирует инкрементальный денежный поток от внедрения ИИ) до 1 (отсутствие модельного риска). Его использование обусловлено необходимостью учета несовершенства технологий ИИ, которые не всегда реализуют функцию соразмерно результатам человеческой деятельности [16].

На рис. 3 показаны сравнительные результаты различных моделей по отношению к результатам человеческой деятельности для разных видов задач: распознавание текста, речи и изображений, понимание языка и прочитанного. Несмотря на то что современные модели, например, SQuAD и Glue, достаточно быстро развивали свою точность, разработчикам потребовалось не менее года, чтобы улучшить модель. А на примере ImageNet можно видеть сравнительное ухудшение результатов работы, что и является примером модельного риска.

В соответствии с методологией учета модельного риска каждый момент времени в период с  $t + 1$  до  $t + n$  имеет встроенный опцион на сокращение работы модели с целью проверки гипотезы. Так как не всегда соблюдается прямая взаимосвязь между статистической метрикой качества, которую можно объективно пронаблюдать, и инкрементальным экономическим эффектом от внедрения ИИ, то возвращение на этап проверки гипотезы является оптимальным решением менеджмента.

Таким образом инвестиционный проект внедрения ИИ-решения представляется сложным опционом, который содержит в себе разные виды опциональностей. На рис. 4 представлена упрощенная схема взаимосвязи денежных потоков, этапов жизненного цикла модели и встроенных опциональностей. В отличие от проектного подхода, используемого ранее для определения жизненного цикла модели ИИ, подход

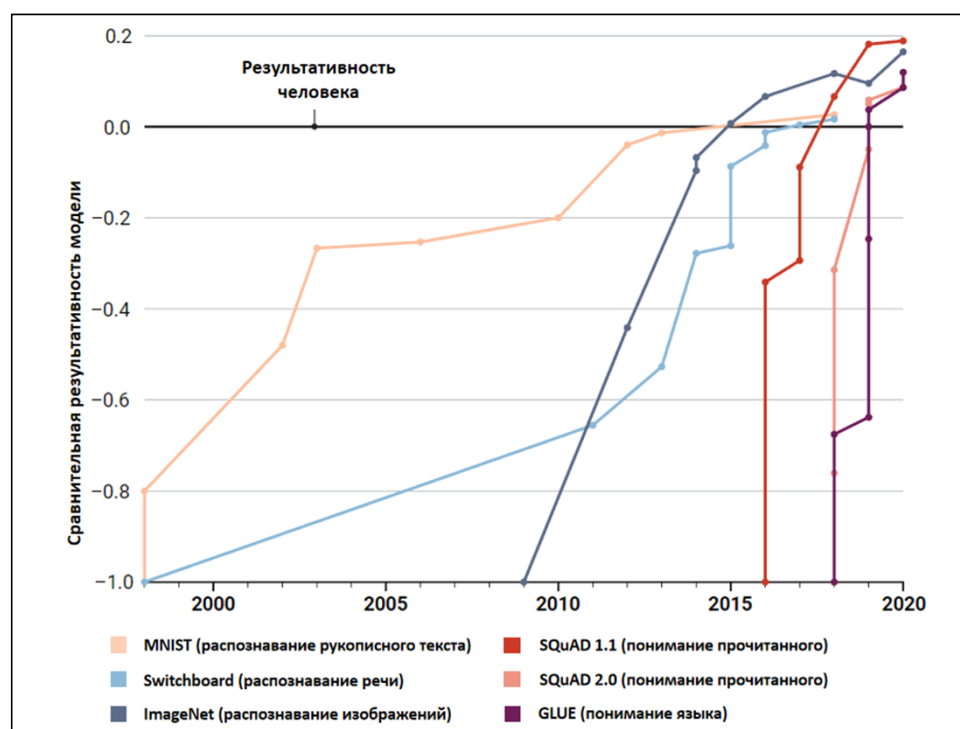


Рис. 3 / Fig. 3. Развитие результативности моделей, решающих различные задачи машинного обучения / Development of the Effectiveness of Models that Solve Various Machine Learning Problems

Источник / Source: составлено на основании [17] / Compiled by the author based on [17].

с точки зрения создания стоимости позволяет определить дискретность денежных потоков от ИИ проекта, а также учесть их опциональный характер. Элементы представленного нами жизненного цикла зависят от организационной практики разработки и внедрения ИИ-моделей: например, отсутствие процесса валидации не позволит нам учитывать опцион отказа, так как предприятие не сможет определить, в какой момент эффективность модели стала отрицательной, а отсутствие практики проверки гипотез не позволяет учитывать опцион на расширение.

### ПРИМЕНЕНИЕ КОНЦЕПЦИИ ЖИЗНЕННОГО ЦИКЛА ДЛЯ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Особенности больших языковых моделей в контексте жизненного цикла, описанного выше, вносят изменения в его этапы. Это связано с тем, что классические модели машинного обучения решают одну определенную задачу. Например, задача регрессии направлена на прогнозирование конкретного значения показателя, тогда как БЯМ могут решать широкий набор задач. Современные англоязычные статьи, исследующие работу GPT-4, отмечают, что большие языковые модели являются шагом к общему искусственному интеллекту, так как, помимо высокого качества работы с текстом и его содержанием, способны решать следующие задачи: генерация

изображений, написание музыки, решение простых математических задач, написание программного кода, понимание контекста медицинских, правовых и психологических проблем. С частью этих задач GPT-4 справляется не хуже, чем человек [18].

БЯМ представляют из себя набор нейронных сетей для обучения на последовательных данных без учителя на неаннотированном тексте. Кроме того, большие языковые модели обладают долгой кратковременной памятью, что позволяет им строить ответ на запрос на основании ранее полученных вводных, а также прогнозировать следующее предложение. Единица текста, на которой обучается модель и которую она выдает в качестве результата, называется токеном. Очевидно, что для обучения БЯМ требуется значительный объем данных, а для их обработки — аппаратное обеспечение. На рис. 5 представлен объем условных единиц данных, необходимых для обучения различных моделей ИИ. Условная единица данных — это такая единица, которая не учитывает тип данных: например, изображение и слово могут быть равны одной единице условных данных. На рисунке видно, что значительный рост объема используемых данных произошел в 1990-е гг. и достиг 3 трлн ед. данных в 2023 г. для языковой модели Palm2. Это связано с тремя основными факторами:

1. Накопление данных в цифровой форме.
2. Развитие технологий хранения данных.

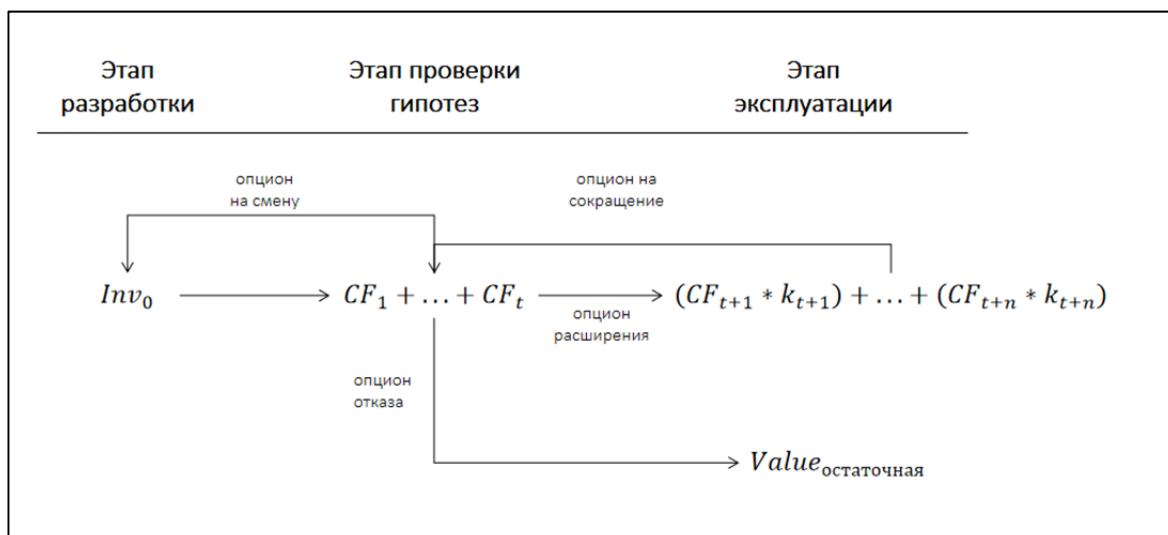


Рис. 4 / Fig. 4. Схема встроенных опциональностей в жизненном цикле модели ИИ с точки зрения создания стоимости / Scheme of Built-In Options in the Life Cycle of an AI Model from a Value Creation Perspective

Источник / Source: составлено автором / Compiled by the author.

3. Развитие технологий обработки больших массивов данных.

С точки зрения применения в бизнес-процессах можно выделить следующие направления:

1. Альтернатива традиционному поиску.

Компании активно внедряют БЯМ в качестве альтернативы поисковому движку, что позволяет пользователю быстрее находить необходимую информацию или агрегировать ее из нескольких источников, не обращаясь непосредственно к сайтам. Ключевым изменением является предоставление не набора ссылок, а готовый ответ пользователю. Примеры применений: поисковики Yandex, Google, Bing.

2. Копилоты (Copilots).

Является инструментом для увеличения производительности сотрудников и представляет из себя БЯМ, дообученный на конкретном наборе данных, необходимых специалисту. Известным примером является JARVIS, разработанный российской компанией Сбер на базе GPT-3. Она позволяет программистам без дополнительных запросов в интернете писать простейший код и находить ошибки в текущем.

3. Умный ассистент.

В отличие от копилотов, которые направлены на увеличение производительности работы сотрудников за счет взаимодействия с ним, умные ассистенты позволяют заместить прямое общение клиента и сотрудника. Умные ассистенты могут быть встроены в рекомендательные системы, позволяющие быстро получить информацию о товаре или услуге, а иногда и получить эту услугу. Например, в медицине БЯМ могут использоваться для проведения первичного

осмотра и сбора анамнеза у пациента. Мультимодальность БЯМ, т.е. возможность работать как с текстом, так и с изображением, позволяет использовать их и в качестве копилотов для врачей [20].

4. Генеративный контент.

БЯМ, особенно мультимодальные, позволяют заменить ряд рутинных операций по созданию контента по запросу. Например, БЯМ способны написать связанный текст на определенную тематику или сгенерировать картинку. Данное направление широко применяется в маркетинге и рекламе и позволяет увеличить производительность сотрудников при создании контента [21].

5. Реализация результатов работы модели за подписку.

Данный подход может быть использован компаниями, которые сами разрабатывают БЯМ модели и имеют патент на данную разработку, что позволяет продавать пользователям подписку на использование модели. Данное направление реализуется, например, OpenAI, которое за определенную плату предоставляет ограниченный объем токенов для запросов ChatGPT.

Проиллюстрируем оценку денежных потоков от одного из направлений внедрения БЯМ. Предположим, что ИТ-компания внедряет БЯМ в процесс разработки программных продуктов в формате копилота. Средняя стоимость труда одного разработчика в год составляет 1,5 млн руб. в год. На предприятии работает 1000 разработчиков. Без учета косвенных затрат годовой денежный поток составляет 1500 млн руб. в год.



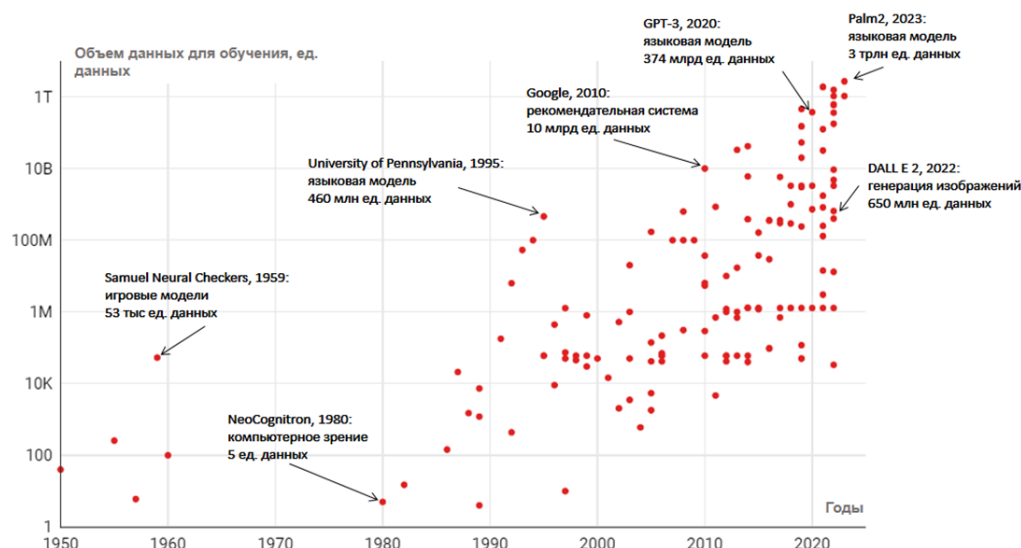


Рис. 5 / Fig. 5. Объем условных единиц данных, необходимых для обучения разных моделей ИИ / The Volume of Conventional Data Units Needed to Train Different AI Models

Источник / Source: составлено на основании [19] / Compiled by the author based on [19].

1. Для того чтобы подтвердить гипотезу об эффективности внедрения модели и положительном влиянии на производительность, предприятие проводит пилот и определяет по 100 сотрудников в целевую и контрольную группы. В результате фиксируется 56%-ное увеличение производительности труда, что позволяет снизить издержки (за счет сокращения штата) на кратную величину<sup>3</sup>. Годовой денежный поток пилота в контрольной группе составляет 150 млн руб., а в целевой группе — 84 млн руб. В этом случае, в соответствии с формулой (1), компания получает значение инкрементального денежного потока, равное 66 млн руб.

2. Для оценки денежных потоков от эксплуатации компания использует предпосылку, что качество результатов модели снижается вдвое. Корректирующий коэффициент в первый год составляет  $k_1 = 0,5$ , а во второй год работы модели  $k_2 = 0,25$ . По истечении двух лет эксплуатация прекращается. Годовой инкрементальный денежный поток на этапе эксплуатации без учета ухудшения качества модели составляет 660 млн руб., так как распространяется на всех 1000 сотрудников. За весь этап эксплуатации денежный поток составит:

$$\begin{aligned} \sum CF_{\text{эксплуатации}} &= 660 \text{ млн руб.} \cdot 0,5 + \\ &+ 660 \text{ млн руб.} \cdot 0,25 = 495 \text{ млн руб.} \end{aligned} \quad (3)$$

<sup>3</sup> Прим. автора: используется реальный рост производительности труда, зафиксированный в исследовании, посвященном оценке роста производительности труда разработчиков программного обеспечения за счет использования модели GitHub Copilot [22].

Таким образом, принятие решения о внедрении строится на основании инкрементальных денежных потоков от всех потенциально возможных направлений реализации данного инструмента. На рис. 6 представлена схема жизненного цикла БЯМ для компании, занимающейся ее разработкой и внедрением в собственные операционные процессы, а также продаж ее в качестве продукта.

Этап разработки включает в себя два направления инвестиций: разработка общей языковой модели ( $Inv_0$ ) и инвестиции, связанные с обучением модели под конкретную задачу, требующие определенных данных, которые в открытом доступе отсутствуют ( $Inv_i$ ). Этот этап в разработке ИИ называется «тонкая настройка» (от англ. fine-tuning) [23]. Например, компания обучает БЯМ для того, чтобы работать с внутренней юридической документацией или документацией клиента. Процедура тонкой настройки позволяет повысить эффективность модели для решения специфических задач. Объем таких инвестиций определяется на основании каждого  $i$ -го направления внедрения, требующего дообучения.

Этап проверки гипотез аналогичен процессу в машинном обучении с той особенностью, что пилотирование проводится для каждого направления и на основании него принимается решение о доработке или вводе в эксплуатацию. Соответственно при оценке денежных потоков учитывается инкремент от каждого направления ( $CF_{t,i}$ ).

На этапе эксплуатации усложняется процедура мониторинга, так как требует оценки эффектив-

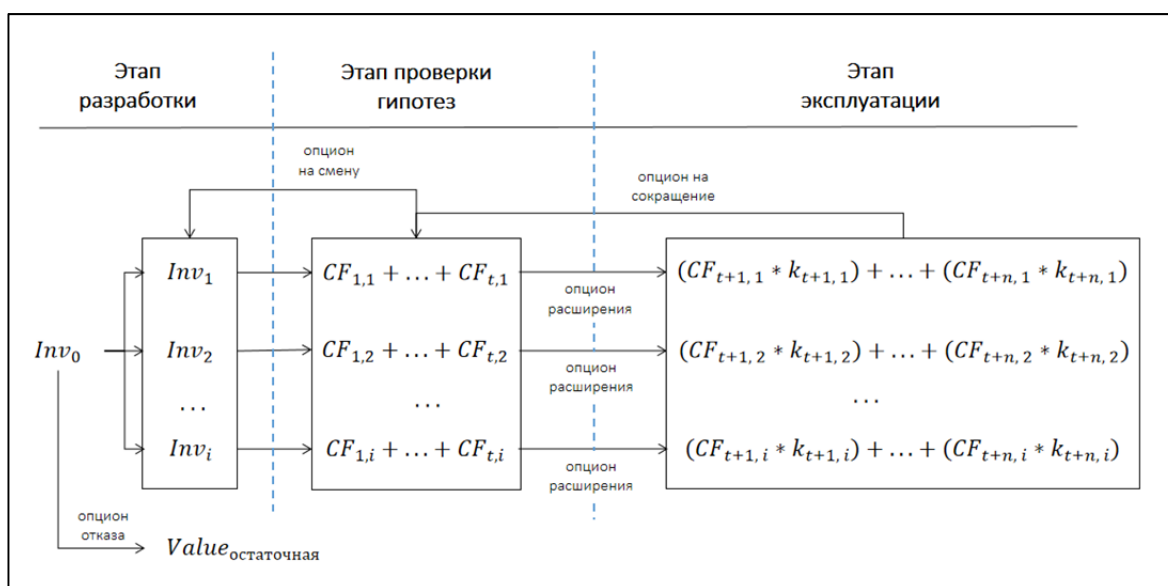


Рис. 6 / Fig. 6. Жизненный цикл большой языковой модели с точки зрения создания стоимости / Life Cycle of Large Language Models in Terms of Value Creation

Источник / Source: составлено автором / Compiled by the author.

ности выполнения моделью задач каждого направления. Ранее мы приводили пример исследования GPT, демонстрирующие значительное изменение в правильности выполнения ряда задач — написание кода, понимание изображений. Для переоценки денежных потоков с учетом наступления модельного риска необходимо использовать свой корректирующий коэффициент ( $k_{t,i}$ ) в зависимости от того, какую задачу решает модель.

## ВЫВОДЫ

Существующие концепции жизненного цикла модели (CRISP-DM, TDSP и другие) позволяют определить содержание процесса разработки ИИ технологий как проекта, однако они не раскрывают экономическое содержание инвестиций в ИИ. Предложенная нами методология жизненного цикла модели не противоречит проектному подходу и позволяет раскрыть его экономическое содержание с точки зрения создания стоимости. Описанные нами этапы и переходы между ними определяют существующие опциональности — расширение, сокращение, смену или отказ от инвестиций, использование которых влияет на итоговую оценку инвестиций. Использование данной методологии расширяет потенциал оценочного инструментария и дальнейших исследований: инвестиционный проект в технологии ИИ может рассматриваться как система встроенных реальных опционов.

Еще одним новшеством методологии является выделение процедуры проверки гипотез в качестве

самостоятельного этапа. Ранее оценка эффективности модели ограничивалась проверкой статистического качества работы модели: нами была предложена практика, распространенная в оценке инноваций, — проведение пилотирования. Применение данного подхода к проверке гипотез позволяет малыми затратами получить фактическое знание не только о статистическом качестве, но и операционном процессе и его экономике. Специфика проведения пилотов для оценки инвестиций в технологии ИИ также может быть предметом будущих исследований.

Методология ЖЦМ была апробирована на конкретном типе моделей — больших языковых моделях. Так как в отличие от подмножества моделей ИИ, относящихся к машинному обучению, БЯМ являются значительным шагом к общему искусственному интеллекту в части разнообразия решаемых задач, то были предложены следующие корректировки:

1. Разделение этапа разработки на две инвестиционные фазы — инвестиции в разработку общей языковой модели и инвестиции в тонкую настройку модели.
2. Разделение жизненного цикла на  $i$ -е количество направлений внедрения с выделением отдельных инкрементальных потоков для каждого из них.
3. Коэффициент учета модельного риска является уникальным для каждого направления внедрения БЯМ.

Предложенная методология может быть использована при оценке целесообразности инвестиций в разработку БЯМ крупными предприятиями.

## СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Maslej N., Fattorini L., Brynjolfsson E., et al. Artificial intelligence index report 2023. Stanford, CA: Institute for Human-Centered AI, Stanford University; 2023. 386 p. URL: [https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI\\_AI-Index-Report\\_2023.pdf](https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf) (дата обращения: 01.08.2023).
2. Chen L., Zaharia M., Zou J. How is ChatGPT's behavior changing over time? *Harvard Data Science Review*. 2024;(6.2):1–47. DOI: 10.1162/99608f92.5317da47
3. Shearer C. The CRISP-DM model: The new blueprint for data mining. *Journal of Data Warehousing*. 2000;5(4):13–22. URL: <https://mineraodados.wordpress.com/wp-content/uploads/2012/04/the-crisp-dm-model-the-new-blueprint-for-data-mining-shearer-colin.pdf>
4. Tabladillo M. The Team Data Science Process lifecycle. 2017;552–554. URL: <https://learn.microsoft.com/pdf?url=https%3A%2F%2Flearn.microsoft.com%2Fen-us%2Fazure%2Farchitecture%2Fai-ml%2Ftoc.json> (дата обращения: 01.08.2023).
5. Haakman M., Cruz L., Huijgens H., van Deursen A. AI lifecycle models need to be revised. *Empirical Software Engineering*. 2021;26:95. DOI: 10.1007/s10664-021-09993-1
6. De Silva D., Alahakoon D. An artificial intelligence life cycle: From conception to production. *Patterns*. 2022;3(6):100489. DOI: 10.1016/j.patter.2022.100489
7. Голубев А.А. Жизненный цикл инновации и ресурсное обеспечение инновационной деятельности. *Современные проблемы науки и образования*. 2015;(2–2):414. URL: <https://science-education.ru/ru/article/view?id=22026> (дата обращения: 05.08.2023).  
Golubev A.A. The life cycle of innovation and resource support innovation. *Sovremennye problemy nauki i obrazovaniya = Modern Problems of Science and Education*. 2015;(2–2):414. URL: <https://science-education.ru/ru/article/view?id=22026> (accessed on 05.08.2023). (In Russ.).
8. Городнова Н.В. Применение искусственного интеллекта в бизнес-сфере: современное состояние и перспективы. *Вопросы инновационной экономики*. 2021;11(4):1473–1492. DOI: 10.18334/vinec.11.4.112249  
Gorodnova N.V. Application of artificial intelligence in the business sphere: Current state and prospects *Voprosy innovatsionnoi ekonomiki = Russian Journal of Innovation Economics*. 2021;11(4):1473–1492. (In Russ.). DOI: 10.18334/vinec.11.4.112249
9. Кашеварова Н.А., Панова Д.А. Анализ современной практики применения технологии искусственного интеллекта в финансовой сфере и его влияния на трансформацию финансовой экосистемы. *Креативная экономика*. 2020;14(8):1565–1580. DOI: 10.18334/ce.14.8.110708  
Kashevarova N.A., Panova D.A. Analysis of the current practice of applying artificial intelligence in the financial sector and its impact on the transformation of the financial ecosystem. *Kreativnaya ekonomika = Journal of Creative Economy*. 2020;14(8):1565–1580. (In Russ.). DOI: 10.18334/ce.14.8.110708
10. Устинова О.Е. Искусственный интеллект в менеджменте компаний. *Креативная экономика*. 2020;14(5):885–904. DOI: 10.18334/ce.14.5.102145  
Ustinova O.E. Artificial intelligence in company management. *Kreativnaya ekonomika = Journal of Creative Economy*. 2020;14(5):885–904. (In Russ.). DOI: 10.18334/ce.14.5.102145
11. Åström J., Reim W., Parida V. Value creation and value capture for AI business model innovation: A three-phase process framework. *Review of Managerial Science*. 2022;16(7):2111–2133. DOI: 10.1007/s11846-022-00521-z
12. Moro-Visconti R. The valuation of artificial intelligence. In: *The valuation of digital intangibles: Technology, marketing, and the metaverse*. Cham: Palgrave Macmillan; 2022:265–282. DOI: 10.1007/978-3-031-09237-4\_8
13. Никитин Н.А. Финансовая оценка проектов с искусственным интеллектом в банковском секторе. *Финансовый бизнес*. 2022;(5):122–125.  
Nikitin N.A. Financial evaluation of projects with artificial intelligence in the banking sector. *Finansovyi biznes = Financial Business*. 2022;(5):122–125. (In Russ.).
14. Agarwal N., Moehring A., Rajpurkar P., Salz T. Combining human expertise with artificial intelligence: Experimental evidence from radiology. NBER Working Paper. 2023;(31422). URL: [https://www.nber.org/system/files/working\\_papers/w31422/w31422.pdf](https://www.nber.org/system/files/working_papers/w31422/w31422.pdf) (дата обращения: 09.08.2023).
15. Помулев А.А. Искусственный интеллект как объект стоимостной оценки. *Имущественные отношения в Российской Федерации*. 2022;(6):42–56. DOI: 10.24412/2072-4098-2022-6249-42-56  
Pomulev A.A. Artificial intelligence as an object of valuation. *Imushchestvennye otnosheniya v Rossiiskoi Federatsii = Property Relations in the Russian Federation*. 2022;(6):42–56. (In Russ.). DOI: 10.24412/2072-4098-2022-6249-42-56

16. Никитин Н. А. Вероятностные методы учета модельных рисков при оценке инвестиций в технологии искусственного интеллекта. *Инновационное развитие экономики*. 2023;(2):123–134. DOI: 10.51832/2223798420232123  
Nikitin N.A. Probabilistic methods for accounting model risks in assessing investments in artificial intelligence technologies. *Innovatsionnoe razvitie ekonomiki = Innovative Development of Economy*. 2023;(2):123–134. (In Russ.). DOI: 10.51832/2223798420232123
17. Kiela D., Bartolo M., Yixin Nie Y., et al. Dynabench: Rethinking benchmarking in NLP. In: Proc. 2021 conf. North American Chapter of the Association for Computational Linguistics: Human language technologies. Stroudsburg, PA: Association for Computational Linguistics; 2021:4110–4124. URL: <https://aclanthology.org/2021.naacl-main.324.pdf> (дата обращения: 11.08.2023).
18. Bubeck S., Chandrasekaran V., Eldan R., et al. Sparks of artificial general intelligence: Early experiments with GPT-4. Cornell University. arXiv:2303.12712 [cs.CL]. 2023. DOI: 10.48550/arXiv.2303.12712
19. Sevilla J., Heim L., Ho A., et al. Compute trends across three eras of machine learning. In: 2022 Int. joint conf. on neural networks (IJCNN). (Padua, July 18–23, 2022). Piscataway, NJ: IEEE; 2022. DOI: 10.1109/IJCNN 55064.2022.9891914
20. Waisberg E., Ong J., Masalkhi M., et al. GPT-4: A new era of artificial intelligence in medicine. *Irish Journal of Medical Science*. 2023;192(6):3197–3200. DOI: 10.1007/s11845-023-03377-8
21. Rivas P., Zhao L. Marketing with ChatGPT: Navigating the ethical terrain of GPT-based chatbot technology. *AI*. 2023;4(2):375–384. DOI: 10.3390/ai4020019
22. Dohmke T., Iansiti M., Richards G. Sea change in software development: Economic and productivity analysis of the AI-powered developer lifecycle. New Hyde Park, NY: Keystone; 2023. 30 p. URL: <https://github.blog/wp-content/uploads/2023/06/Sea-Change-in-Software-Dev.pdf>
23. Chung H.W., Hou L., Longpre S., et al. Scaling instruction-finetuned language models. *Journal of Machine Learning Research*. 2024;25:1–53. URL: <https://www.jmlr.org/papers/volume25/23-0870/23-0870.pdf>

## ИНФОРМАЦИЯ ОБ АВТОРЕ / ABOUT THE AUTHOR



**Никита Александрович Никитин** — аспирант департамента финансового и инвестиционного менеджмента, факультет «Высшая школа управления», Финансовый университет, Москва, Россия

**Nikita A. Nikitin** — postgraduate student, Department of Financial and Investment Management, Graduate School of Management, Financial University, Moscow, Russia  
<https://orcid.org/0000-0003-1217-713X>  
[nikitanrus@mail.ru](mailto:nikitanrus@mail.ru)

*Конфликт интересов: автор заявляет об отсутствии конфликта интересов.*

*Conflicts of Interest Statement: The author has no conflicts of interest to declare.*

*Статья поступила в редакцию 16.08.2023; после рецензирования 18.09.2023; принята к публикации 25.09.2023.*

*Авторы прочитали и одобрили окончательный вариант рукописи.*

*The article was submitted on 16.08.2023; revised on 18.09.2023 and accepted for publication on 25.09.2023.*

*The authors read and approved the final version of the manuscript.*