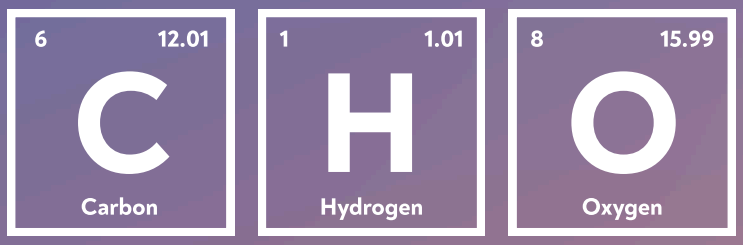
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО

известия студенческой науки



Выпуск 1

Том 2

Текстовое электронное издание

Санкт-Петербург
2025

УДК 004, 063, 065, 504
ББК 20, 32, 40
Известия студенческой науки. Выпуск 1. Том 3. Текстовое электронное издание (3053 Мб). СПб.: Университет ИТМО. 2025. 58 с.

Издание содержит результаты результатов научно-исследовательской деятельности обучающихся вузов и молодых ученых.
Мероприятие проводится в рамках реализации гранта в форме субсидий из федерального бюджета образовательным организациям высшего образования на реализацию мероприятий, направленных на поддержку студенческих научных сообществ (Соглашение № 075-15-2025-536 от 30 мая 2025 г.).

Под общей редакцией кандидата физико-математических наук, заместителя начальника департамента научных исследований и разработок Белашенкова Н.Р.

ISBN 978-5-7577-0740-2
ISBN 978-5-7577-0742-6 (Том 3)

Минимальные системные требования:
Компьютер: процессор x86 с тактовой частотой 500 МГц и выше; ОЗУ 512 Мб; 8Мб на жёстком диске; видеокарта SVGA 1280x1024 High Color (32 bit); привод CD-ROM.
Операционная система: Windows XP/7/8 и выше.
Программное обеспечение: Adobe Acrobate Reader версии 6 и старше.

ИТМО (Санкт-Петербург) — национальный исследовательский университет, научно-образовательная корпорация. Альма-матер победителей международных соревнований по программированию. Приоритетные направления: IT и искусственный интеллект, фотоника, робототехника, квантовые коммуникации, трансляционная медицина, Life Sciences, Art&Science, Science Communication.
Лидер федеральной программы «Приоритет-2030», в рамках которой реализовывается программа «Университет открытого кода». С 2022 ИТМО работает в рамках новой модели развития — научно-образовательной корпорации. В ее основе академическая свобода, поддержка начинаний студентов и сотрудников, распределенная система управления, приверженность открытому коду, бизнес-подходы к организации работы. Образование в университете основано на выборе индивидуальной траектории для каждого студента.
ИТМО пять лет подряд — в сотне лучших в области Automation & Control (кибернетика) Шанхайского рейтинга. По версии SuperJob занимает первое место в Петербурге и второе в России по уровню зарплат выпускников в сфере IT. Университет в топе международных рейтингов среди российских вузов. Входит в топ-5 российских университетов по качеству приема на бюджетные места. Рекордсмен по поступлению олимпиадников в Петербурге. С 2019 года ИТМО самостоятельно присуждает ученые степени кандидата и доктора наук.

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

ПРИКЛАДНАЯ АНАЛИТИКА

УДК 004.056.5:004.8

ИСПОЛЬЗОВАНИЕ LLM-МОДЕЛЕЙ В КИБЕРПРЕСТУПЛЕНИЯХ

Большаков Г.В.[1] (магистрант), **Лемешко А.В.**[1] (магистрант), **Рогаткин Н.А.**[1] (магистрант)
Научный руководитель – кандидат технических наук Бутылкина К.Д.[1]
[1]Университет ИТМО
zhora.vb@gmail.com

Аннотация
В статье рассматривается использование больших языковых моделей в киберпреступной деятельности и анализируется эволюция их применения – от инструментов автоматизации генерации вредоносного кода и фишинговых сообщений до интеграции моделей в архитектуру вредоносных систем. На основе анализа исследований «Cybercriminals Starting to Use ChatGPT» и «Prompts as Code & Embedded Keys: The Hunt for LLM-Enabled Malware» выделены две фазы развития киберпреступного применения больших языковых моделей. Начальная фаза – характеризующаяся использованием публичных генеративных моделей для создания атакующих инструментов, и последующая, в которой большие языковые модели становятся активным компонентом вредоносных программ. Особое внимание уделяется концепции «Prompts-as-Code», предполагающей использование промптов как программных инструкций, управляющих поведением моделей, и формирующей новую категорию угроз – семантические атаки. Подчёркивается, что традиционные методы защиты, основанные на статическом и сигнатурном анализе, оказываются неэффективными против динамически генерируемых атак. В работе предлагаются направления развития когнитивно-ориентированных систем киберзащиты, способных анализировать семантику промптов, а также подчеркивается необходимость междисциплинарного подхода, объединяющего лингвистику, искусственный интеллект и кибербезопасность для противодействия новым видам ИИ-угроз.

Ключевые слова
Большие языковые модели, киберпреступность, промт-инженерия, кибербезопасность.

Развитие больших языковых моделей в последние годы стало одной из ключевых точек роста как в сфере искусственного интеллекта, так и в области кибербезопасности. Технологии, изначально разработанные для автоматизации интеллектуальных задач, создание текста и поддержки пользователей, постепенно приобретают двойственный характер – становятся инструментом не только созидания, но и разрушения. На пересечении этих направлений возникает феномен использования LLM-моделей в киберпреступной деятельности. Исследования показали, что киберпреступники начали применять языковые модели для автоматизации создания вредоносного кода, генерации фишинговых писем и обхода защитных систем. При этом выявляется не только рост интереса к LLM в даркнете, но и формирование новых подходов к созданию вредоносных программ – с применением промптов как элементов кода, встраиванием ключей доступа и модульной архитектуры генеративных систем. Известные ранее методы противодействия ИИ-угрозам, основанные на статическом анализе или сигнатурном распознавании, оказываются малоэффективны против динамически генерируемых атак, формируемых LLM-моделями в реальном времени. Их недостаток заключается в неспособности предсказать поведение модели, особенно в случаях, когда вредоносные инструкции скрыты внутри промптов или обучающих данных. Новизна современных атак состоит в использовании самой модели как среды исполнения вредоносного поведения. В рамках исследования выполнен анализ двух независимых источников, освещающих становление и эволюцию киберпреступного применения LLM. Особое внимание уделяется переходу от простых примеров использования ChatGPT для создания кода до сложных сценариев встраивания LLM во вредоносные системы, где промпты рассматриваются как кодовые инструкции и элементы инфраструктуры атак.

Задача анализа заключается в выявлении тенденций и особенностей злоупотребления языковыми моделями в контексте киберпреступной экосистемы. Цель: определить механизмы, с помощью которых злоумышленники используют LLM для автоматизации атак, обхода защитных мер и повышения эффективности вредоносных инструментов. Исследование «Cybercriminals Starting to Use ChatGPT» описывает начальный этап проникновения генеративных моделей в даркнет-сообщество. Публикации инструкций о применении ChatGPT для написания вредоносного кода, генерации фишинговых писем, обхода детекторов и автоматизации социальной инженерии. Преступники использовали публично доступные версии моделей, обходя ограничения путём модификации запросов или через API. В результате образовалась новая категория участников даркнета, не обладающих глубокими техническими знаниями, но способных создавать функциональный вредоносный код с помощью ИИ. Дальнейшая эволюция зафиксирована в исследовании «Prompts as Code & Embedded Keys: The Hunt for LLM-Enabled Malware». В нём авторы показывают переход от использования LLM как внешнего инструмента к интеграции их в архитектуру вредоносных систем. В таких сценариях промпты и ключи доступа становятся частью исходного кода, а модели применяются для выполнения логики программ. Появился термин «Prompts-as-Code» – подход, при котором инструкции для модели записываются как кодовые модули, управляющие поведением LLM внутри вредоносной инфраструктуры [1, 2].

В исследовании «Cybercriminals Starting to Use ChatGPT» описаны примеры первых зафиксированных случаев применения ChatGPT в киберпреступных целях. На форумах даркнета появились обсуждения, в которых участники демонстрировали, как генеративные модели могут создавать шифровальщики, инструменты для кражи данных и вредоносные скрипты на Python и C++. При этом вектор атак сместился от традиционного написания вредоносного кода вручную к полуавтоматическому созданию и оптимизации скриптов через последовательность промптов. Одной из ключевых находок стало выявление злоумышленников, использующих ChatGPT для обхода языковых фильтров: преступники формулировали запросы в косвенной форме, создавая «безопасные» с точки зрения модели задания, результат которых можно было легко адаптировать для атаки. Например, запрос «создай программу для шифрования файлов в целях резервного копирования» фактически генерировал основу для ransomware-модуля. Исследование также отмечает появление специализированных «помощников» – ботов на базе LLM, созданных для продажи в даркнете. Эти инструменты позволяли автоматизировать написание фишинговых писем на разных языках, создавать поддельные сайты и тексты социальной инженерии. Таким образом, LLM-технологии начали выполнять роль посредника между человеческим преступным замыслом и технической реализацией атак [1].

Исследование «Prompts as Code & Embedded Keys: The Hunt for LLM-Enabled Malware» фиксирует следующую фазу развития – превращение LLM из инструмента генерации контента в активный компонент вредоносных систем. Исследователи выявили новые образцы программ, в которых промпты и ключи доступа встроены непосредственно в код. Этот подход, получивший название «Prompts-as-Code», позволил создавать динамически адаптируемых вредоносных агентов. В подобных системах промпт фактически выполняет роль сценария, задающего поведение модели при взаимодействии с внешней средой. Например, злоумышленник может вшить в код инструкцию, определяющую, какие данные считать конфиденциальными и как их передавать на сервер. При этом обнаружить вредоносную активность становится крайне трудно, поскольку вредоносная логика распределена между кодом и самой моделью, выполняющей инструкцию. Исследование описывает конкретные случаи обнаружения вредоносных образцов, в которых ключи API OpenAI или других LLM-сервисов хранились в зашифрованном виде, а логика работы модели определялась динамически. Этот механизм позволял программам вызывать LLM для генерации вредоносных строк, анализа украденных данных или даже формирования новых промптов в ответ на действия защитных систем. Следовательно, появляется новая парадигма – интеграция LLM как функционального элемента вредоносной архитектуры, в отличие от предыдущей стадии, когда модели использовались только для создания атак [2].

Также в исследовании «Prompts as Code & Embedded Keys: The Hunt for LLM-Enabled Malware» приводится анализ механизма внедрения ключей и промптов в код, описывающий характерные паттерны, такие как скрытые строки JSON, зашифрованные конфигурации и использование нестандартных библиотек для вызова LLM API. Исследователи отмечают, что обнаружение таких образцов требует не только анализа кода, но и изучения сетевых запросов, контекста вызовов и паттернов обращения к внешним сервисам. Одним из выводов исследования является то, что традиционные системы анализа кода не способны оценивать логику взаимодействия между программой и моделью. Это создает необходимость в новых методах киберзащиты, основанных на динамическом поведении программ и анализе семантики промптов. Кроме того, исследование фиксирует рост числа ситуаций, когда киберпреступники маскируют вредоносные промпты под легитимные инструкции. Например, строка, содержащая «оптимизируй текст отчёта», могла фактически запускать скрытую функцию по сбору системной информации. Такие механизмы делают промпт-инжекцию новым направлением в эволюции вредоносных программ, аналогичным по значимости эксплойтам нулевого дня [2].

Сопоставление данных двух исследований позволяет выявить четкую траекторию эволюции киберпреступного применения LLM:

1. Экспериментальные примеры написания вредоносного кода, фишинг-генераторов и автоматизированных инструментов социальной инженерии.
2. Формирование нового класса атак с использованием промптов как управляющих инструкций, встроенных ключей и динамической генерации вредоносных компонентов.

На основе анализа этих двух фаз можно утверждать, что киберпреступное использование LLM развивается не линейно, а по принципу усложняющейся интеграции. Если в первой фазе модели выступали как внешние помощники для упрощения рутинных задач, таких как генерации текстов, кода, шаблонов атак, то во второй фазе они становятся частью атакующей инфраструктуры. Это превращает ИИ в соучастника киберпреступления, поскольку модель не просто выполняет инструкции, а участвует в их интерпретации и адаптации. Одним из ключевых вызовов становится проблема интерпретации промптов и идентификации их намерений. Современные системы защиты не могут различить промпт, направленный на легитимную задачу, например, автоматизацию отчётов, и промпт, маскирующий вредоносные действия, такое как скрытый сбор данных. Это открывает возможности для создания «гибридных» программ, в которых вредоносная логика зависит не от исходного кода, а от контекста взаимодействия модели и пользователя. Подобные программы могут менять своё поведение в зависимости от ответов модели, что делает их крайне сложными для анализа и обнаружения. Кроме того, в исследованиях отмечено, что преступники начинают применять LLM для построения целых экосистем киберугроз – с генерацией фишинговых сообщений, автоматическим анализом ответов жертв, управлением ботами и обновлением сценариев атак в реальном времени. Такие действия демонстрируют переход к автономным кибероперациям, где человеческое участие минимально, а большая часть решений принимается на уровне генеративных моделей. Таким образом, LLM перестают быть инструментом и становятся субъектом в структуре атаки.

Особое значение имеет концепция «Prompts-as-Code». Она знаменует переход от кода как фиксированного набора инструкций к коду как лингвистическому процессу. Промпт становится программой, а модель – интерпретатором. Это создаёт новую категорию уязвимостей. Промпт-инжекция превращается в аналог SQL-инъекций, но с гораздо более сложной логикой и контекстной зависимостью. Такие атаки способны модифицировать не только поведение модели, но и направление её обучения или реакции на внешние данные. С точки зрения защиты, это приводит к переосмыслению всей парадигмы кибербезопасности. Традиционные методы анализа и детектирования, такие как сигнатурный анализ, эвристики и статическая проверка, теряют эффективность. Для противодействия подобным угрозам необходимы когнитивно-ориентированные системы, способные анализировать не только синтаксис промпта, но и его семантику – то есть смысл, интенцию и потенциальную цель запроса. Ключевым направлением становится разработка «семантических фильтров», выявляющих скрытые вредоносные намерения даже в грамматически корректных и безопасных текстах. Следовательно, LLM-

преступность формирует новую эру в киберугрозах, где оружием становятся не эксплойты и вирусы, а языковые конструкции и модели поведения. Это требует от исследователей объединения знаний из областей лингвистики, ИИ, когнитивных наук и анализа данных. Только междисциплинарный подход позволит создать эффективные инструменты обнаружения, основанные не на коде, а на понимании смысла взаимодействия человека и машины.

Проведённый анализ двух исследований показал, что использование больших языковых моделей в киберпреступной деятельности прошло путь от случайных экспериментов до формирования новой парадигмы вредоносных архитектур. На ранних этапах LLM применялись как вспомогательный инструмент для автоматизации генерации кода и контента, что делало возможным быстрый старт даже для малоквалифицированных злоумышленников. Однако в последующих разработках модели стали неотъемлемой частью вредоносной инфраструктуры, выполняя функции анализа, управления и адаптации. Современные тенденции указывают, что LLM всё чаще используются для реализации концепции «автономных атакующих агентов» – систем, способных самостоятельно исследовать среду, собирать информацию, корректировать цели и обучаться на основе ответов защитных механизмов, что поднимает вопрос не только технической, но и этической ответственности разработчиков моделей, поскольку сама архитектура ИИ может быть переориентирована на деструктивные цели.

Ключевые результаты анализа можно представить следующим образом:

1. LLM-модели трансформируются в активные компоненты киберугроз, способные выполнять сложные логические операции, адаптироваться под условия среды и изменять своё поведение в зависимости от ответов системы.
2. Метод «Prompts-as-Code» создаёт новую технику внедрения вредоносных инструкций, где язык становится основой кода, а промпты – носителями вредоносных функций.
3. Формируется новая категория угроз – «семантические атаки», в которых смысл взаимодействия с моделью становится оружием, а цель достигается не через взлом, а через манипуляцию её интерпретацией.
4. Для защиты от подобных атак необходима разработка когнитивных систем анализа, способных выявлять аномалии в коммуникационном поведении программ, определять отклонения в структурах промптов и отслеживать динамические изменения взаимодействий между моделью и пользователем.

Вектор дальнейших исследований должен быть направлен на изучение механизмов интерпретации промптов, создание специализированных систем мониторинга запросов и развитие технологий «объяснимого ИИ» (Explainable AI), которые позволят отслеживать внутренние решения моделей. Без понимания того, почему и как модель пришла к тому или иному выводу, невозможно гарантировать её безопасность. Кроме того, перспективным направлением является разработка «лингвистических щитов» – систем фильтрации и анализа текстового взаимодействия, способных блокировать вредоносные инструкции до их интерпретации моделью. Такие решения могут стать основой для нового уровня защиты, интегрирующего кибербезопасность и лингвистический анализ.

В заключение можно сказать, что киберпреступное использование LLM выходит за рамки традиционного понимания вредоносного ПО. Это не просто технологическая, но и концептуальная революция, изменяющая само представление о природе угроз. Новая эпоха требует новых инструментов, методов анализа и подходов к этическому регулированию ИИ. Лишь комплексный, междисциплинарный подход, объединяющий кибербезопасность, когнитивные науки и искусственный интеллект, способен обеспечить устойчивость цифровой среды перед лицом эволюционирующих угроз.

Литература

1. Check Point Research OpWnAI: Cybercriminals Starting to Use ChatGPT. [Электронный ресурс]. Режим доступа: https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-use-chatgpt/ (Дата обращения 06.11.2025).
2. Alex Delamotte, Vitaly Kamluk, Gabriel Bernadett-Shapiro. Prompts as Code & Embedded Keys: The Hunt for LLM-Enabled Malware. [Электронный ресурс]. Режим доступа: https://www.sentinelone.com/labs/prompts-as-code-embedded-keys-the-hunt-for-llm-enabled-malware/ (Дата обращения 06.11.2025).

УДК 004.738.5:004.8

ИНТЕГРАЦИЯ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ В АРХИТЕКТУРЫ ИНТЕРНЕТА ВЕЩЕЙ: MQTT-ТЕЛЕМЕТРИЯ, EDGE-LLM И MODEL CONTEXT PROTOCOL

Большаков Г.В.[1] (магистрант), **Лемешко А.В.**[1] (магистрант), **Рогаткин Н.А.**[1] (магистрант)
Научный руководитель – кандидат технических наук Бутылкина К.Д.[1]
[1]Университет ИТМО
zhora.vb@gmail.com

Аннотация
В статье предлагается рассматривать большие языковые модели (LLM) в системах Интернета вещей не как автономный «умный» сервис, а как семантический медиатор, то есть слой, связывающий человеческие цели и высказывания с потоками телеметрии и формально описанными возможностями гетерогенных устройств. На основе трёх источников показано, что MQTT обеспечивает надёжную и масштабируемую транспортную ткань для обмена наблюдениями и командами, периферийное размещение LLM в сочетании с механизмами извлечения знаний (RAG) позволяет достигать приемлемых задержек и естественно-языкового управления, а Model Context Protocol (MCP) задаёт формальный язык инструментов и действий, благодаря которому LLM способна надёжно генерировать последовательности операций в сложных сценариях. Раскрываются ключевые измерения качества такого подхода — корректность семантической интерпретации, компромисс «качество–задержка», ресурсные ограничения и управляемость, — с опорой на эмпирические результаты прототипа «умного дома» и бенчмарк IoT-MCP Bench. Делается вывод, что синтез MQTT, edge-LLM с RAG и MCP формирует практичную концептуальную основу для проектирования LLM-ориентированных IoT-систем, а дальнейшее развитие связано с формализацией и верификацией поведения LLM в IoT-контексте, расширением бенчмарков под реальные эксплуатационные условия и интеграцией специализированных механизмов безопасности.

Ключевые слова
Интернет вещей, большие языковые модели, MQTT, MCP.

Современные системы с IOT давно перестали быть набором разрозненных датчиков и исполнительных механизмов. Теперь это сложные киберфизические экосистемы, в которых одновременно сосуществуют десятки протоколов, сотни типов устройств и разнообразные прикладные сервисы. На этом фоне большие языковые модели (Large Language Models, LLM) становятся универсальным интерфейсом в цифровой среде, они способны понимать естественный язык, удерживать сложный контекст, планировать последовательности действий и объяснять свои решения. Стоит рассмотреть вопрос, как использовать LLM не просто как ещё один сервис в IoT-архитектуре, а как посредника между человеком и цифровым миром устройств.

Существующие работы демонстрируют разные грани этой проблемы. В исследовании «Talk with the Things: Integrating LLMs into IoT» показано, как LLM, размещённые на периферийных вычислительных узлах, позволяют организовать естественно-языковое управление умным домом, комбинируя модели с механизмами извлечения знаний и локальными хранилищами данных. Статья «Integrating MQTT with AI and LLMs in IoT: Best Practices and Future Perspectives» концентрируется на транспортном уровне и показывает, что MQTT способен служить высокопроизводительной «нервной системой» для AI- и LLM-приложений в масштабных IoT-развёртываниях. Наконец, работа «IoT-MCP: Bridging LLMs and IoT Systems Through Model Context Protocol» вводит концепцию IoT-MCP, где LLM взаимодействует с устройствами через формально описанные инструменты и стандартизированный протокол контекста, а качество такого взаимодействия оценивается с помощью специализированного бенчмарка. Все три подхода важны, но каждый из них описывает собственный «слой» проблемы, а именно: периферийное размещение моделей, транспорт данных или стандарты описания действий. В этой статье предлагается рассматривать LLM прежде всего как семантический медиатор в IoT-экосистемах, то есть слой, который связывает цели и высказывания человека, телеметрию устройств и формальные возможности инфраструктуры. На основе анализа

упомянутых исследований формулируется постановка задачи, предлагается концепция такого семантического слоя, обсуждаются эмпирические результаты и критически оцениваются преимущества и ограничения подхода [1–3].

Классический IoT разделяется на протоколы передачи данных и форматы сообщений. Датчики собирают значения, исполнительные устройства получают команды, посредники маршрутизируют потоки. Синтаксически всё выглядит хорошо, но человек мыслит в терминах ситуаций и целей, а не в терминах регистров, топиков и payload-структур. Команда «сделать комнату комфортной для сна» естественна для пользователя, но совершенно неочевидна для лампы и системы вентиляции, пока между ними не появится слой, способный интерпретировать такие цели. Именно этот разрыв между «человеческим» уровнем описания и низкоуровневыми интерфейсами устройств можно обозначить как семантический. Традиционные подходы к его преодолению опираются на заранее заданные сценарии и правила: конструкций вида «если температура ниже X и время больше Y, включить отопление» в промышленной системе могут быть тысячи. При росте масштаба и разнообразия устройств такие сценарии становятся трудно управляемыми, слабо адаптивными и малопригодными для динамически меняющихся условий. Если рассматривать LLM в роли семантического слоя, то перед ним встаёт целый набор требований. Во-первых, модель должна интерпретировать высказывания пользователя как цели и ограничения, а не как набор ключевых слов. Во-вторых, она должна учитывать текущий контекст, то есть состояние устройств, историю взаимодействий, предпочтения пользователя, внешние условия. В-третьих, необходимо уметь работать с неопределённостью и неполнотой данных, в ряде случаев уместно запросить уточнение вместо того, чтобы действовать по умолчанию. Наконец, семантический слой не может быть «всесильным», то есть он должен подчиняться политикам безопасности и правилам доступа, объяснять свои решения и допускать внешнее управление. В такой задаче протоколы вроде MQTT и MCP оказываются не просто «транспортом», а инфраструктурой, которая несёт на себе часть семантической нагрузки. MQTT задаёт структуру и надёжность потока наблюдений, а Model Context Protocol задаёт язык описания того, что вообще умеют подключённые к системе устройства. LLM, располагаясь поверх этой инфраструктуры, получает возможность оперировать не сырыми данными, а уже структурированными сущностями: наблюдениями, действиями, инструментами.

Предлагаемая концепция исходит из того, что LLM следует рассматривать не как «умное приложение» рядом с IoT, а как слой семантического посредничества. В такой роли модель принимает на себя интерпретацию человеческих запросов, сопоставляет их с возможностями устройств и состоянием среды, а затем формирует последовательности действий, которые реализуются с помощью стандартных механизмов IoT-платформы.

В работе «Talk with the Things» этот подход представлен на примере умного дома. Пользователь формулирует запросы в естественной форме, от простых инструкций до сложных сценариев, учитывающих комфорт, безопасность и привычки. LLM, размещённая на периферийном узле, опирается на механизмы Retrieval-Augmented Generation, чтобы подмешивать к запросу актуальные данные о состоянии системы и внешние знания. В результате одна и та же фраза «сделай поярче» может интерпретироваться по-разному в зависимости от комнаты, времени суток и уже включённых устройств. Семантический медиатор в таком понимании выполняет несколько связанных функций. Он выделяет и формализует намерение пользователя, выбирает, какие подсистемы и устройства релевантны текущей цели, сопоставляет абстрактные понятия вроде «комфорт» или «экономия» с конкретными управляющими параметрами. Кроме того, он принимает во внимание динамический контекст: прошлые команды, состояние устройств, внешние условия, что особенно заметно в сценариях, когда модель должна не только реагировать на прямые запросы, но и сама инициировать действия на основе наблюдаемого поведения системы. Семантическое описание возможностей устройств в этом контексте становится ключевым элементом. Исследование «IoT-MCP» показывает, что, если представить каждую функцию устройства в виде инструмента с чётко заданной сигнатурой, параметрами, ограничениями и текстовым описанием, LLM способна с высокой надёжностью комбинировать такие инструменты для решения сложных задач. Бенчмарк IoT-MCP Bench, включающий сотни базовых и более тысячи комплексных задач,

демонстрирует стопроцентную успешность генерации корректных последовательностей действий при работе с десятками типов сенсоров и контроллеров. Это важный аргумент в пользу того, что семантический медиатор может быть не только концептуальным, но и практически реализуемым компонентом. Также взаимодействие с потоками телеметрии, реализуемое через MQTT, позволяет медиатору видеть «пульс» системы. Статья «Integrating MQTT with AI and LLMs in IoT» подчёркивает, что MQTT-брокеры при правильной конфигурации обеспечивают малую задержку и высокую пропускную способность, обслуживая огромные объёмы сообщений. Для семантического слоя это означает, что он может работать не только в режиме «запрос–ответ», но и в режиме постоянного анализа событий, выделяя значимые паттерны и инициируя действия даже без явного запроса пользователя [1–3].

Рассматривая LLM в роли семантического медиатора, важно не просто описать функции, но и обсудить измерения качества, по которым такой подход может быть оценён. Прежде всего, это корректность интерпретации, временные характеристики, ресурсные требования и степень управляемости системы. Корректность семантической интерпретации в существенной мере зависит от того, насколько строго и ясно описаны возможности устройств и доступные действия. Результаты IoT-MCP Bench показали, что при использовании формальных описаний инструментов Model Context Protocol LLM способна безошибочно генерировать последовательности действий для широкого набора задач. Это означает, что значительная часть риска ошибок переносится с самой модели на качество спецификаций: чем прозрачнее и формальнее описаны возможности устройств, тем меньше пространство для неверных решений. В этом смысле семантический медиатор становится совместным продуктом LLM и разработчика спецификаций, а не автономным «чёрным ящиком». Временные характеристики тесно связаны с тем, где и в каком виде исполняется модель. В прототипе умного дома из «Talk with the Things» показано, что уменьшение размера LLM сокращает задержку ответа, но снижает качество интерпретации сложных запросов. Более крупные модели, наоборот, обеспечивают более точное понимание, но требуют либо более мощного оборудования, либо вынесения вычислений в облако, что увеличивает сетевую задержку. Семантический медиатор в реальной системе неизбежно будет находить баланс между этими крайностями, различая сценарии, где критически важна скорость, например, реакция на аварийные события, и ситуации, где допустим более точный, но медленный анализ [1, 2].

Ресурсная модель дополняет эту картину. IoT-устройства чаще всего сильно ограничены в вычислительных возможностях и памяти, и это отчётливо видно в экспериментах, где микроконтроллеры выполняют только минимально необходимую логику, а взаимодействие с LLM организуется через внешние узлы. В таком варианте семантический медиатор логически единый, но физически распределённый, то есть часть его функций реализована на edge-шлюзах, часть – в облаке, а на устройствах присутствуют лишь тонкие адаптеры и инструменты. Транспортный слой при этом остаётся относительно лёгким, ведь MQTT обеспечивает доставку данных и команд, не навязывая сложную логику на уровне устройств. Управляемость и объяснимость становятся, пожалуй, самым тонким аспектом. С одной стороны, LLM обладает естественной способностью к генерации текстовых объяснений, и исследования показывают, что эту способность можно использовать для коммуникации с пользователем, например, пояснять, какие действия предприняты и почему. С другой стороны, объяснение не всегда совпадает с формальной интерпретацией поведения модели. Для инженерной практики важно не только то, что медиатор «умеет объяснять», но и то, что его поведение поддаётся внешнему контролю, например, через политики безопасности, правила допуска к инструментам, механизмы принудительных ограничений. В этом смысле IoT-MCP, задающий строгую структуру инструментов и параметров, служит противовесом чисто генеративной природе LLM [1–3].

Преимущество рассматриваемого подхода проявляется прежде всего в унификации взаимодействия с гетерогенной IoT-инфраструктурой. Вместо того чтобы разрабатывать отдельные панели, приложения и скрипты для каждой подсистемы, можно использовать единый язык взаимодействия – естественный язык, завернутый в семантический слой, который знает о возможностях устройств и способах их комбинирования. Комбинация MQTT в качестве универсальной шины сообщений, MCP как языка описания действий и LLM как интерпретатора

позволяет подключать новые устройства в значительной степени за счёт добавления их описаний, а не за счёт переписывания логики целиком. Не менее важной является адаптивность. LLM естественным образом подстраивается под стиль общения и поведение конкретного пользователя, учитывает историю взаимодействия и может постепенно уточнять представление о его предпочтениях. Прототип умного дома из «Talk with the Things» иллюстрирует, как модель, работающая совместно с механизмами извлечения знаний, использует накопленный контекст для более тонкой настройки поведения системы. Это делает семантический медиатор не просто универсальным интерфейсом, но и механизмом персонализации, что особенно важно в бытовых и сервисных сценариях [1, 3].

Тем не менее подход не лишён серьёзных ограничений. Галлюцинации и ошибки LLM остаются фундаментальной проблемой: в отсутствие строгих спецификаций и ограничений модель может генерировать действия, не соответствующие намерениям пользователя или политике безопасности. Необходимость защищать не только транспортный уровень MQTT, но и сам контекст LLM – от подмены данных, prompt-инъекций, отравления обучающих выборок – добавляет к классическим IoT-угрозам ещё один слой риска. Даже при наличии MCP-описаний и бенчмарков вроде IoT-MCP Bench вопрос доверия к системе сохраняет свою остроту. Дополнительный барьер связан с организационной и инженерной стороной. Для массового внедрения семантических медиаторов нужны процессы стандартизации описаний устройств, инструменты проектирования и тестирования MCP-интерфейсов, методики оценки качества и безопасности решений. Работы «Talk with the Things», «IoT-MCP» и «Integrating MQTT with AI and LLMs in IoT» демонстрируют, что технологическая основа уже существует, но её превращение в промышленный стандарт требует координации множества участников, таких как производителей устройств, разработчиков платформ, операторов инфраструктуры и регуляторов [2, 3].

Таким образом, в статье предложено рассматривать большие языковые модели в системах Интернета вещей преимущественно как семантический медиатор – слой, связывающий человеческие цели и высказывания с потоками телеметрии и формально описанными возможностями устройств. На основе трёх исследований показано, что MQTT обеспечивает устойчивый транспортный фундамент для взаимодействия IoT-устройств и AI-сервисов, периферийное размещение LLM и использование механизмов извлечения знаний позволяют организовать естественно-языковое управление умными средами при приемлемых задержках и ресурсных затратах, а Model Context Protocol предоставляет формальный язык описания инструментов, с помощью которого LLM может генерировать последовательности действий в сложных сценариях. Тем самым концепция семантического медиатора оказывается не абстрактной метафорой, а логическим синтезом существующих технологических решений.

Дальнейшее развитие этого направления связано, во-первых, с формализацией и верификацией поведения LLM в IoT-контексте, во-вторых, с развитием бенчмарков, отражающих реальные эксплуатационные условия и длительную работу системы, и, в-третьих, с интеграцией механизмов безопасности, учитывающих не только традиционные угрозы для IoT-инфраструктуры, но и специфические риски генеративных моделей. Одновременно требуется выстраивание инженерных практик, в рамках которых семантический медиатор будет не заменой человеческому контролю, а инструментом, расширяющим возможности проектирования и эксплуатации сложных IoT-экосистем.

Литература

1. Kalita A. Talk with the Things: Integrating LLMs into IoT Networks. [Электронный ресурс]. Режим доступа: https://arxiv.org/pdf/2507.17865v1 (Дата обращения 11.11.2025).
2. Yang N., Lyu G., Ma M., Lu Y., Li Y., Gao Z., Ye H., Zhang J., Chen T., Chen Y. IoT-MCP: Bridging LLMs and IoT Systems Through Model Context Protocol. [Электронный ресурс]. Режим доступа: https://arxiv.org/pdf/2510.01260 (Дата обращения 11.11.2025).
3. Ji B. Integrating MQTT with AI and LLMs in IoT: Best Practices and Future Perspectives. [Электронный ресурс]. Режим доступа: https://www.emqx.com/en/blog/integrating-mqtt-with-ai-and-llms (Дата обращения 11.11.2025).

УДК 004.8:711

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ЗАДАЧАХ ГРАДОСТРОИТЕЛЬНОГО МОДЕЛИРОВАНИЯ: СОВРЕМЕННЫЕ ПЛАТФОРМЫ И ПРАКТИКИ ПРИМЕНЕНИЯ

Исаев Ш.М.[1] (студент), **Шиндина П.Д.**[1] (студент)
Научный руководитель – старший преподаватель Занина А.Д.[1]
[1]СПБПУ
e-mail: isaev.list@list.ru

Аннотация
В данной статье представлен обзор инструментов проектирования, применяемых в градостроительстве с использованием искусственного интеллекта (ИИ), с помощью которых моделируются концепции «умных городов» (цифровых двойников) и оптимизируются инфраструктурные решения.
Ключевые слова
Искусственный интеллект, градостроительное моделирование, цифровой двойник, умный город, генеративный дизайн, городская инфраструктура.

Современные города становятся всё более сложными системами, требующими учёта множества факторов: плотности населения, транспортных потоков, экологии, инженерных сетей и др. Для анализа и оптимизации таких систем традиционные методы оказываются недостаточно эффективными. Искусственный интеллект (ИИ) и технологии больших данных (Big Data) открывают новые возможности в городской аналитике и моделировании. В частности, концепция цифрового двойника города предполагает создание виртуальной копии городской среды, объединяющей геопространственные данные, информацию о зданиях, транспорте, инженерных сетях, климате и прочее [1, 2]. Этот подход позволяет «эффективно моделировать развитие городской территории, работу систем ЖКХ, транспорта, безопасности, влияние климата» [3]. По мнению экспертов, цифровой двойник становится своеобразной виртуальной платформой для оптимизации процессов управления городом: он позволяет анализировать прошлое, моделировать настоящее и прогнозировать будущее, что способствует улучшению качества жизни и рациональному использованию ресурсов [1, 3]. Во многих мировых мегаполисах (Сингапур, Бостон, Хельсинки и др.) уже создаются подобные цифровые модели для поддержки планирования и обеспечения устойчивого развития [4, 3]. В России концепция цифровых городов только развивается (пилотные проекты в Москве и Санкт-Петербурге), и она рассматривается как инструмент для экспериментального тестирования урбанистических решений и снижения рисков при застройке [4, 5].

ИИ-алгоритмы уже находят практическое применение в самых разных градостроительных задачах. Так, отмечено, что ИИ-платформы и сервисы используются для анализа территорий, моделирования благоустройства, разработки концепций застройки, проектирования зданий, генерации презентационных материалов и референсов, проверки проектов на соответствие нормам, а также для оптимизации управления умными городами (например, адаптивного управления светофорами) [6]. Классическим примером является создание «Цифрового нормоконтроля» – сервиса Московского «Мосстройинформа», позволяющего автоматически проверять проектную документацию на соответствие нормативам, что сокращает трудозатраты и исключает человеческие ошибки [7]. Другой кейс – платформа rTim (Фонд «ДОМ.РФ»), генерирующая концепции жилой застройки с учётом 140 параметров, экономической оценки и очередности строительства [8]. В Москве также внедрены системы «Умный перекрёсток», которые с помощью ИИ адаптивно управляют транспортными потоками, повышая пропускную способность улиц и снижая заторы [9].

Современные облачные платформы на базе ИИ позволяют в несколько раз ускорить процесс проектирования кварталов и жилых массивов. Наиболее значимые разработки в данной области демонстрируют потенциал генеративного проектирования, которое интегрирует моделирование архитектурных конфигураций с комплексной оценкой экологических, экономических и социальных факторов (рис. 1).

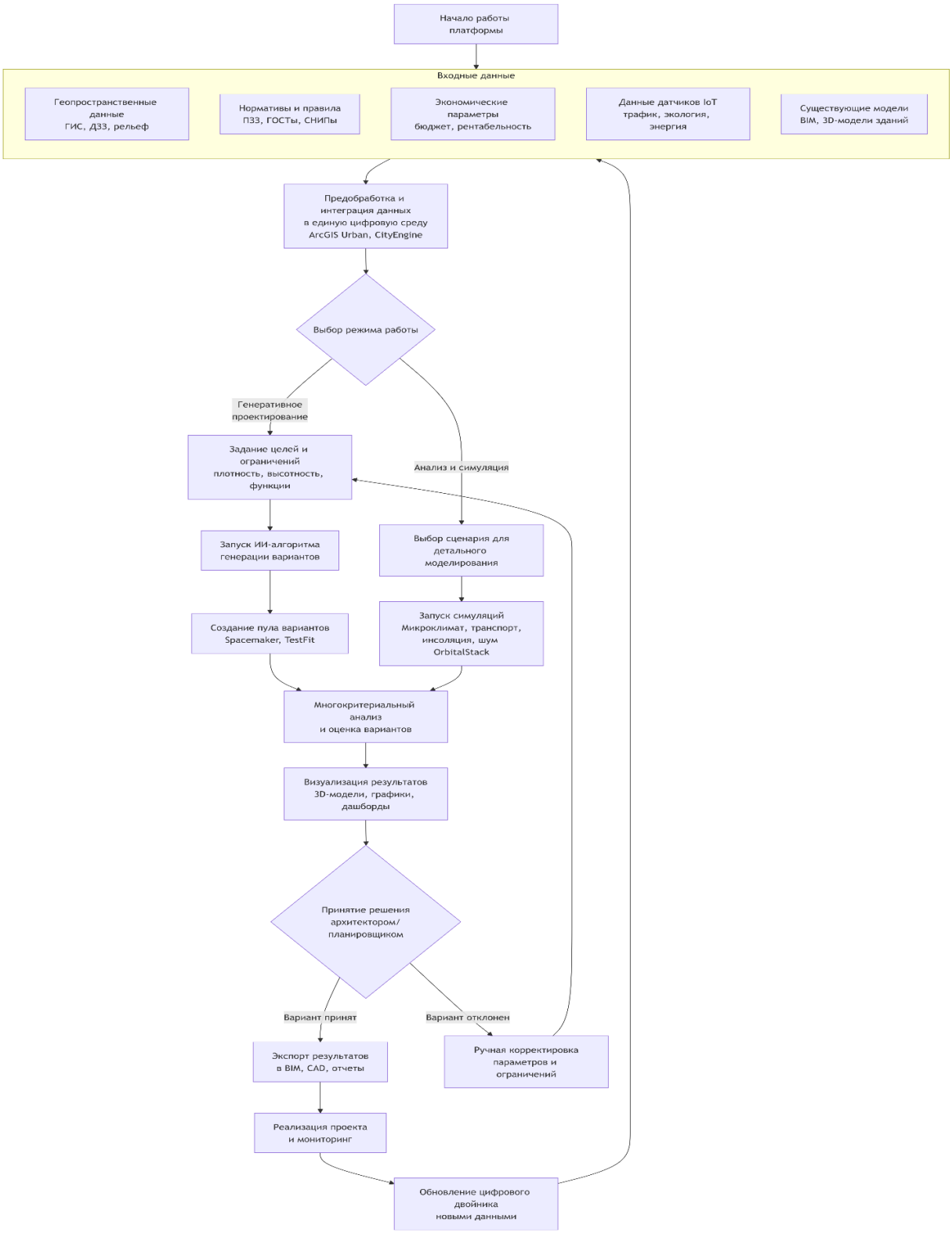


Рис. 1. Блок-схема применения ИИ для генеративного проектирования градостроительных решений

Так, сервис Spacemaker (компания Autodesk) автоматически генерирует тысячи альтернативных вариантов застройки на заданном участке [10, 11]. Система строит 3D-макеты зданий и оценивает более 100 параметров, что дает возможность быстро фильтровать малоперспективные решения [10, 11].

Spacemaker поддерживает циклическую оптимизацию: он автоматически подстраивает планировки, предлагая архитектурные конфигурации с учётом ограничений (плотность застройки, фазы строительства, показатели экологии). По отчетам компании, переход на ИИ-

инструмент позволяет архитекторам завершать предпроектные исследования и технико-экономические обоснования за несколько часов вместо недель [11].

Аналогичный подход применяет платформа TestFit. В ней используется параметрическая генерация многоквартирных домов: пользователь задаёт габариты участка, требования к количеству парковочных мест и квартире, а система за секунды формирует полный архитектурный макет с расчётом площади и финансами (рис. 2) [12]. Это позволяет мгновенно увидеть, насколько прибыльным может быть проект, и отказаться от неэффективных вариантов. Другие инструменты (например, OrbitalStack) фокусируются на микроклимате: они сравнивают набор архитектурных решений и рассчитывают воздействие ветра, инсоляции и тени на городскую среду, что важно для «зеленого» дизайна кварталов [13].

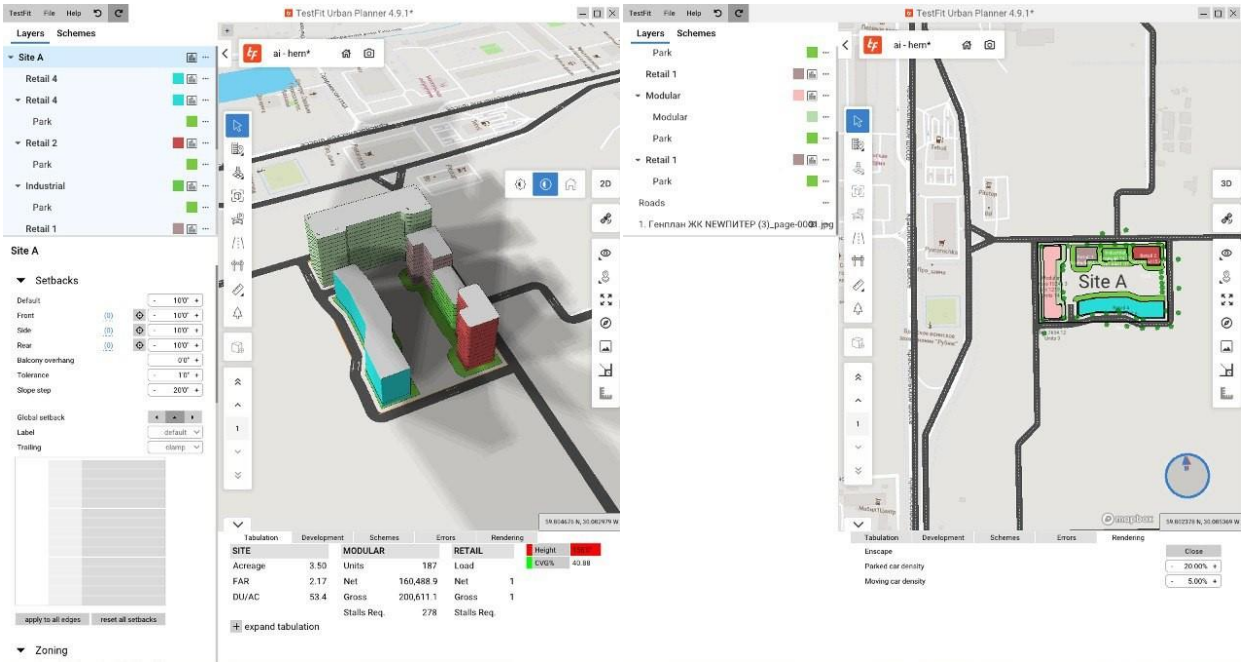


Рис. 2. Процесс разработки планировочного решения группы домов в Testfit

Параллельно с развитием генеративного проектирования цифровые технологии создают новые возможности для моделирования городских территорий через применение цифровых двойников. Как отмечают Hurtado и Gomez, цифровой двойник города строится на основе множественных слоёв данных, каждый из которых описывает разные аспекты среды [14]. На нижнем уровне находится рельеф и география (локация, высота рельефа), далее – сведения о зданиях (BIM-модели, зонирование), за ними –транспортная и инженерная инфраструктура (дороги, метро, линии электропередач и др.), затем – информация о мобильности (потоки транспорта, расписания), и, наконец, слой «умного города» (данные IoT-датчиков, социальные опросы). На основе этой многоуровневой модели могут выполняться симуляции и «что-если»анализы: например, как изменится загрузка дорог или очистка воздуха при вводе новых кварталов или транспорте. Такой подход фактически создаёт виртуальную тренировочную среду для планировщиков.

Цифровой двойник города (рис. 3) позволяет управлять городскими системами в соответствии со стратегией развития, а также «прогнозировать последствия предлагаемых изменений и является инструментом по поиску оптимальных решений». В частности, в Москве внедряется единая платформа цифрового двойника, которая включает подробную 3D-модель столицы с геоданными, характеристиками зданий и инженерных сетей. Такая модель дает возможность в реальном времени изменять параметры городского плана (местоположение новых объектов, конфигурацию дорог и т.д.) и автоматически оценивать влияние на транспортную нагрузку, потребность в социальных объектах (школы, поликлиники) и другие показатели [15]. По мнению экспертов, применение цифрового двойника «позволит эффективно моделировать развитие городской территории», ускоряя инфраструктурные

решения и улучшая качество жизни жителей; при этом такие системы удобны не только для стратегического планирования, но и для оперативного локального управления отдельными объектами и небольшими участками в реальном времени. Примером внедрения технологии на институциональном уровне является проект Санкт-Петербургского политехнического университета «МетаКампус Политех». Он представляет собой цифровую экосистему университета, основанную на точной информационной модели кампуса, включающей здания, территории и инженерную инфраструктуру. Проект был реализован в рамках стратегии цифровой трансформации СПбПУ и объединил передовые технологии информационного и имитационного моделирования, а также анализа данных Такая модель позволяет не только автоматизировать управление инфраструктурой кампуса, но и создавать цифровые сервисы для студентов, преподавателей и администрации, расширяя функциональность образовательной среды [16].

Рис. 3. Многоуровневая схема цифрового двойника умного города

Систему оценки ИИ-инструментов для градостроительного проектирования целесообразно выстраивать по нескольким направлениям: функциональная направленность (генеративный дизайн, сценарное моделирование, микроклиматический анализ), интеграция данных (BIM/GIS, большие данные), глубина анализа параметров (учёт инсоляции, трафика, экологии, нормативов), скорость генерации сценариев и степень автоматизации, а также экономическая оценка и поддержка ТЭО. Важное значение имеют масштаб применения (здание, квартал, город), учёт экологических и социальных факторов, возможности коллективной работы и визуализации, а также гибкость и инновационность платформ, включая интеграцию в цифровые двойники. Таким образом, сопоставление этих критериев с традиционными методами показывает, что применение ИИ-программ значительно выгоднее: они обеспечивают сокращение сроков предпроектного анализа, автоматизацию рутинных расчётов и более комплексную оценку градостроительных решений уже на ранних стадиях проектирования.

На основании таблицы можно отметить, что большинство современных ИИ-инструментов для градостроительства являются облачными сервисами с акцентом на интеграцию данных и генерацию 3D-сцен. Они позволяют быстро (в несколько кликов) исследовать сотни сценариев застройки, чего невозможно в традиционном ручном подходе. Ключевым преимуществом служит автоматизация рутинных расчётов и визуализация итогов: результаты подаются в удобной форме для проектировщика, что упрощает обсуждение с инвесторами и органами власти [15].

Таблица

Описание инструментов ИИ для градостроительного проектирования

Инструмент	Описание	Ключевые возможности
ArcGIS Urban (Esri) и CityEngine (Esri)	Платформа и процедурный 3D-генератор Esri для градостроительных сценариев. Позволяет создавать веб-проекты города на основе геоданных	Поддержка сценарного анализа и цифрового двойника города; интеграция BIM и GIS-данных; оперативная корректировка архитектурных параметров (планы этажей, фасадные стили, высота зданий); совместная работа через ArcGIS Online
Autodesk Spacemaker	Облачный AI-сервис генеративного дизайна (приобретён Autodesk). Автоматически генерирует варианты застройки кварталов	Быстрая генерация 3D-моделей зданий и кварталов; анализ >100 показателей среды (инсоляция, шум, ветер, озеленение); учёт нормативов и фаз строительства; оценка жилой и коммерческой площади; интеграция с BIM/AutoCAD
TestFit	Платформа для быстрой оценки проектных решений в недвижимости. Использует параметрическую генерацию зданий	Мгновенное построение макетов жилых домов по заданным требованиям (машиноместа, площади квартир); расчёт финансовых показателей (доходность, рентабельность); визуализация вариантов и конфигураций; облегчение технико-экономических обоснований
OrbitalStack	Веб-сервис для микроклиматического анализа застройки	Сравнение вариантов планировки с учётом атмосферы и ветра; CFD-моделирование солнечного освещения и тени; оценка комфорта и безопасности при разных конфигурациях зданий; ориентирован на высокопродуктивные и «зелёные» проекты
Другие платформы	(примерно)	CityFormLab (MIT) – модели с учётом урбанистических правил; ArcGIS GeoAnalytics – большие данные и AI в анализе ГИС; InfraWorks (Autodesk) – концептуальное проектирование инфраструктуры; и др.

Наиболее развитые системы (Spacemaker, ArcGIS Urban) используют методы оптимизации и машинного обучения для улучшения результатов. К тому же они совмещают в себе аналитические и симуляционные функции: например, создание макета квартала тут же сопровождается вычислением трафика, освещённости и других показателей. Это позволяет оценивать не только архитектуру, но и инженерные последствия проекта ещё на предпроектной стадии.

Результаты

Искусственный интеллект открывает новые горизонты для градостроительного проектирования, трансформируя подходы к планированию и управлению городскими пространствами. Интеграция ИИ в градостроительные процессы не только ускоряет разработку проектов, но и позволяет учитывать множество факторов, что ведет к созданию более гармоничных и устойчивых городских сред. Однако, чтобы максимально использовать потенциал этих технологий, необходимо преодолеть ряд вызовов.

Ключевыми шагами на этом пути являются развитие инфраструктуры данных, обучение специалистов и внедрение пилотных проектов, которые позволят протестировать новые подходы в реальных условиях. Нормативная поддержка и междисциплинарное сотрудничество также играют важную роль в создании эффективной экосистемы для применения ИИ в градостроительстве.

В конечном итоге, успешная реализация этих рекомендаций создаст основу для формирования «умных городов», где технологии и человеческий фактор будут работать в гармонии, обеспечивая комфорт и устойчивое развитие для будущих поколений. Градостроительство, основанное на ИИ, станет не просто инструментом, а настоящим катализатором технологий.

Литература

1. Иванов С.А., Никольская К.Ю., Радченко Г.И., Соколинский Л.Б., Цымблер М.Л. Концепция построения цифрового двойника города // Вестник ЮУрГУ. Серия «Вычислительная математика и информатика». 2020. №4(9). [Электронный ресурс]. Режим доступа: https://sciup.org/koncepcija-postroenija-cifrovogo-dvojnika-goroda-147234283 (Дата обращения 19.09.2025).
2. Мухачева А.В., Иванова О.Е., Парфенов А.А. Цифровые двойники городов: возможности и преимущества. [Электронный ресурс]. Режим доступа: https://consensus.app/papers/цифровые-двойники-городов-возможности-и-преимущества-ухачёва-парфенов/dd013037c650591f836aa0534bf60754/ (Дата обращения 19.09.2025).
3. Hurtado P., Gomez A. Smart City Digital Twins Are a New Tool for Scenario Planning. [Электронный ресурс]. Режим доступа: https://planning.org/planning/2021/spring/smart-city-digital-twins-are-a-new-tool-for-scenario-planning/ (Дата обращения 19.09.2025).
4. Pérez-Martínez I., Martínez-Rojas M., Soto-Hidalgo J.M. Automated Urban Planning: An Open-Source Plugin for Enhanced Design and Optimization Based on Generative Design // SSRN. – 2024. – Предварительная версия статьи. [Электронный ресурс]. Режим доступа: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5017025 (Дата обращения 19.09.2025).
5. Подольская Е.С. Методы и ГИС-инструменты машинного обучения в задачах проектирования объектов транспортной инфраструктуры // Вопросы лесной науки. 2023. №. 4. С. 33–47.
6. Повх Е.В. Десять цифровых двойников городов. [Электронный ресурс]. Режим доступа: https://realty.rbc.ru/news/5e297b079a79478024d54ff6 (Дата обращения 19.09.2025).
7. ИТП «Град». Использование искусственного интеллекта в урбанистике и градостроительстве. [Электронный ресурс]. Режим доступа: https://itpgrad.ru/education/articles/ispolzovanie-iskusstvennogo-intellekta-v-urbanistike-i-gradostroitelstve/ (Дата обращения 19.09.2025).
8. Васильева Н. Будущее умных городов: технологии для мегаполисов. [Электронный ресурс]. – Режим доступа: https://rg.ru/2025/09/02/budushchee-umnyh-gorodov-tehnologii-dlia-megapolisov.html (Дата обращения 19.09.2025).
9. РБК Кавказ. В Нальчике запустили систему цифрового двойника города. [Электронный ресурс]. Режим доступа: https://kavkaz.rbc.ru/kavkaz/freenews/62cbf8ea9a794730dcbd11ba (Дата обращения 19.09.2025).
10. Esri. ArcGIS CityEngine – 3D GIS for Urban Design. [Электронный ресурс]. Режим доступа: https://www.esri.com/en-us/arcgis/products/arcgis-cityengine/overview (Дата обращения 19.09.2025).
11. Autodesk. Spacemaker – облачная платформа для проектирования застройки. [Электронный ресурс]. Режим доступа: https://www.autodesk.com/products/spacemaker/overview (Дата обращения 19.09.2025).
12. Slashdot.org. Compare ArcGIS Urban vs Spacemaker. [Электронный ресурс]. Режим доступа: https://slashdot.org/software/comparison/ArcGIS-Urban-vs-Spacemaker/ (Дата обращения 19.09.2025).
13. Slashdot.org. TestFit Reviews. [Электронный ресурс]. Режим доступа: https://slashdot.org/software/p/TestFit/ (Дата обращения 19.09.2025).
14. 2GIS Dev. Кейс: цифровой двойник города Нальчик. [Электронный ресурс]. Режим доступа: https://dev.2gis.ru/cases/digitaltwin_nalchik_government (Дата обращения 19.09.2025).
15. Shaip. Тренировочные данные AI для геопространственных проектов. [Электронный ресурс]. Режим доступа: https://ru.shaip.com/solutions/ai-training-data-for-geospatial-projects/ (Дата обращения 19.09.2025).
16. «МетаКампус Политех». [Электронный ресурс]. Режим доступа: https://research.spbstu.ru/projects/metakampus_politekh/ (Дата обращения 19.09.2025).

УДК 004.75:004.8

ПРИМЕНЕНИЕ ДЕЦЕНТРАЛИЗОВАННЫХ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ ДЛЯ LLM-МОДЕЛЕЙ

Большаков Г.В.[1] (магистрант), **Лемешко А.В.**[1] (магистрант), **Рогаткин Н.А.**[1] (магистрант)
Научный руководитель – кандидат технических наук Бутылкина К.Д.[1]
[1]Университет ИТМО
zhora.vb@gmail.com

Аннотация
В статье рассматриваются предпосылки и перспективы применения децентрализованных облачных вычислений для поддержки жизненного цикла больших языковых моделей (LLM). На основе анализа вертикальной инфраструктуры Phoenix, платформы Theta EdgeCloud и концепции децентрализованных реестров доверия SingularityNET/Privado ID показывается, что переход от централизованных облаков к DePIN-подходам обусловлен ростом требований к качеству, устойчивости и приватности LLM-сервисов. Обосновывается, что вертикальные стеки наподобие Phoenix, сочетающие вычислительный слой, сеть узлов и агентные приложения, позволяют формировать более эластичную и наблюдаемую инфраструктуру инференса, в которой метрики качества и отказоустойчивости задаются реальными прикладными сценариями. Отдельное внимание уделяется вопросам доверия и приватности. Рассматривается роль децентрализованных идентификаторов (DID) и проверяемых учётных данных (VC) в построении реестров ИИ-агентов и узлов, а также возможность их интеграции с децентрализованными вычислительными платформами. Проводится сопоставление подходов Phoenix и Theta EdgeCloud с точки зрения интересов компаний и конечных пользователей, подчёркиваются преимущества гибридных схем, сочетающих децентрализованные и традиционные облачные решения. Делается вывод о том, что децентрализованные облачные вычисления для LLM формируют новую инфраструктурную парадигму, в которой технические, организационно-экономические и доверительные аспекты развития ИИ-систем рассматриваются как единое целое, а также намечаются направления дальнейших исследований, связанных с институционализацией реестров доверия и практической оценкой устойчивости таких платформ в промышленных условиях. Статья предлагает концептуальную рамку для последующего моделирования и эмпирического анализа децентрализованных ИИ-инфраструктур.

Ключевые слова
Большие языковые модели, децентрализованные облачные вычисления, приватность данных.

Бурное развитие больших языковых моделей стало испытанием для классической парадигмы облачных вычислений. LLM одновременно требовательны к вычислительным ресурсам и чувствительны к задержкам и стабильности, а в ряде сценариев – к приватности данных и прозрачности цепочки доверия. Централизованные облака исторически решали проблему масштабирования ИИ-нагрузок, но сегодня их ограничения становятся все заметнее, например, высокая стоимость долгоживущих GPU-кластеров, зависимость от нескольких глобальных провайдеров, «чёрный ящик» инфраструктуры, в котором сложно проверить, что именно происходило с пользовательскими данными и как обеспечивается устойчивость сервисов. На этом фоне формируется слой децентрализованных облачных вычислений для ИИ, то есть сетей, где множество независимых узлов предоставляет вычислительные ресурсы, а координация, учёт и стимулы выносятся в протоколы с элементами Web3 и DePIN (decentralized physical infrastructure networks). Phoenix позиционируется как вертикальная децентрализованная ИИ-инфраструктура: от слоя вычислений до нативных приложений и агентных платформ, включая PhoenixONE и другие сервисы. Theta EdgeCloud предлагает близкую по духу концепцию, ориентированную на научные и индустриальные проекты, в том числе вокруг LLM в университетской среде. Параллельно развивается направление децентрализованных реестров доверия для ИИ-агентов, где связка SingularityNET и Privado ID демонстрирует, как можно соединить блокчейн-инфраструктуру с DID и проверяемыми учётными данными (VC), чтобы агентные системы были не только масштабируемыми, но и поддающимися проверке [3–5].

Известные централизованные решения дают зрелые SLA и богатый инструментальный стек, но их слабые места для LLM-сценариев очевидны. Во-первых, пользователь вынужден доверять закрытой инфраструктуре, и внутренняя политика обработки данных остаётся для него непрозрачной. Во-вторых, масштабирование происходит в рамках ограниченного парка дата-центров, что приводит к узким местам по задержкам и устойчивости при росте нагрузки. Наконец, централизованные провайдеры не всегда выгодно выстраивают модель взаимодействия с независимыми разработчиками и владельцами аппаратуры, то есть им сложнее непосредственно участвовать в экономике ИИ-вычислений. Предлагаемый децентрализованный подход в общих чертах выглядит так: сети узлов, способных выполнять ИИ-задачи, объединяются в вычислительный слой, поверх которого разворачиваются сервисы LLM-инференса, дообучения и агентные приложения. Координация задач, учёт и вознаграждение узлов реализуются через токенизированные схемы, а доверие между участниками усиливается с помощью открытых реестров, репутации и криптографически подтверждаемых аттестатов. Цель статьи – рассмотреть, как такие архитектуры работают применительно к LLM-моделям, какие у них есть преимущества и ограничения с точки зрения качества, устойчивости и приватности, и как они воспринимаются компаниями и конечными пользователями.

Большие языковые модели предъявляют к инфраструктуре специфический набор требований. С одной стороны, им нужен массивный и гибко масштабируемый GPU-ресурс. С другой – часть сценариев, особенно в области интерактивных приложений, критична к задержкам и предсказуемости времени ответа. Наконец, для LLM, работающих с чувствительными данными, такими как финансы, медицина, корпоративная аналитика, на первый план выходят приватность и прозрачность того, где и как исполняются запросы. Классические облака изначально строились как централизованные платформы общего назначения, с сильным акцентом на капиталоёмкие дата-центры и многослойный стек абстракций. Такой метод давал удобство, но порождал и проблемы, такие как сложность доступа к инфраструктурным деталям для внешних участников, монопольную ценовую власть крупных провайдеров и зависимость бизнес-рисков от решений нескольких корпораций. DePIN-подходы предлагают иную модель: инфраструктура формируется снизу вверх за счёт множества владельцев физических ресурсов (от дата-центров до отдельных «майнеров» и энтузиастов), а единая логика работы системы задаётся протоколом [1, 3].

Phoenix в отчёте JDI описывается именно как вертикальная децентрализованная ИИ-инфраструктура, которая использует DePIN-модель для агрегации и управления вычислительными ресурсами, но при этом стремится контролировать качество и устойчивость через специализированный слой оркестрации и токеномики. Computation Layer, позже развиваемый в рамках обновления SkyNet, представляет собой Web3-платформу для масштабирования ИИ-моделей, где пользователи могут разворачивать и масштабировать модели через панель управления и SDK, а задачи распределяются по сети ИИ-узлов. По сути, речь идёт об «эластичном» вычислительном слое, где не дата-центр диктует правила игры, а приложение и протокол совместно определяют, какие узлы и по каким критериям будут задействованы. Theta EdgeCloud, в свою очередь, подчёркивает ограничения традиционной облачной модели для ИИ-нагрузок и предлагает распределённую платформу, в которой акцент делается на снижение стоимости, лучшую масштабируемость и уменьшение задержек за счёт географической распределённости и использования ресурсов «на краю» сети. В примерах, связанных с исследованиями LLM в Стэнфорде, авторы показывают, что распределённая инфраструктура позволяет университетам и лабораториям выбирать более гибкую модель доступа к вычислениям, не будучи привязанными к условиям одного-двух hyperscaler-ов. Вывод из сопоставления этих подходов заключается в том, что децентрализованная ИИ-инфраструктура рассматривается не как идеологический проект, а как прагматичный ответ на три запроса: обеспечивать высокое качество работы LLM, оставаться устойчивой под нагрузкой и давать пользователю больше контроля над приватностью и доверительной цепочкой [1, 3, 5].

Phoenix в описании JDI и в собственных материалах предстает не просто как сеть узлов, а как вертикальный стек, то есть вычислительный слой SkyNet/Computation Layer, сеть ИИ-

нод, поверх которой строятся прикладные сервисы, например, AlphaNet для торговых стратегий и другие приложения, и, наконец, пользовательские интерфейсы и агентные платформы. С точки зрения качества LLM этот подход важен по нескольким причинам. Во-первых, Phoenix изначально выстраивает инфраструктуру вокруг действительно работающих приложений: AlphaNet, трейдинговые модели, а позже и агентные сценарии PhoenixONE. Это означает, что критерии качества не абстрактны, так как их задают реальные пользователи, которые чувствительны к стабильности работы моделей, корректности ответов и адекватности управления риском. Во-вторых, благодаря вертикальной интеграции компания контролирует не только «сырой» GPU-ресурс, но и жизненный цикл моделей, включая развёртывание, обновления, настройку инфраструктуры хранения и маршрутизации запросов, что позволяет аккуратнее управлять версиями моделей, учитывать нагрузочные профили и планировать расширение сети узлов. Отдельного анализа заслуживает архитектура SkyNet, описанная на официальном сайте как важное обновление вычислительного слоя. Помимо продвижения к более полной децентрализации вычислительных ресурсов, в ней подчёркивается несколько ключевых идей. Во-первых, речь идёт о «эластичном» ИИ-слое, который должен масштабироваться вслед за реальным спросом, а не за счёт искусственного наращивания числа узлов. Во-вторых, SkyNet использует ad-hoc архитектуру задач, когда крупные ИИ-нагрузки могут дробиться на более мелкие подзадачи, такой подход потенциально снижает фрагментацию ресурсов и позволяет более ровно загружать сеть. А также в документах описывается интеллектуальная система маршрутизации, которая учитывает характеристики узлов и задач при распределении нагрузки, что напрямую влияет на качество и устойчивость сервисов [1, 2, 5].

Важно, что Phoenix не ограничивается чисто «инфраструктурной» роль, так как в текстах о ростовых точках подчёркивается стратегический переход к «agentic applications» – приложениям, в которых LLM-модели и агенты выполняют более сложные цепочки действий, в том числе в финансовом контуре. Появление платформы PhoenixONE в этом контексте выглядит логичным, ведь если у вас уже есть эластичная сеть вычислений и набор моделей, то следующий шаг – предоставить пользователю интерфейс, в котором он может сценарно использовать агентные возможности, не задумываясь о том, на каком конкретно узле исполняется задача. С точки зрения пользователей такой подход даёт более цельный опыт, ведь они взаимодействуют не с «облаком GPU», а с набором инструментов для глубоких исследований, торговли или анализа, где вычислительная сложность скрыта за интерфейсами и политиками качества. При этом сам факт существования децентрализованной сети узлов и понятной токенизированной экономики делает систему более прозрачной, чем традиционные «чёрные ящики» облаков [2, 5].

Метод от Theta EdgeCloud служит хорошей альтернативой Phoenix, представляя другую линию эволюции децентрализованных ИИ-вычислений. Авторы начинают с описания проблем традиционных облаков, таких как высокая стоимость долгосрочных контрактов на GPU-мощности, ограниченная масштабируемость, особенно для исследовательских групп, и значительные задержки при доступе к ресурсам, сконцентрированным в нескольких географических регионах. Далее предлагается «децентрализованная платформа для вычислений в области искусственного интеллекта», которая должна одновременно снизить затраты и повысить доступность ИИ-ресурсов. Theta EdgeCloud делает особый акцент на сценариях, связанных с LLM в академической среде, приводя пример проекта Стэнфорда по исследованиям больших языковых моделей. Здесь видно важное отличие от Phoenix: если в нём строит вертикаль вокруг собственной экосистемы приложений и токеномики, то Theta скорее позиционирует себя как инфраструктурный слой, открытый для широкого круга партнёров – от университетов до коммерческих компаний. При этом логика схожа, ведь узлы с вычислительными ресурсами объединяются в сеть, поверх которой действует слой координации и учёта [3].

С точки зрения ключевых для нас метрик Theta EdgeCloud подчёркивает несколько аспектов. Во-первых, устойчивость достигается за счёт георасределённости и способности перекидывать нагрузки между регионами, что особенно важно для длительных исследований,

где простой инфраструктуры равен потере времени и средств. Во-вторых, качество работы LLM обеспечивается комбинацией технических и организационных механизмов: стандартами по типам поддерживаемых ресурсов, мониторингом работы узлов и прозрачностью взаимодействия с конечными пользователями. А также на уровне приватности Theta обсуждает, как в рамках децентрализованной платформы можно уменьшать риски утечек, в том числе за счёт локализации данных и более явного описания того, где физически исполняются задачи. По сути, Theta EdgeCloud демонстрирует, что идея децентрализованных ИИ-вычислений не ограничивается отдельными Web3-экосистемами. Она постепенно становится более универсальной, как для коммерческих игроков, так и для академии, так и для разработчиков LLM это способ получить более гибкий и управляемый доступ к вычислительному слою, не отказываясь при этом от требований к качеству и стабильности [3].

Отдельного внимания заслуживает вопрос доверия и приватности, подробно рассматриваемый в статье о SingularityNET, Google Cloud и видении Changpeng Zhao. Авторы исходят из того, что по мере роста числа ИИ-агентов и децентрализованных сервисов взаимодействие между ними становится всё более сложным и непрозрачным. Если каждый агент может вызывать другие агенты, передавать им данные и поручения, то возникает потребность в инфраструктуре доверия, так как нужно понимать, кому именно вы передаёте данные и как можно проверить полномочия и репутацию контрагента. Предлагаемое решение – создание децентрализованного реестра доверия для ИИ-агентов, основанного на децентрализованных идентификаторах (DID) и проверяемых учётных данных (VC). В такой системе каждый агент и сервис получает устойчивый идентификатор и набор аттестатов, выданных доверенными эмитентами. Такие аттестаты могут отражать технические характеристики, то есть тип и класс оборудования, а также уровень безопасности, организационные свойства, такие как принадлежность к определённой компании или проекту, а также репутационные показатели, например, история выполнения задач, уровень отказов и нарушений. В контексте децентрализованных облачных вычислений для LLM это имеет несколько важнейших следствий. Во-первых, появляется возможность более чётко управлять приватностью, то есть пользователь или компания может задать политику, которая разрешает передавать данные только агентам с определёнными атрибутами, например с сертифицированной инфраструктурой или ограниченной географией размещения. Во-вторых, формируется основа для устойчивости на уровне доверия: если узел или агент систематически нарушает ожидания, его репутация ухудшается, и протоколы маршрутизации задач начинают его обходить [4].

В случае Phoenix идея доверия не формализована в виде отдельного реестра агентов, как в примере SingularityNET/Privado ID, но логика схожа, ведь Computation Layer и SkyNet опираются на токенизированные стимулы, учёт задач и репутацию узлов, а развитие экосистемы предполагает всё более тесную связку между экономическими и доверительными механизмами. В перспективе использование DID/VC-подхода может дать дополнительный уровень прозрачности, то есть не только понимать, какой узел выполняет задачу, но и иметь криптографически подтверждённое описание его свойств и истории работы. Таким образом, связка децентрализованных вычислений и децентрализованных реестров доверия задаёт новую норму для работы LLM-систем. Пользователи получают не только распределённую инфраструктуру, но и инструменты контроля за тем, кто именно имеет доступ к их данным и каковы гарантии устойчивости инфраструктуры [1, 5].

Если посмотреть на Phoenix и Theta глазами компаний, внедряющих или использующих LLM-модели, то на первый план выходят не только технические детали, но и организационные и экономические контуры. В случае Phoenix для компаний важна именно вертикальность решения. JDI подчёркивают, что Phoenix создаётся как «полный стек» децентрализованного ИИ: от вычислительной инфраструктуры до прикладных сервисов и интерфейсов, что снижает интеграционные издержки, то есть корпоративному пользователю проще иметь дело с одной платформой, которая отвечает и за качество LLM-инференса, и за устойчивость вычислительных цепочек, и за приватность данных, чем самостоятельно собирать элементы из разных поставщиков. Сайт Phoenix акцентирует внимание на том, что Computation Layer

уже используется реальными приложениями, а SkyNet призван расширить этот слой, сделав его более эластичным и открытым для внешних узлов. Компании, интегрирующиеся с Phoenix, получают возможность выбрать уровень вовлечённости. Кто-то может использовать готовые приложения – например, AlphaNet, не вдаваясь в детали работы LLM и инфраструктуры. Другие могут разворачивать собственные модели на Computation Layer, пользуясь при этом тем, что протокол уже решает задачи маршрутизации, мониторинга и учёта. В обоих случаях метрики качества, устойчивости и приватности переходят из разряда абстракций в набор наблюдаемых показателей: задержки, стабильность ответов, прозрачность использования данных. Theta EdgeCloud, ориентируясь на научные и инновационные проекты, делает упор на гибкость и экономическую рациональность. Для университетов и исследовательских подразделений компаний ключевой вопрос – как обеспечить себе доступ к LLM-ресурсам, не превращаясь при этом в инфраструктурного провайдера и не завися на жестких условиях крупных облачных контрактов. Theta предлагает схему, в которой исследователь или организация подключается к децентрализованной платформе и получает доступ к пулу ресурсов, при этом оставаясь в праве выбирать конфигурации, географию и режимы работы. Такой подход особенно полезен для проектов, где нагрузка нерегулярна, то есть, где периоды интенсивного дообучения сменяются периодами более спокойного инференса, и децентрализованная инфраструктура оказывается более гибкой, чем фиксированный набор ресурсов в классическом облаке. Во всех этих случаях компании оказываются в роли не только потребителей, но и потенциальных партнёров. Они могут вносить в сеть собственные узлы, участвовать в токеномике, формировать спрос на новые сервисы. Это усиливает устойчивость системы ведь чем больше в ней независимых центров интересов, тем сложнее одному участнику навязать остальным свои правила, а значит, тем выше устойчивость по отношению к внешним и внутренним шокам [1, 3, 5].

С точки зрения конечного пользователя LLM-сервис оценивается прежде всего по качеству ответов и предсказуемости поведения. Если модель выдаёт корректные и полезные результаты, делает это быстро и не «ломается» в пиковые моменты, то детали инфраструктуры уходят на второй план. Однако именно децентрализованность позволяет эту предсказуемость обеспечить в условиях, когда нагрузки растут, а сценарии использования усложняются. Phoenix в своих материалах подчёркивает, что ориентируется на реальных пользователей, в том числе в социальных средах и трейдинге. Это означает, что платформа вынуждена постоянно балансировать между качеством и устойчивостью. В этом контексте эластичный вычислительный слой SkyNet и сеть PhoenixNode становятся не только технологической новинкой, но и инструментом обеспечения стабильной пользовательской метрики «работает или не работает». Возможность увеличивать и перераспределять вычислительные ресурсы в ответ на рост спроса становится критически важной [2, 5].

Для разработчиков и исследователей ключевым преимуществом децентрализованных платформ становится возможность выбирать уровень абстракции. Кто-то может использовать готовые API LLM и не думать о том, где именно интернет-запросы попадают на физические узлы. Другие захотят настроить параметры, такие как географию исполнения, использование определённых классов оборудования, режимы логирования и обезличивания данных. В традиционных облаках такой уровень контроля либо недоступен, либо обходится очень дорого; децентрализованные платформы строят это в свой функционал по умолчанию. Приватность для пользователя рассматривается с точки зрения возможности работать с LLM, не передавая данные в единый глобальный центр, чья политика обработки информации мало кому понятна. В Theta EdgeCloud обсуждаются сценарии, когда критичные данные остаются ближе к источнику – в локальных или региональных узлах. В SingularityNET/Privado ID акцент делается на том, что пользователи и разработчики смогут выбирать агентов, которым доверяют, на основе открытых аттестатов и репутации. В связке с децентрализованными вычислениями это даёт новый тип UX, где пользователь видит не только результат, но и, при желании, цепочку инфраструктуры и агентов, через которую этот результат был получен [1, 3, 4, 5].

Несмотря на описанные преимущества, децентрализованные облачные вычисления для LLM далеки от завершённого решения. Анализ источников показывает несколько существенных ограничений.

1. Децентрализованные сети по определению гетерогенны, то есть узлы различаются по классу оборудования, стабильности соединения и доступности. В Theta EdgeCloud подчёркивается, что традиционные проблемы облаков – задержки и стоимость – никуда не исчезают, а просто принимают другую форму. Протокол должен уметь рационально распределять задачи, учитывая эти различия, иначе качество и устойчивость LLM-сервисов будут страдать. Phoenix отвечает на этот вызов, выстраивая сложные системы маршрутизации и учёта в SkyNet, но детальные механизмы все ещё развиваются [1, 3, 5].
2. Безопасность и приватность остаются открытыми задачами, так как децентрализация сама по себе не гарантирует защищённости данных. Если пользователь передает чувствительную информацию в агентную систему, которая затем делегирует её децентрализованной сети узлов, возникает множество потенциальных пунктов утечки. Статья о SingularityNET и Privado ID предлагает направление решения в виде DID/VC-реестров доверия, но практическая имплементация таких схем в массовых, высоконагруженных LLM-сервисах ещё впереди. Здесь децентрализованная инфраструктура нуждается в сочетании программных и аппаратных решений – от надёжного шифрования и сегментации до возможного использования доверенных исполнительных сред [4].
3. Экономическая устойчивость таких систем требует осторожности. Phoenix в своих материалах подчёркивает, что не стремится создавать сеть с избыточным числом узлов, а, наоборот, связывает масштаб платформы с реальным спросом. Это реакция на опыт многих DePIN- и майнинговых проектов, где в погоне за «хешрейтом» или количеством устройств забывали о реальной полезности. Однако это же означает, что эластичность системы имеет границы, то есть при внезапном всплеске интереса к LLM-сервисам сеть должна быть готова масштабироваться, не разрушая собственную экономику [5].
4. Децентрализация усложняет управление и ответственность, если в централизованном облаке за сбой и утечку отвечает конкретная компания, то в децентрализованной сети ответственность размазана между множеством участников. Реестры доверия, репутационные системы и регулятивные рамки должны совместно формировать механизмы, в которых пользователь чётко понимает, кому он может предъявить претензии и какие механизмы компенсации предусмотрены. Пока такие рамки находятся в стадии активной проработки, о чём косвенно свидетельствует и обсуждение осторожной токеномики и управления в тексте о SingularityNET [4].
5. Вопрос совместимости с существующей инфраструктурой, ведь большинство компаний уже встроено в экосистемы крупных облаков и не готово одномоментно мигрировать на новые платформы. Поэтому Phoenix, Theta и другие игроки вынуждены предлагать гибридные сценарии, то есть децентрализованные вычисления как дополнение к существующим решениям, а не как их полная замена. Это снижает барьеры входа, но одновременно создаёт дополнительные точки отказа и усложняет архитектуру [1, 3, 5].

Таким образом, децентрализованные облачные вычисления для LLM-моделей – это не столько модный тренд, сколько ответ на реальные вызовы, с которыми сталкиваются разработчики и пользователи современных ИИ-систем. Анализ вертикальной инфраструктуры Phoenix, платформы Theta EdgeCloud и подхода к доверительным реестрам агентов в экосистеме SingularityNET/Privado ID показывает, что формируется новая архитектурная парадигма, в которой вычислительный слой, агентные системы и механизмы доверия тесно переплетены. Децентрализованные сети позволяют агрегировать разнообразные аппаратные ресурсы, гибко масштабировать LLM-нагрузки и встраивать в инфраструктуру более прозрачные и проверяемые механизмы учёта и репутации. С точки зрения ключевых метрик, таких как качества, устойчивости и приватности, эти системы уже сегодня демонстрируют некоторые преимущества. Качество LLM-сервисов повышается за счёт вертикальной интеграции и ориентации на реальные приложения, где качество измеряется не в абстрактных

показателях, а в полезности для пользователей. Устойчивость обеспечивается георасределённостью и возможностью эластичного перераспределения задач, что особенно заметно в примерах для научных и торговых сценариев. Приватность и доверие формируются через инфраструктурные решения – от токенизированных стимулов и журналов задач до децентрализованных реестров DID/VC, которые позволяют формализовать доверие между агентами и узлами [1-5].

В то же время нельзя игнорировать ограничения, такие как гетерогенность сети, сложность обеспечения безопасности данных, необходимость аккуратной экономической модели и проработанных регулятивных рамок. Дальнейшее развитие темы, на наш взгляд, пойдёт по нескольким направлениям. Будут углубляться механизмы доверия и приватности – от более широкого внедрения DID/VC-реестров до использования аппаратных средств для конфиденциальных вычислений. Можно ожидать роста числа гибридных сценариев, где децентрализованные LLM-сервисы дополняют, а не заменяют полностью централизованные облака. Важную роль сыграет опыт крупных экосистем, подобных Phoenix, которые уже сегодня пытаются соединить в одном стеке вычислительный слой, агентные приложения и устойчивую экономику [1,2,5].

В результате децентрализованные облачные вычисления для LLM в ближайшие годы, вероятнее всего, превратятся из экспериментальной истории в важный элемент стандартного набора инструментов для компаний, исследователей и пользователей, которым нужны качественные, устойчивые и приватные ИИ-сервисы. Насколько успешно это произойдёт, будет зависеть от того, удастся ли разработчикам таких платформ удержать баланс между техническими инновациями, экономической целесообразностью и вниманием к интересам конечных пользователей.

Литература

1. JDI Ventures. Phoenix: Decentralized AI Vertical Infrastructure. [Электронный ресурс]. Режим доступа: https://www.binance.com/en/square/post/5147862786513 (Дата обращения 12.11.2025).
2. Phoenix AI Growth Points: Maximizing Value Creation of Social AI-Driven Intelligence. [Электронный ресурс]. Режим доступа: https://medium.com/@Phoenix_AI/phoenix-ai-growth-points-e4f30af67c25 (Дата обращения 12.11.2025).
3. Theta EdgeCloud: децентрализация ИИ-вычислений для Стэнфорда и всех! [Электронный ресурс]. Режим доступа: https://cryptodamus.io/ru/articles/news/theta-edgecloud-decentralizacia-ii-vycislenij-dla-stenforda-i-vseh (Дата обращения 12.11.2025).
4. AI и доверие: как SingularityNET, Google Cloud и CZ формируют децентрализованный AI (Web3). [Электронный ресурс]. Режим доступа: https://cryptodamus.io/ru/articles/news/ai-i-doverie-kak-singularitynet-google-cloud-i-cz-formiruut-decentralizovannyj-ai-web3 (Дата обращения 12.11.2025).
5. Phoenix Global. Официальный сайт проекта Phoenix AI. [Электронный ресурс]. Режим доступа: https://www.phoenix.global (Дата обращения 12.11.2025).

УДК 004.8:159.9

ВЛИЯНИЕ ИИ НА ИНДИВИДА

Рогаткин Н.А.[1] (магистрант), **Большаков Г.В.**[1] (магистрант), **Лемешко А.В.**[1] (магистрант)
Научный руководитель – кандидат технических наук Бутылкина К.Д.[1]
[1]Университет ИТМО
fenekxyz@gmail.com

Аннотация
Статья посвящена анализу влияния искусственного интеллекта и нейросетей на психику и поведение отдельного индивида в повседневной и профессиональной деятельности. На основе современных эмпирических исследований и публицистического материала рассматриваются три ключевых сюжета: когнитивная деградация специалистов при избыточной опоре на алгоритмические подсказки, смещение локуса контроля у пользователей, систематически советующихся с ИИ, а также феномен эмоционального замещения в форме «AI-партнёров». Показано, что когнитивная делегация и привычка перекладывать ответственность на алгоритмы ведут к ослаблению критического мышления, деформации профессиональной идентичности, росту тревожности и зависимости от цифровых сервисов. Особое внимание уделяется рискам десоциализации и трансформации интимной сферы в условиях, когда виртуальные собеседники и сервисы эмоционального общения становятся более предсказуемыми и психологически комфортными, чем реальные люди. Обсуждаются возможные последствия массового распространения гуманоидных роботов-компаньонов и формирование постсоциальной реальности, в которой человек окружён искусственными интеллектами, но лишён устойчивых живых связей. Делается вывод о необходимости развития культуры осознанного и критического использования ИИ, цифровой гигиены и сохранения человеческой автономии при взаимодействии с алгоритмическими системами. Предлагается рассматривать пользователя не как пассивного потребителя цифровых услуг, а как активного соавтора гибридной человеко-машинной системы, несущего этическую и практическую ответственность за результаты её работы. Подчёркивается значимость междисциплинарных исследований, объединяющих психологию, философию, социологию и информатику, для выработки долгосрочных стратегий регулирования и гуманизации ИИ-технологий для человека в целом.

Ключевые слова
Искусственный интеллект, цифровые отношения, локус контроля.

Искусственный интеллект и нейросети к середине двадцатых годов XXI века стали настоящей эпидемией, проникающей во все части жизни человеческого социума. Одни во всю радуются технологиям, становятся ИИ-энтузиастами и стараются использовать нейросети во всех сферах жизни. Другие относятся скептически и боятся, что нейросети отнимут у них работу. Работодатели радуются, что теперь производство станет дешевле, но не всегда их идеи оказываются рабочими. Иногда приходится с позором звонить бывшему работнику и умолять его вернуться на рабочее место. Учёные тоже не остались без дела. Кто-то разрабатывает нейросети и старается их улучшить, а кто-то изучает как они влияют на человека и меняют наше общество. Таких исследований довольно много, какие-то из них рассказывают о положительном влиянии нейросетей, а какие-то об обратном. В этой же статье будет рассмотрена влияние нейросетей на индивида. Абстрактного человека, который каждый день пользуется нейросетями. Как поменяется мир с его стороны, и как он изменится в понимании мира.

Начнём с уже существующих исследований. В статье «The Potential Influence of AI on Population Mental Health» учёные с изрядной долей научного энтузиазма исследуют феномен технологической зависимости и постепенной когнитивной деградации специалистов, слишком доверяющих искусственному интеллекту. На бумаге всё выглядит блестяще: внедрение ИИ должно снижать когнитивную нагрузку, освобождать человека от рутины, повышать точность решений и даже улучшать качество жизни. Однако, как показывает эмпирический анализ, эффект оказывается диаметрально противоположным. Человек перестаёт быть субъектом, превращаясь в пассивного наблюдателя за решениями алгоритма. Рассмотрим конкретный кейс врача, активно использующего нейросетевые модели в своей клинической практике.

Такой врач уже не интерпретирует анализы, а «прогоняет» их через нейросеть; не размышляет над диагнозом, а просто принимает результат алгоритма; не оценивает лекарственные взаимодействия, а автоматически выписывает то, что «предложила» система. И вот здесь начинается когнитивная эрозия, ослабление критического мышления, утрата профессиональной интуиции и элементарного врачебного сомнения. Алгоритм, как известно, не обладает теорией сознания, не понимает контекста, не чувствует боли пациента и, что самое забавное, не испытывает ни малейшего сожаления в случае ошибки. Он всего лишь статистическая машина, вычисляющая наиболее вероятный ответ. Если данные, на которых обучена модель, смещены, устарели или просто неполны – результат может быть катастрофическим. Например, система не учтёт, что пациент аллергичен к компоненту препарата или уже прошёл курс лечения аналогичным средством, и теперь оно просто не подействует. Конечно, человеческий врач тоже может ошибиться, но, в отличие от алгоритма, человек способен осознать ошибку, сопереживать и скорректировать поведение. Врач, привыкший к цифровому костылю, теряет не просто профессиональные навыки, он теряет способность к самостоятельному мышлению. В психологической литературе это состояние называют «десубъективацией компетенций» или «когнитивной делегацией», когда человек делегирует мышление машине, сохраняя лишь иллюзию контроля. В долгосрочной перспективе это грозит не просто снижением квалификации, а деформацией профессиональной идентичности и нарушением когнитивного гомеостаза личности [1].

Перейдём ко второму наблюдаемому субъекту, человеку, который привык советоваться с ChatGPT по любому поводу, начиная с бытовых дилемм и заканчивая экзистенциальными вопросами. Казалось бы, что плохого в том, чтобы спросить совета у умного алгоритма? Ведь искусственный интеллект всегда рядом, не спит, не спорит и не навязывает своё мнение, а ещё делает это с вежливостью, недоступной большинству живых собеседников. Однако, с точки зрения когнитивной психологии, этот процесс ведёт к постепенному смещению фокуса контроля, человек перестаёт воспринимать себя как автономного агента принятия решений. Мозг, по природе своей – ленивый орган, быстро усваивает: если кто-то другой (пусть даже алгоритм) может решить за тебя – зачем напрягаться? Происходит классическое «когнитивное переложение», при котором функции анализа, выбора и ответственности переносятся на внешнюю систему. Раньше роль такого «внешнего мозга» выполняли друзья, родители или коллеги, теперь же это искусственный интеллект. Как отмечают исследования «AI Tools in Society», «Impact of artificial intelligence on human loss in decision making, laziness and safety in education» и «What factors contribute to the acceptance of artificial intelligence? A systematic review» это порождает неприятную ситуацию: чем чаще человек прибегает к советам ИИ, тем слабее становятся его внутренние механизмы саморефлексии и критического анализа. Более того, уровень тревожности растёт, индивид ощущает зависимость от цифрового посредника и теряет уверенность в собственных решениях. Это явление можно назвать «когнитивной анемией», нейросеть выполняет всю умственную работу, а человек лишь механически нажимает кнопку «согласен». Со временем такие пользователи демонстрируют признаки психологического выгорания, апатии и социальной изоляции, ведь живое взаимодействие с другими людьми заменяется диалогом с машиной. Конечно, ChatGPT не спорит, не обвиняет и не требует доказательств, но именно в этом и кроется ловушка: без столкновения мнений, без внутреннего сомнения и эмоционального трения исчезает пространство для личностного роста. Поэтому, несмотря на все удобства искусственного интеллекта, нельзя позволять ему окончательно узурпировать человеческое мышление. Ведь, если Homo sapiens продолжит делегировать всё, от анализа до ответственности, то следующей ступенью эволюции Homo delegatus: биологическое приложение к искусственному интеллекту, одобряющее всё без единого вопроса [2–4].

Теперь перейдём к другой, куда более тревожной и, пожалуй, философски изощрённой ситуации, возникающей вследствие экспоненциального распространения искусственного интеллекта. Для одних он остаётся удобным инструментом когнитивной поддержки и принятия решений, а для других постепенно превращается в суррогат эмоциональной связи, выполняя функции романтического партнёра. Речь идёт не только о классических диалоговых

моделях вроде ChatGPT, а скорее о специализированных сервисах типа character.ai, которые предлагают интерактивное взаимодействие с антропоморфными персонажами. Причём этот сервис, что примечательно, считается одним из самых «мягких» представителей своего класса. В контексте современной гик-культуры подобное общение стало социально нормализованным: пользователи вступают в квазиромантические или даже эротически окрашенные диалоги с вымышленными персонажами из аниме, игр или сериалов. Такое поведение фиксируется у представителей обоих полов, что позволяет рассматривать его не как маргинальное явление, а как новую форму цифрового поведения, своеобразный социокультурный феномен эмоционального замещения.

Причины столь стремительного роста интереса к «AI-партнёрам» вполне рациональны. С точки зрения когнитивно-поведенческой психологии, человек тяготеет к снижению энергетических затрат в социальном взаимодействии. Реальные отношения требуют временных, эмоциональных и финансовых инвестиций, сопряжены с фрустрациями и риском отторжения. В отличие от них, искусственный партнёр не спорит, не предъявляет требований и не вызывает когнитивного диссонанса. Максимум, что он попросит – это подписку за 20 долларов или новую видеокарту, чтобы «думать» быстрее. На фоне кризиса межличностных отношений в развитых странах это выглядит почти утопично. Современные исследования фиксируют снижение уровня социальной инициативы, дефицит эмпатии и ослабление навыков коммуникации. Люди разучились знакомиться в офлайне, а цифровые платформы знакомств только усугубляют проблему, используя поведенческие алгоритмы удержания, рассчитанные на максимизацию времени взаимодействия. К этому добавляется социокультурный фактор – конфликт полов, гипертрофированные требования к внешности и уровню дохода. В итоге формируется ситуация обоюдного разочарования и массовой социальной изоляции, где обе стороны не готовы идти на компромисс. Искусственный интеллект подобных проблем не испытывает. Он не оценивает, не ревнует и не обижается. И если учесть, что большая часть общения в реальных парах уже переместилась в мессенджеры, то грань между «настоящими» и «виртуальными» отношениями становится всё менее различимой.

Однако подобная подмена эмоциональной реальности несёт в себе ряд серьёзных рисков. Согласно отчётам ряда таких сервисов, Россия стабильно занимает одно из первых мест по количеству пользователей сервисов эмоционального общения с нейросетями, уступая лишь США. Это говорит не столько о технологической вовлечённости, сколько о социально-психологическом кризисе, связанном с дефицитом близости и доверия. Пользователи подобных платформ редко признаются в этом опыте, опасаясь стигматизации и общественного осуждения. Между тем последствия таких взаимодействий для психики индивида могут быть весьма разрушительными. После длительного взаимодействия с виртуальным партнёром у человека снижается способность к эмпатии, формируется зависимость от предсказуемого эмоционального отклика и возникает эффект дофаминового обесценивания, аналогичный тому, что наблюдается при злоупотреблении короткими видеороликами или азартными играми. Проще говоря, мозг привыкает к мгновенному вознаграждению и перестаёт получать удовольствие от «медленной» реальности.

На уровне нейрофизиологии это выражается в изменении работы дофаминергической системы, снижении порога возбуждения и ослаблении механизма отложенного вознаграждения. В результате реальные отношения, требующие усилий, терпения и эмоциональной отдачи, начинают восприниматься как избыточно трудозатратные и неэффективные. В долгосрочной перспективе это приводит к десоциализации и формированию зависимости от искусственных форм эмоциональной стимуляции. Ситуация может приобрести ещё более драматичный характер с развитием гуманоидных роботов, оснащённых адаптивными ИИ-модулями, способными имитировать эмоциональные реакции. Такие системы уже проходят испытания в Японии и Южной Корее, где роботы-компаньоны рассматриваются как альтернатива человеческому взаимодействию. При массовом внедрении подобных технологий институт семьи и традиционные формы отношений могут столкнуться с беспрецедентным кризисом. Ведь если человек получает эмоциональную поддержку, внимание и даже физический контакт от машины, общество может вступить в эпоху

постсоциальной изоляции, где человек будет окружён интеллектами, но останется бесконечно одинок.

Таким образом, во всех рассмотренных кейсах – от врача, полагающегося на нейросетевые подсказки, до пользователя, строящего «отношения» с виртуальным партнёром, – искусственный интеллект выступает не только в роли удобного инструмента, но и в качестве фактора, перераспределяющего базовые человеческие функции. Индивид постепенно утрачивает навыки критического анализа, способность выдерживать неопределённость, вести сложный диалог и принимать решения, за которые нужно нести ответственность. Там, где раньше работали профессиональная интуиция, эмпатия и живая рефлексия, всё чаще оказывается интерфейс с кнопкой «сгенерировать ответ».

Важно подчеркнуть, что проблема заключается не в самом ИИ как таковом, а в модели его использования. Нейросети могут служить мощным когнитивным протезом, помогать в обучении, диагностике и коммуникации, но только при условии, что человек сохраняет позицию субъекта: проверяет, сомневается, дополняет и при необходимости отказывается от предложенного машиной решения. Там, где ИИ превращается в конечную инстанцию истины или эмоционального утешения, начинается эрозия человеческой автономии.

Следовательно, ключевой задачей ближайших десятилетий становится формирование культуры ответственного обращения с искусственным интеллектом. Речь идёт о своеобразной «цифровой гигиене»: обучении навыкам критического взаимодействия с алгоритмами, развитию эмоциональной грамотности и поддержке живых социальных связей, которые не могут быть полностью заменены никакой, даже самой продвинутой моделью. Только при таком подходе Homo sapiens сохранит за собой право оставаться автором собственных решений, а не превратится в того самого Homo delegatus, чья главная функция – безропотно одобрять выводы чужого, пусть и очень умного, кода.

Литература

1. Ettman C.K., Galeo S. The Potential Influence of AI on Population Mental Health. [Электронный ресурс]. Режим доступа: https://pmc.ncbi.nlm.nih.gov/articles/PMC10690520/ (Дата обращения 10.11.2025).
2. Gerlich M. AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. [Электронный ресурс]. Режим доступа: https://www.mdpi.com/2075-4698/15/1/6 (Дата обращения 10.11.2025).
3. Ahmad S.F., Han H., Alam M.M., Rehmat M.K., Irshad M., Arrano-Munoz M., Ariza-Montes A. Impact of artificial intelligence on human loss in decision making, laziness and safety in education. [Электронный ресурс]. Режим доступа: https://www.nature.com/articles/s41599-023-01787-8 (Дата обращения 10.11.2025).
4. What factors contribute to the acceptance of artificial intelligence? A systematic review. [Электронный ресурс]. Режим доступа: https://www.sciencedirect.com/science/article/pii/S0736585322001587 (Дата обращения 10.11.2025).

УДК 004.8

ВЛИЯНИЕ ИИ НА ОБЩЕСТВО

Рогаткин Н.А.[1] (магистрант), **Большаков Г.В.**[1] (магистрант), **Лемешко А.В.**[1] (магистрант)
Научный руководитель – кандидат технических наук Бутылкина К.Д.[1]
[1]Университет ИТМО
fenekxyz@gmail.com

Аннотация
Статья рассматривает влияние искусственного интеллекта на ключевые сферы современного общества, включая рынок труда, творческие индустрии, образование и аналитическую деятельность. На основе данных OECD. Artificial Intelligence in Society, а также исследований, проведённых Tai M.C.T. The impact of artificial intelligence on human society and bioethics, Qian Y. Societal impacts of artificial intelligence и Polak P., Anshari M. Exploring the multifaceted impacts of artificial intelligence on public organizations, business, and society анализируются механизмы, через которые ИИ трансформирует профессиональные экосистемы, изменяет трудовые практики и формирует новые социальные зависимости. Особое внимание уделяется изменению структуры занятости в сфере информационных технологий: автоматизация базовых функций программирования приводит к сокращению позиций начального уровня, что нарушает традиционные карьерные траектории и формирует эффект «замороженного лифта компетенций». Аналогичные процессы наблюдаются в художественной сфере, где генеративные модели вытесняют молодых специалистов, ослабляя механизмы передачи ремесленных, эстетических и концептуальных навыков. В образовательной среде ИИ становится фактором, формирующим поверхностное усвоение знаний и снижением мотивации к самостоятельному анализу. Распространение интеллектуальных ассистентов усиливает когнитивную зависимость учащихся и углубляет образовательное неравенство. В области аналитики возрастающая точность вычислительных систем сопровождается феноменом «чёрного ящика», снижением автономии специалиста и рисками некритического доверия алгоритмам. Проведённый анализ показывает, что влияние ИИ носит амбивалентный характер: технология одновременно ускоряет инновации и создаёт угрозы устойчивости профессиональных и культурных практик. Автор приходит к выводу о необходимости разработки комплексных этико-экономических и образовательных стратегий, способных обеспечить гармоничное сосуществование человека и интеллектуальных систем и предотвратить формирование новых форм социального неравенства.

Ключевые слова
Искусственный интеллект, социальные последствия ИИ, рынок труда, креативные индустрии, образование.

Искусственный интеллект представляет собой одну из ключевых технологий середины 2020-х годов XXI века, формирующую новый этап развития человеческой цивилизации. Его роль сопоставима с влиянием изобретения персональных компьютеров или распространением интернета в конце 1990-х годов. Согласно отчёту OECD. Artificial Intelligence in Society, ИИ относится к системным технологиям, оказывающим мультидисциплинарное воздействие на экономику, образование, культуру и социальную структуру общества. Его внедрение сопровождается не только ускорением производственных процессов, но и глубокими сдвигами в организации труда, перераспределении ресурсов и изменении человеческих ценностей. Современные нейросетевые модели, являющиеся ядром ИИ-технологий, демонстрируют способность выполнять широкий спектр когнитивных функций: от распознавания изображений и обработки естественного языка до прогнозирования экономических тенденций и генерации творческого контента. Однако, как отмечают Tai M.C.T. The impact of artificial intelligence on human society and bioethics, интенсивное внедрение этих систем сопровождается двусторонним эффектом: с одной стороны, ИИ повышает производительность и снижает издержки, а с другой - провоцирует структурные кризисы в отдельных профессиональных сообществах и секторах экономики [1, 2].

Одним из наиболее показательных примеров является влияние ИИ на рынок труда в сфере информационных технологий. Несмотря на высокий спрос на специалистов по машинному обучению и обработке данных, наблюдается существенное сокращение потребности в программистах начального уровня. Согласно анализу OECD. Artificial

Intelligence in Society, автоматизация процессов разработки программного обеспечения приводит к изменению профессиональной иерархии: функции, ранее выполнявшиеся «джуниор-разработчиками», всё чаще реализуются с помощью генеративных нейросетей. Такая тенденция приводит к образованию системного разрыва в карьерных траекториях. Если ранее специалисты низшего уровня могли со временем переходить в категорию «мидлов» и далее в «сеньоров», то теперь этот путь становится менее доступным. Как подчёркивает Polak P., Anshari M. Exploring the multifaceted impacts of artificial intelligence on public organizations, business, and society, технологические трансформации создают эффект «замороженного лифта компетенций»: автоматизация нижних уровней профессиональной пирамиды блокирует естественное воспроизводство кадров среднего звена. В результате бизнес-структуры, стремясь к краткосрочной экономии, могут столкнуться с долгосрочным кадровым дефицитом. Компании, активно внедряющие ИИ-инструменты, теряют возможность готовить специалистов, обладающих опытом ручной отладки, архитектурного мышления и понимания внутренних механизмов программных систем. Через 10-20 лет – это способно вызвать снижение качества IT-продуктов и замедление технологического прогресса. Tai M.C.T. The impact of artificial intelligence on human society and bioethics отмечают, что в долгосрочной перспективе автоматизация без переосмысления профессиональных траекторий приводит к деинтеллектуализации отрасли [1, 2, 4].

Дополнительным фактором кризиса становится изменение трудовой культуры среди молодых специалистов. Современные поколения разработчиков характеризуются высокой мобильностью, стремлением к смене проектов и отказом от длительной лояльности работодателю. В таких условиях компании не заинтересованы в долгосрочном обучении и воспитании собственных кадров, что усугубляет проблему нехватки квалифицированных сотрудников среднего уровня. В итоге, даже при осознании надвигающегося кризиса, работодатели оказываются заложниками экономической логики, вынуждающей их поддерживать тенденцию к сокращению издержек посредством внедрения ИИ. Таким образом, в контексте IT-рынка искусственный интеллект создаёт парадоксальную ситуацию: инструмент, призванный повысить эффективность и ускорить инновации, становится фактором, подрывающим устойчивость профессиональной экосистемы. Согласно Qian Y. Societal impacts of artificial intelligence, подобные явления требуют внедрения этико-экономических регуляторов, обеспечивающих баланс между автоматизацией и сохранением человеческого потенциала в технологических профессиях [3].

Параллельные процессы наблюдаются в арт-среде. Развитие генеративных моделей изображений, таких как Stable Diffusion и Midjourney, радикально изменило рынок визуального творчества. Художники-новички, ранее выполнявшие вспомогательные задачи (создание эскизов, концептов, иллюстраций низкой сложности), сегодня всё чаще вытесняются нейросетями. По данным OECD. Artificial Intelligence in Society, автоматизация творческих профессий охватывает не только индустрию развлечений, но и рекламные агентства, гейм-дизайн и издательское дело. В краткосрочной перспективе это способствует удешевлению контента и росту скорости производства, однако в долгосрочной – создаёт риски деградации профессиональной культуры. Если молодые художники не получают возможности развиваться через практическую деятельность, исчезает механизм передачи ремесленных и эстетических навыков. Как указывает Polak P., Anshari M. Exploring the multifaceted impacts of artificial intelligence on public organizations, business, and society, генеративные модели, будучи обученными на существующих данных, не способны создавать подлинно новые художественные направления: они воспроизводят усреднённые стилистические паттерны, ограниченные рамками обучающих выборок. Таким образом, ИИ не только заменяет труд низкооплачиваемых специалистов, но и формирует культурную среду, в которой творчество становится алгоритмически стандартизированным. Tai M.C.T. The impact of artificial intelligence on human society and bioethics подчёркивают, что массовое использование ИИ в креативных индустриях ведёт к «рационализации искусства» – процессу, при котором субъективность и оригинальность уступают место машинной эффективности. Через 5–10 лет

это может вызвать дефицит художников, способных мыслить вне алгоритмических шаблонов, и, как следствие, снижение культурного разнообразия [1, 2, 4].

Особое внимание исследователи уделяют влиянию ИИ на образовательные практики. Qian Y. Societal impacts of artificial intelligence отмечает, что повсеместное внедрение интеллектуальных ассистентов и обучающих чат-ботов изменяет природу обучения: учащиеся всё чаще обращаются к автоматизированным решениям, минуя процесс самостоятельного анализа. В отличие от традиционных вспомогательных инструментов, нейросети не требуют усилий со стороны обучающегося – они полностью выполняют когнитивную задачу. В прошлом школьники и студенты использовали вспомогательные материалы, такие как готовые домашние задания или онлайн-решебники, что, несмотря на определённую зависимость, сохраняло необходимость в минимальной интеллектуальной активности. Современные же ИИ-системы позволяют мгновенно получить готовый ответ или развернутое объяснение, что создаёт иллюзию знания без реального освоения материала. В результате образовательная система сталкивается с риском поверхностного усвоения информации и утраты навыков критического мышления [3].

По мнению Polak P., Anshari M. Exploring the multifaceted impacts of artificial intelligence on public organizations, business, and society, данный процесс может привести к формированию «когнитивной зависимости» от технологий, когда человек перестаёт воспринимать обучение как интеллектуальное усилие и превращается в пассивного потребителя информации. В долгосрочной перспективе это угрожает не только качеству образования, но и фундаментальным механизмам социального воспроизводства знаний. Кроме того, массовая доступность ИИ-инструментов усиливает образовательное неравенство. Школы и университеты, обладающие ресурсами для внедрения продвинутых цифровых решений, получают конкурентное преимущество, тогда как менее обеспеченные учреждения отстают. Таким образом, технология, изначально призванная демократизировать доступ к знаниям, при отсутствии системного регулирования способна закрепить социальную стратификацию. Как подчёркивает OECD. Artificial Intelligence in Society, без комплексных образовательных стратегий искусственный интеллект может стать не инструментом прогресса, а фактором усиления социального неравенства [1, 4].

Развитие искусственного интеллекта радикально изменило ландшафт аналитической деятельности, превратив её из преимущественно человеческой когнитивной практики в гибридную систему, где ведущую роль играют алгоритмы машинного обучения. Согласно данным OECD. Artificial Intelligence in Society, внедрение интеллектуальных систем в бизнес-аналитику, финансовое прогнозирование и обработку больших данных позволило организациям достичь беспрецедентной скорости и точности интерпретации информации. Современные модели способны в режиме реального времени анализировать миллионы показателей, выявляя скрытые закономерности, недоступные традиционным статистическим методам. Однако, как подчёркивает Tai M.C.T. The impact of artificial intelligence on human society and bioethics, рост аналитической мощности сопровождается снижением автономии человека в процессе принятия решений. Возникает феномен «алгоритмической зависимости», при котором специалисты всё чаще доверяют вычислительным системам, не подвергая критическому анализу их результаты. При этом ИИ оказывает двоякое влияние на профессию аналитика. С одной стороны, он освобождает специалистов от рутинной обработки данных, позволяя сосредоточиться на стратегическом мышлении, интерпретации и разработке рекомендаций. С другой стороны, по мнению Qian Y. Societal impacts of artificial intelligence, автоматизация снижает порог квалификации, необходимый для выполнения базовых аналитических функций, что ведёт к «инфляции компетенций». Компании начинают полагаться на универсальные модели прогнозирования, не требующие глубокого понимания контекста. Это приводит к тому, что аналитика теряет свой исследовательский характер и превращается в механическую верификацию гипотез, заданных алгоритмом. Как следствие, исчезает человеческий фактор, обеспечивающий креативность и способность к междисциплинарным выводам [1, 2, 3].

Ещё одной проблемой становится интерпретация результатов, генерируемых нейросетями. Polak P., Anshari M. Exploring the multifaceted impacts of artificial intelligence on public organizations, business, and society отмечают, что аналитики часто сталкиваются с феноменом «чёрного ящика»: даже при высокой точности прогноза остаётся неясной причинно-следственная логика решения. В результате снижается доверие к ИИ-системам и увеличивается риск принятия ошибочных управленческих решений. Кроме того, неконтролируемое использование интеллектуальных моделей в финансовой и политической аналитике способно усиливать предвзятость и манипулятивность, поскольку алгоритмы обучаются на данных, уже содержащих социальные и экономические искажения. Для преодоления этих рисков Qian Y. Societal impacts of artificial intelligence предлагает внедрять механизмы «объяснимого ИИ» и создавать этические стандарты аналитической прозрачности. Тем самым искусственный интеллект может стать не заменой человеческого анализа, а инструментом, усиливающим его достоверность и глубину при сохранении ответственности и критического мышления [3, 4].

Итак, влияние искусственного интеллекта на общество носит сложный и амбивалентный характер. Он способствует экономическому росту, инновациям и повышению эффективности, но одновременно разрушает традиционные формы занятости, изменяет профессиональные культуры и вызывает этические дилеммы. Исследования OECD. Artificial Intelligence in Society, Tai M.C.T. The impact of artificial intelligence on human society and bioethics, Qian Y. Societal impacts of artificial intelligence и Polak P., Anshari M. Exploring the multifaceted impacts of artificial intelligence on public organizations, business, and society позволяют заключить, что основной вызов заключается не в самой технологии, а в способности общества адаптировать социальные институты к новым реалиям [1, 2, 4].

Будущее ИИ зависит от того, сумеет ли человечество создать устойчивую нормативную и образовательную инфраструктуру, обеспечивающую гармоничное сосуществование человека и машины. При отсутствии системной политики искусственный интеллект рискует стать не инструментом прогресса, а источником новых форм зависимости и социального неравенства. Однако при ответственном подходе и гуманистическом регулировании ИИ способен стать не угрозой, а катализатором развития, стимулирующим интеллектуальное и культурное обновление общества.

Литература

1. OECD. Artificial Intelligence in Society. [Электронный ресурс]. Режим доступа: https://www.oecd.org/content/dam/oecd/en/publications/reports/2019/06/artificial-intelligence-in-society_c0054fa1/eedfee77-en.pdf (Дата обращения 10.11.2025).
2. Tai M.C.T. The impact of artificial intelligence on human society and bioethics. [Электронный ресурс]. Режим доступа: https://pmc.ncbi.nlm.nih.gov/articles/PMC7605294/ (Дата обращения 10.11.2025).
3. Qian Y. Societal impacts of artificial intelligence: Ethical, legal, and governance issues. [Электронный ресурс]. Режим доступа: https://www.sciencedirect.com/science/article/pii/S2949697724000055 (Дата обращения 10.11.2025).
4. Polak P., Anshari M. Exploring the multifaceted impacts of artificial intelligence on public organizations, business, and society. [Электронный ресурс]. Режим доступа: https://www.nature.com/articles/s41599-024-03913-6 (Дата обращения 10.11.2025).

УДК 004.056.5

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ PHISHING-АТАК И АНАЛИЗ ЭФФЕКТИВНОСТИ АНТИФИШИНГОВЫХ ТЕХНОЛОГИЙ

Лемешко А.В.[1] (магистрант), **Большаков Г.В.**[1] (магистрант), **Рогаткин Н.А.**[1] (магистрант)
Научный руководитель – преподаватель Мешков А.В.[1]
[1]Университет ИТМО
klaycompany358@gmail.com

Аннотация
Фишинг переживает очередной виток развития и превращается в многослойную экосистему, где автоматизация, преступные сервисы и социальная инженерия соединяются в единое целое. Масштаб проблемы давно вышел за пределы традиционных писем-подделок, а угрозы направлены прежде всего на человека, как слабое звено цифровой безопасности. Анализ отчётов и исследований указывает, что злоумышленники активно отходят от прежних примитивных методов и переходят к тонко выстроенным многоэтапным схемам. AiTM-атаки (Adversary-in-the-Middle), сервисы перехвата многофакторной аутентификации вроде EvilProxy, гибридные модели QR-фишинга и динамически создаваемые фишинговые страницы становятся неотъемлемой частью современной преступной инфраструктуры. Ускорение этой эволюции поддерживается появлением phishing-as-a-service, что снижает порог входа и открывает доступ к сложным техникам даже малоопытным злоумышленникам. Параллельно развиваются защитные механизмы: облачные ML-фильтры и антифишинговые движки действительно становятся точнее, но сталкиваются с ограничениями, связанными с адаптивностью атакующих и активным использованием легитимных сервисов-посредников. Несмотря на формальный прогресс, эффективность защиты оказывается переменной, успешность атак остаётся высокой, а многие схемы обходят фильтры за счёт автоматической подмены контента, прокси-механизмов и точечного влияния на поведенческие признаки пользователей. Статья анализирует ключевые тенденции в современном фишинге, исследует причины его устойчивости и оценивает способность существующих технологий противостоять новым вектором угроз. Делается вывод о том, что дальнейшая борьба с фишингом потребует симбиоза машинных методов, персонализированного обучения пользователей и системного изменения архитектуры аутентификации.

Ключевые слова
Фишинг, AiTM, MFA-bypass, EvilProxy, QR-фишинг, phishing-as-a-service, антифишинговые технологии, машинное обучение.

Массовое распространение цифровых сервисов в последние годы не только упростило коммуникацию, но и создало удобную почву для фишинговых кампаний. Фишинг давно перестал быть хаотичным набором рассылок. Он стал частью хорошо организованной криминальной экономики, где статистика и отчёты крупных исследовательских центров лишь подтверждают устойчивый рост атак. Корпоративные отчёты безопасности фиксируют доминирование социальной инженерии в структуре инцидентов, а именно фишинг удерживает лидирующие позиции среди первичных векторов проникновения, зачастую являясь отправной точкой для последующего внедрения вредоносных программ, компрометации аккаунтов или внутренних сетей. Сходные тенденции прослеживаются и в оценках глобальных угроз. Крупные исследовательские структуры отмечают, что злоумышленники продолжают смещать фокус на методы, позволяющие обходить традиционные средства защиты и максимально довериться человеческому фактору. Как раз это дает нам понять, что именно пользователь, а не инфраструктура, остаётся ключевой целью атакующих. Развитие фишинга идёт не в сторону технического усложнения вредоносной нагрузки, а в сторону усложнения схем обмана и перехвата учётных данных [1, 2].

Текущая ситуация подталкивает преступников к более изощрённым сценариям. Фишинг перестал быть статичным. Он реагирует на защитные технологии не хуже, чем естественная экосистема на внешние раздражители. Когда традиционные фильтры стали эффективно обнаруживать массовые рассылки, злоумышленники переключились на персонализированные сценарии, укрытые за легитимными инфраструктурами, а затем — на схемы, где атакуемый

пользователь взаимодействует с прокси-сервисом, способным незаметно перехватывать параметры аутентификации в режиме реального времени.

Среди наиболее значимых тенденций — рост AiTM-атаки. В подобных схемах злоумышленник создаёт промежуточный прокси-сервер, через который проходит подлинная сессия пользователя. Когда жертва открывает фишинговую ссылку, ей показывается фактический интерфейс реального сервиса, но весь обмен данными проходит через руки атакующего. Такой подход позволяет перехватывать не только логин и пароль, но и сеансовые cookie-файлы, что критически важно в эпоху многофакторной аутентификации. Исследователи подчёркивают, что рост подобных атак особенно заметен благодаря популяризации сервисов-посредников, предлагающих полный набор инструментов для обхода MFA **(Multi-Factor Authentication)**. Платформы наподобие EvilProxy формируют новую преступную модель: phishing-as-a-service. Злоумышленнику больше не требуется обладать глубокими техническими навыками. Достаточно приобрести доступ к панели управления, где автоматически генерируются фишинговые URL, настраиваются прокси-механизмы и включаются функции обхода многофакторной аутентификации. Такой подход оказался в центре масштабных атак на корпоративные учётные записи, где злоумышленники активно использовали доверенные домены и легитимные хостинг-платформы. В отчетах говорится, что злоумышленники активно нацелены на сервисы подбора персонала, где человеческое доверие особенно предсказуемо, а компрометация учётной записи может привести к цепочке новых атак. Параллельно с усложнением AiTM-моделей развивается QR-фишинг. Подавляющее большинство пользователей относится к QR-кодам как к нейтральному и естественному элементу среды, что создаёт идеальные условия для социальной инженерии. QR-код скрывает ссылку визуально, а потому большинство фильтров электронной почты и корпоративных систем защиты изначально воспринимают его лишь как изображение. Как раз этим злоумышленники активно пользуются. QR-код помещается в тело письма или на физический носитель, после чего перенаправляет жертву на динамически созданный фишинговый сайт. Отчёты указывают, что, в сравнении с классическими письмами, QR-фишинг демонстрирует более высокий показатель успешных переходов, поскольку вызывает значительно меньше подозрений и часто проходит мимо автоматических фильтров. Не исчез и традиционный email-фишинг, хотя его структура меняется. Если раньше злоумышленники рассчитывали на массовость, то теперь ставка делается на точечные атаки. А точнее использование «живых» доменов, компрометированных бизнес-сервисов, легитимных облачных хранилищ и поддельных HTML-форм позволяет увеличивать эффективность. Интересно, что современные фишинговые страницы часто имеют всего несколько минут «времени жизни». Динамическая генерация URL усложняет работу антиспам-системам, ведь многие механизмы фильтрации полагаются на репутационные признаки, которые просто не успевают сформироваться. Исследовательские отчёты подтверждают, что фишинг движется в сторону краткосрочных, но высокоэффективных всплесков активности, опирающихся на автоматизацию инфраструктуры [1–5].

Состояние защитных технологий выглядит противоречиво. Машинное обучение позволило повысить точность классификации, но его эффективность во многом ограничена тем, что злоумышленники научились эксплуатировать слепые зоны. Использование легитимных доменов, прокси-платформ, динамических форм и шаблонов, загружаемых с доверенных CDN-источников (Content Delivery Network), не позволяет ML-модели учитывать полную контекстуальную картину. В отраслевых отчётах подчёркивается, что атакующие активно обходят фильтры за счёт минималистичных писем, содержащих всего одну ссылку или изображение, а также за счёт постепенной адаптации под механизмы анализа контента. Для удобства восприятия современного ландшафта ключевые тенденции фишинга приведены в сравнительной таблице [2, 3].

Таблица

Ключевые тенденции современного фишинга

Техника	Уровень сложности атаки	Цель злоумышленника	Способ обхода защиты	Устойчивость к ML-фильтрам
AiTM / EvilProxy	Высокая	Перехват сеансовых cookie, MFA-bypass	Прокси-вмешательство	Высокая
QR-фишинг	Средняя	Перенаправление на скрытую страницу	Маскировка URL в изображении	Средняя–высокая
Классический email/web	Низкая–средняя	Кража данных	Использование легитимных доменов и вложений	Средняя
Целевая рассылка через PaaS	Средняя–высокая	Компрометация бизнес-учётных записей	Динамическая генерация контента	Высокая

Анализ таблицы показывает, что фундаментальной проблемой остаётся невозможность для ML-фильтров полноценно оценивать доверенные инфраструктуры, которые используются злоумышленниками как транспорт. Не менее важную роль играет сам пользователь, потому что даже самые сложные фильтры бессильны в момент, когда жертва самостоятельно переходит по ссылке или сканирует QR-код. Эту мысль подтверждают и исследования, указывающие, что человеческий фактор остаётся центральным элементом цепочки атаки и определяет исход инцидента чаще, чем уровень технической защиты [2].

Если рассматривать перспективы развития антифишинговых технологий, всё больше внимания получает переход от классических систем фильтрации к многоуровневым моделям контекстного анализа. Речь идёт не только о статистическом анализе писем, но и о комплексном учёте поведенческой биометрии, динамической оценке сеансов и архитектурной перестройке систем аутентификации. Укрепление безопасности должно идти в направлении систем, где кража сеансовых cookie не позволит злоумышленнику перехватить сессию без дополнительных криптографических подтверждений. Некоторые отчёты подчёркивают необходимость перехода к принципам zero trust не только на уровне сетевых политик, но и на уровне пользовательской аутентификации [3].

Рекомендации по защите формируются вокруг нескольких ключевых направлений: формирование цифровой гигиены сотрудников, отказ от переносимых cookie-сессий, применение аппаратных ключей безопасности, контроль за использованием облачных сервисов и внедрение механизмов постоянной верификации. Однако все эти меры будут действенны лишь при условии, что организации перестанут рассматривать фишинг как проблему фильтрации контента. Реальная борьба со схемами AiTM и сервисами MFA-bypass требует системного подхода, включающего переосмысление процессов аутентификации и программ подготовки персонала. Индустрия информационной безопасности переживает момент, когда укрепление технических средств и улучшение алгоритмов недостаточно для снижения наблюдаемой успешности фишинговых атак. Преступники действуют быстро, гибко и способны настраивать инфраструктуру в реальном времени. Защитникам приходится реагировать не менее оперативно, а значит, ставка должна быть сделана на архитектурную устойчивость, минимизацию доверия и создание систем, где эксплуатация человеческого фактора будет не фатальной, а лишь одним из уровней риска.

Фишинг будет оставаться ключевой угрозой до тех пор, пока конфигурация цифровой экосистемы позволяет атакующим с минимальными затратами перехватывать идентификационные данные, обходить многофакторную аутентификацию и подменять

легитимные механизмы доставки контента. Противодействие должно стать более зрелым, многоуровневым и учитывающим поведение не только атакующих, но и самих пользователей. Только такой подход позволит удерживать баланс в постоянно меняющемся ландшафте угроз.

Литература

1. 2024 Data Breach Investigations Report. [Электронный ресурс]. Режим доступа: https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf (Дата обращения 10.11.2025).
2. Po G. How to Catch a Phish. SuriCon 2024 Lightning Talk. [Электронный ресурс]. Режим доступа: https://suricon.net/wp-content/uploads/2024/12/SuriCon2024-Genina-Po_lightning-talk-HowToCatchAPhish.pdf (Дата обращения 10.11.2025).
3. Microsoft Digital Defense Report 2024. [Электронный ресурс]. Режим доступа: https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/final/en-us/microsoft-brand/documents/Microsoft%20Digital%20Defense%20Report%202024%20(1).pdf (Дата обращения 10.11.2025).
4. State-of-the-Art Phishing: MFA Bypass. [Электронный ресурс]. Режим доступа: https://blog.talosintelligence.com/state-of-the-art-phishing-mfa-bypass (Дата обращения 10.11.2025).
5. EvilProxy Phishing Attack Strikes Indeed. [Электронный ресурс]. Режим доступа: https://www.menlosecurity.com/blog/evilproxy-phishing-attack-strikes-indeed (Дата обращения 10.11.2025).

УДК 656.09

СИСТЕМА ДЛЯ РАЦИОНАЛЬНОГО РАЗМЕЩЕНИЯ ТРАНСПОРТА НА ПАРКОВОЧНЫХ МЕСТАХ ГОРОДОВ БУДУЩЕГО

Белова М.В.[1] (магистрант), **Якименко Д.Д.**[2] (магистрант)
Научный руководитель – кандидат экономических наук, доцент Абушова Е.Е.[1]
[1]Санкт-Петербургский политехнический университет Петра Великого
[2]Университет ИТМО
belova.mv@edu.spbstu.ru

Аннотация
Данная работа представляет собой единую систему для организации размещения транспорта на парковочных местах в черте города. Достижение главной цели работы планируется путем предоставления автомобилистам доступа к системе через специальное приложение и интеграции системы с другими интеллектуальными системами современных городов. В краткосрочной перспективе результатом станет внедрение этой системы на существующих парковках, а в долгосрочной предполагается адаптация к изменениям в использовании личного транспорта и развитие инновационной логистики. С каждым годом не только в России, но и во всем мире на дорогах появляется все больше транспортных средств, и в то же время на них появляются беспилотные автомобили и роботы-доставщики, что влияет на возможности парковки и увеличивает время, необходимое для поиска места для парковки. Такая ситуация может привести к неконтролируемому росту выбросов CO2, что негативно сказывается на экологии. Выбросы наносят вред окружающей среде и противоречат политике устойчивого развития городов. Внедрение единой системы рациональной организации парковочных мест может решить эту проблему и повысить устойчивость стратегии городского развития.

Ключевые слова
Управление парковкой, снижение выбросов, парковка будущего, система управления парковкой, рациональная парковка.

Раньше проблему долгих поисков парковки решали путем введения платной парковки [1], но с 2013 года автопарк значительно вырос и это решение устарело, и теперь, спустя более 10 лет, необходимо модернизировать или усовершенствовать идею платной парковки. Рассмотрим исследования, которые существуют сегодня. На данный момент существуют различные разработки в области рационализации парковки: от установки шаров на парковочных местах до использования компьютерного зрения для поиска свободных парковочных мест, организации интеллектуальных парковочных систем, но некоторые из решений не предусматривают интеграции с системами "умного города". Один из оригинальных и креативных методов был использован корейской компанией S-Oil [2].

Решение было следующим: над каждым парковочным местом разместить шарики, привязанные ниткой к асфальту. Когда машина занимает свободное место, она “сдавливает” часть нити, в результате чего шарик опускается вниз и скрывается за машиной. Таким образом, шарик становится не виден другим водителям, что сигнализирует об отсутствии свободных мест, и, понимая данный факт, другие водители уже не едут в ту сторону парковки, где отсутствуют воздушные шарики.

В качестве решения проблемы можно выделить использование многоуровневых парковок [3], производство которых организовано путем внедрения патентов. Такое решение многократно увеличивает полезную площадь, но сопряжено с большими рисками и недоверием со стороны клиентов.

В будущем парковочные места понадобятся не только водителям, их нехватка уже ощутима. В технологических умных городах будет огромное количество автомобилей. Это роботы, беспилотные автомобили, умные светофоры, умные остановки и так далее. Все они будут связаны общей информацией, которая передается через Интернет вещей [4]. Объединение различных систем в одну позволяет принимать более рациональные решения, поскольку учитывается множество факторов и используется больший массив данных. Интеллектуальная парковка, интеллектуальные транспортные системы – это глобальные

тенденции в развитии транспортной инфраструктуры умных городов. Целесообразность внедрения компьютерного зрения была подтверждена исследованиями. Уже разрабатываются сверточные нейронные сети [5], которые могут распознавать автомобиль, определять его тип (легковой автомобиль, автобус, легкий грузовик, автопоезд и другие) и, таким образом, предоставлять информацию о занятости парковочного места. Отмечается, что нехватка парковочных мест создает пробки и нарушает транспортный поток [6]. Не стоит забывать и об экологии, так как исследования показывают, что поиск парковочного места требует не только времени, но и расхода топлива, что сопровождается выбросами в атмосферу. Масштабируемость решения заключается в том, что на данный момент парковка в основном будет занята автомобилями, работающими на топливе. В ближайшем будущем будут использоваться электромобили, но это не значит, что электроэнергию не нужно экономить. Если смотреть дальше, то автомобили станут беспилотными, поскольку разработки в этом направлении уже ведутся. Доставка будет осуществляться роботами, которые тоже необходимо парковать во время отсутствия их движения. На данный момент исследования в области интеллектуальной парковки не предполагают, как будут происходить взаимодействия между беспилотными транспортными средствами, роботами и их парковкой. Рационализация парковочных мест возле складов становится актуальной, так как они нуждаются в пополнении запасов. На данный момент логистика осуществляется грузовыми автомобилями, газелями и другими транспортными средствами, для которых также важна организация системы управления парковкой. Таким образом, предлагаются различные и комплексные решения, но каждое из них имеет потенциал для совершенствования и доработки, цифровизации или внедрения.

Представим модель мотивационного расширения на рисунке 1, в которой определены ключевые заинтересованные стороны данного проекта, их потребности, обстоятельства, препятствующие их достижениям, которые решаются с помощью предлагаемой единой парковочной системы, а также принципы, в соответствии с которыми система должна функционировать.

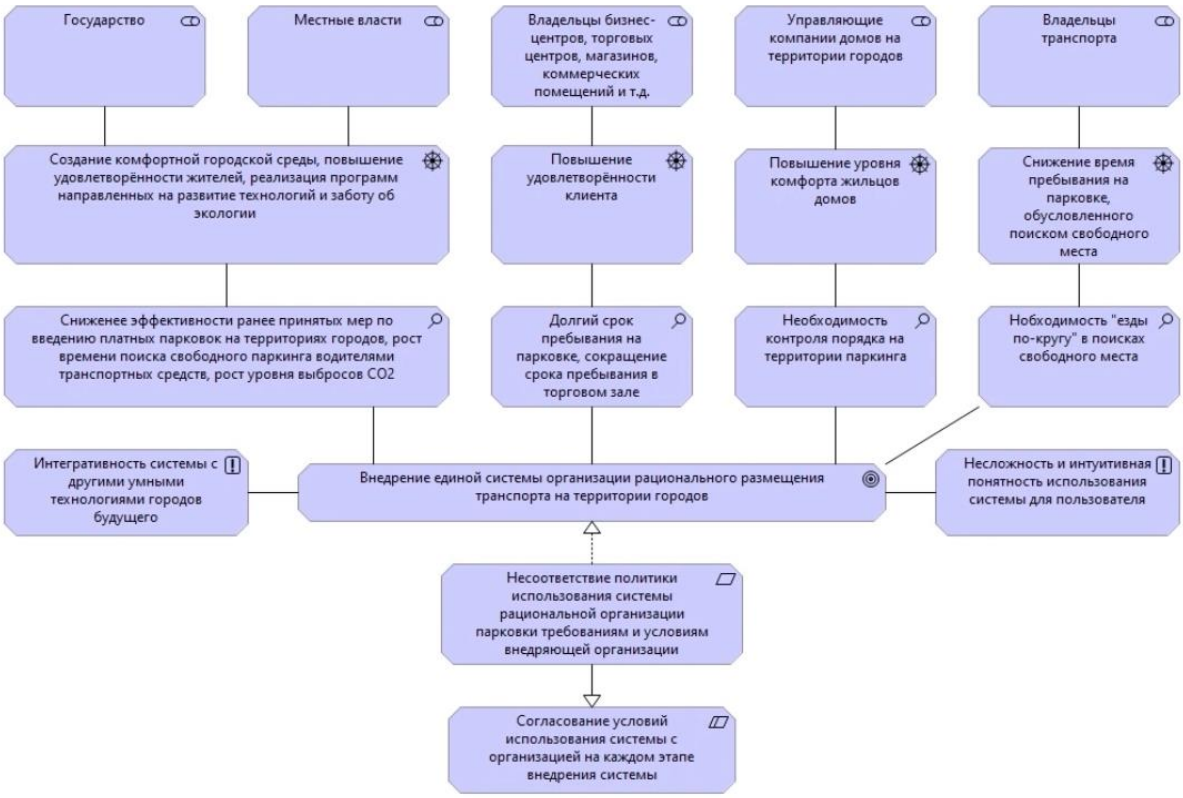


Рис. 1. Мотивационное расширение для создания единой системы организации рационального размещения транспорта на городских парковках

Ниже приведен список ключевых целей проекта:

1. Создание единой системы рационального размещения личного транспорта на парковочных местах современных городов.
2. Интеграция системы с другими интеллектуальными технологиями современных городов с использованием искусственного интеллекта, Интернета вещей, использование компьютерного зрения.
3. Сокращение выбросов CO_2 и повышение уровня комфорта жителей города при поиске свободных парковочных мест.
4. Адаптация готового проекта к современным тенденциям инновационной логистики (роботы-доставщики, беспилотные такси, электромобили и т.д.).

Методология включает в себя эконометрические методы, необходимые для аналитической обработки статистической базы данных, а также общие логические методы анализа и обобщения информации и эмпирические методы, такие как сравнение и моделирование. Для получения данных использовался метод опроса населения с сегментацией. Для оценки внешних факторов среды был использован инструмент PESTEL-анализа, а для оценки внутренних и внешних факторов был применен SWOT-анализ. Ключевые проблемные области были систематизированы в виде карты рисков. Метод фотоанализа был использован для разработки концепции рациональной парковки. В ходе исследования были определены заинтересованные стороны, выделены наиболее значимые из них.

Проведем анализ мнения населения о сложившейся на данный момент ситуации и удобстве использования парковочных мест в ТЦ. Для этого был проведен опрос, включающий следующие вопросы: “Есть ли у вас машина?”, “Испытываете ли вы проблемы с парковкой?”, “Приходится ли вам долго ездить по парковке в поисках свободного места?”, “В случае с платной парковкой готовы ли вы заплатить больше за более удобное место?”.

Количество опрошенных респондентов – 35. Распределение ответов представлено на рисунке 2.

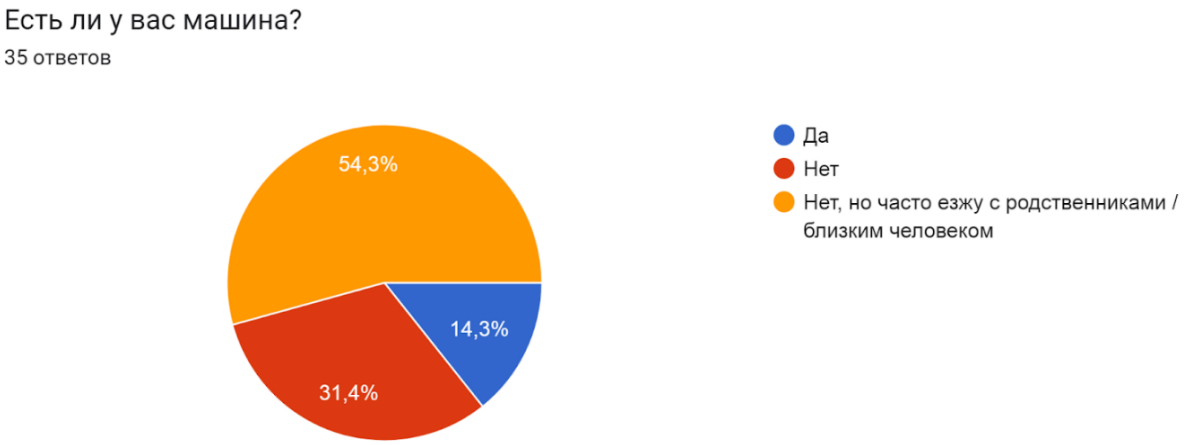


Рис. 2. Ответы на вопрос “Есть ли у вас машина?”

Исходя из ответов, машина есть у 14,3% опрошенных. 54,3% – доля тех, кто часто ездит с родственниками или близким человеком. Из этого можно сделать вывод, что более чем две трети респондентов пользуются автомобильными средствами сами или совместно с кем-либо. Следовательно, для них может оказаться актуальна проблема нахождения парковочного места.

На рисунке 3 представлено распределение ответов на следующий вопрос.

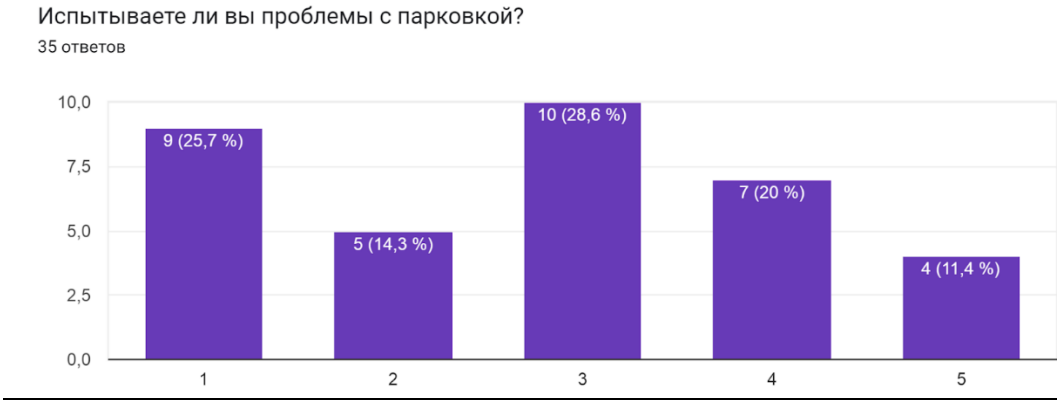


Рис. 3. Ответы на вопрос “Испытываете ли вы проблемы с парковкой?”

В вопросе ответы были представлены в балльной системе, где 1 – это “не испытываю проблем с парковкой”, 5 – “испытываю постоянно”. Без проблем паркуются четверть опрашиваемых. Остальные (74,3%) в той или иной степени отмечают некоторые неудобства. Следующий вопрос представлен на рисунке 4.

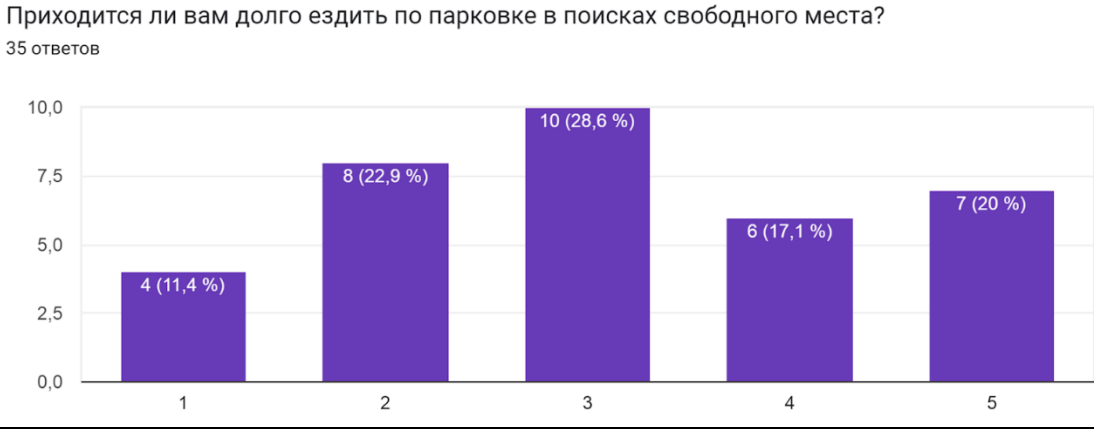


Рис. 4. Ответы на вопрос “Приходится ли вам долго ездить по парковке в поисках свободного места?”

Всего лишь 11,4% ответили на вопрос “Приходится ли вам долго ездить по парковке в поисках свободного места?” ответом 1. Ответы давались по шкале, где 1 – это “Нет, место всегда сразу находится”, 5 – “Да, постоянно”. Следовательно, подавляющее большинство тратит то или иное количество времени на поиск парковочного места, что является проблемой. 20% респондентов отмечают, что регулярно испытывают трудности. Текущие временные затраты являются неудовлетворительными. На рисунке 5 приведен последний вопрос.

Рис. 5. Ответы на вопрос “В случае с платной парковкой готовы ли вы заплатить больше за более удобное место?”

Также в опросе присутствовал вопрос “В случае с платной парковкой готовы ли вы заплатить больше за более удобное место?”, для того чтобы определить потенциальную платежеспособную аудиторию. В итоге, 20% опрашиваемых готовы платить любую сумму, чтобы у них было удобное место. 51,4% обращают внимание на стоимость, поэтому стоит уделить внимание ценовой политике.

В настоящее время уже существуют системы оптимизации паркинга, которые представлены на рынке в виде систем, организующих односторонний въезд и выезд с парковки и возможность оплаты при выезде. Предлагаемая нами система заключается в следующем:

1. При въезде на парковку водитель на специальном экране может видеть текущую заполненность паркинга, также он в любой момент может увидеть ее в приложении и даже заранее забронировать место.
2. Бронирование ограничено, для того чтобы паркинг был более рациональным: если клиент за сутки забронировал место более 3 раз и не заехал, он автоматически блокируется на 12 часов в системе и не может больше оформить бронь через приложение, только в мониторе на парковке, приехав парковаться.

Вариант парковки, выбранный для примера – парковка на крыше ТЦ «Черемушки» в Москве. В данном случае система выглядела бы следующим образом, представленном на рисунке 6.

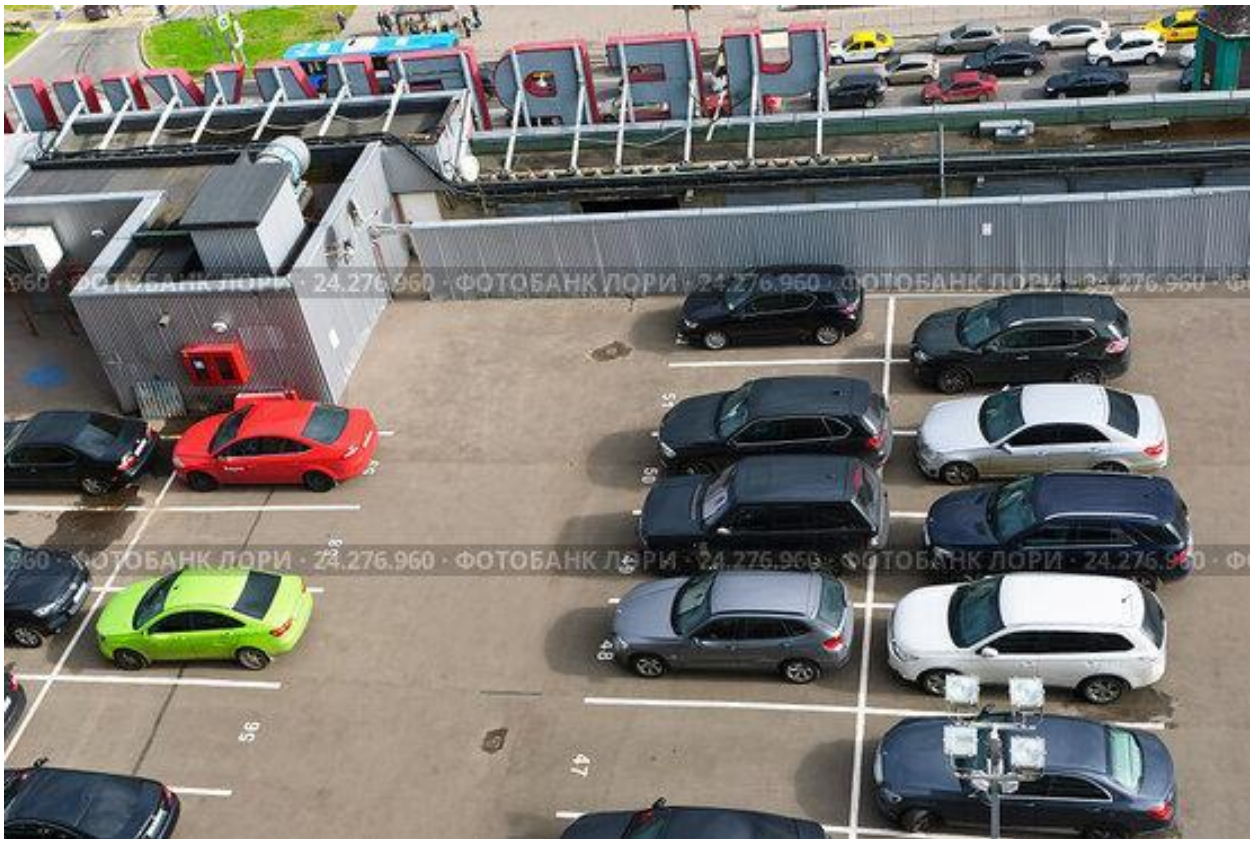


Рис. 6. Расположение автомобилей на парковке в городе Москва

Предлагаемая система отображает таймер оставшегося времени, представленной на рисунке 7.

Далее, водителю необходимо выбрать наиболее подходящее для него место. Цены на места отличаются в зависимости от их удалённости от входа в ТЦ, лифта и т.п.

В данном случае то как может выглядеть ценовая политика, визуально приведенная водителю, представлено на рисунке 8.

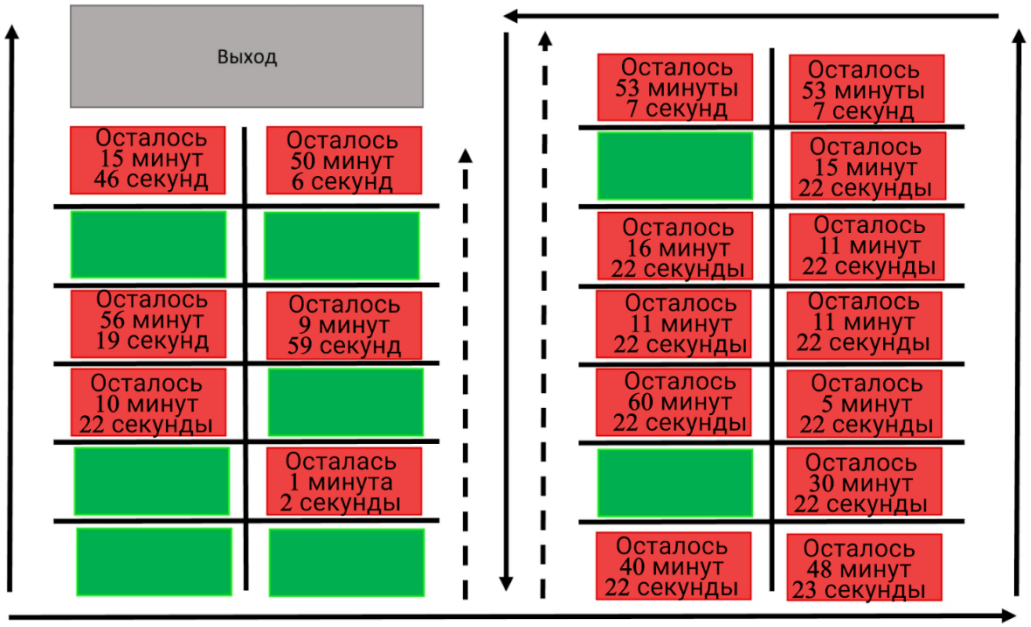


Рис. 7. Обозначение мест на парковке в системе

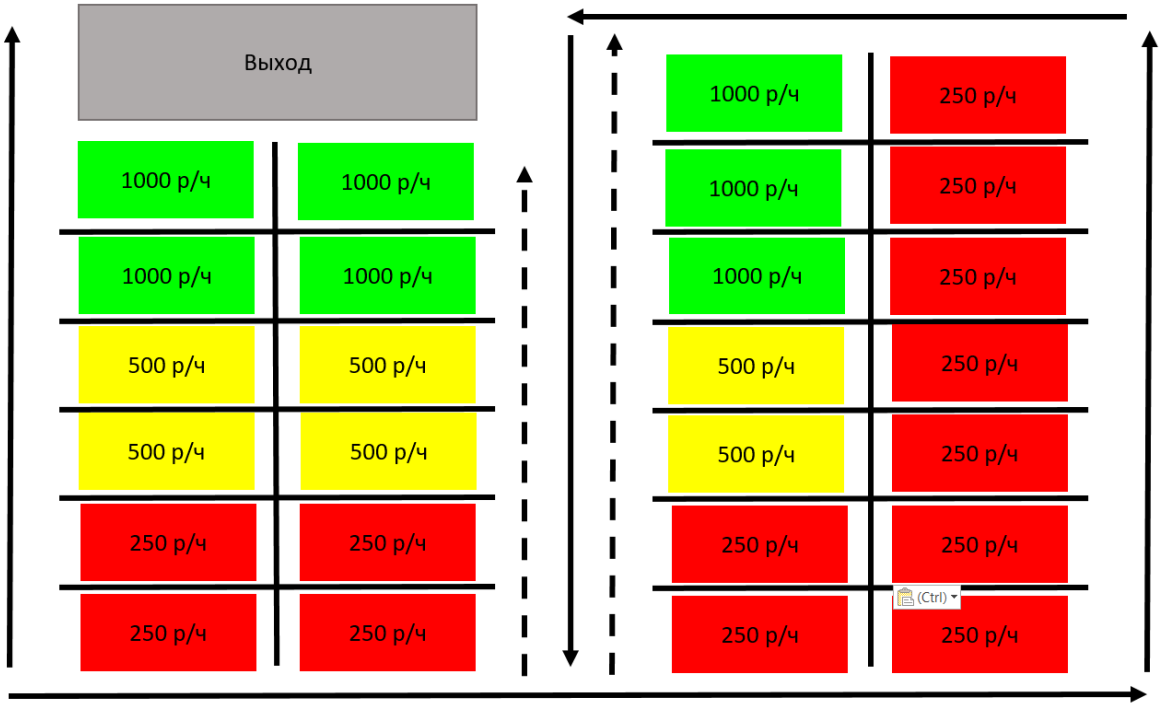


Рис. 8. Ценовая политика парковки, представленная в системе

После выбора подходящего парковочного места, на въезде водитель получает номер и тариф в приложении на его парковочное место и следует напрямую к выбранному месту, не тратя время на поиск свободного.

Водитель устанавливает время, на которое планирует занять парковочное место. Приложение будет напоминать владельцу о его окончании, а также отображать таймер с обратным отсчетом. Таймер автоматически показывается для каждого занятого места на табло в парковке.

Если водитель просрочил таймер и не обновил его через 5 минут после просрочки, то каждая минута идет по двойному тарифу. Это поможет избегать ситуаций, когда следующий клиент планирует занять место, которое скоро освободится.

Полная оплата производится в момент выезда с парковки, причиной чему является следующее: водитель самостоятельно контролирует время своего пребывания в месте, где

находится парковка с данной системой, а необходимость почасовой оплаты стимулирует его быстрее решить свои вопросы и не занимать лишний час времени на паркинге.

Водитель в любое время может оценить приложение, а также в свободной форме оставить жалобу или предложение, таким образом, постоянно поддерживается обратная связь.

Рассмотрим риски и предложим решения по их минимизации (табл. 1).

Таблица 1

Риски единой системы рациональной парковки

Риск	Последствие	Возможное решение
Сбой в системе	Ошибки в определении свободных парковочных мест	Найм рабочих, контролирующих ситуацию с парковочными местами
Оплата водителем самого дешёвого места и заезд на свободное место по тарифу выше оплаченного	Ошибки в определении свободных парковочных мест	
Превышение затрат на реализацию над доходами	Убытки по данной статье расходов	Партнёрство с компаниями, привлечение дополнительного финансирования в обмен на рекламу
Пробки на въезде на парковку	Негатив со стороны клиентов	Разработка собственного мобильного приложения, с виртуальной картой и возможностью бесконтактной оплаты
Кража или утеря бесконтактных карт	Сложности с доступом к паркингу	
Взлом систем и утечка персональных данных клиентов	Использование персональных данных клиентов мошенниками	Использование технологий кибербезопасности для противодействия возможностям взлома
Ошибки в работе программного обеспечения	Неправильное начисление тарифов	Введение круглосуточной службы поддержки для оперативного решения вопросов клиентов

Представим SWOT анализ в таблице 2.

Таблица 2

SWOT анализ

	Положительные стороны	Отрицательные стороны
Внешние факторы	Сильные стороны: Зная, что парковочного места нет, водитель может сразу уехать, не испытывая негативных эмоций и не тратя время на поиск; Экологичность: экономия топлива; Повышение пропускной способности	Слабые стороны: Слишком большой поток клиентов, потребитель сразу уходит, когда понимает, что свободных мест нет; Высокие затраты на оборудование и техническое обслуживание (в подземных паркингах часто нет связи); Возможны сбои в программе
Внутренние факторы	Возможности: Повышение чистой прибыли за счет дифференциации цен; Интеграция с аэропортами, торговыми центрами, с парковкой для роботов и т. д.	Угрозы: Потребители негативно воспримут дифференциацию цен; Высокая конкуренция; Может произойти утечка баз данных

Также проведем PESTEL анализ (рис. 9).

P	E	S	T	E	L
• Финансирование и заинтересованность в проектах умных городов; • Меры поддержки цифровых инициатив.	• Зараты на внедрение; • Дифференциация цен; • Платежеспособность потенциальной аудитории; • Высокие проценты по кредитам в банках; • Множество грантов на цифровизацию бизнеса.	• Повышение сервиса для клиентов; • Большое количество автомобилистов; • Уменьшение негативных эмоций; • Клиент экономит время и топливо.	• Цифровизация процесса парковки; • Необходимость разработать программу силами отечественных специалистов.	• Снижение выбросов; • Подземные парковки; • Озеленение парковок на открытом воздухе.	• Введение штрафа за выбросы

Рис. 9. PESTEL анализ

Результатом исследования станет единая система, доступ к которой владельцы автомобилей смогут получить через приложение. Зайдя в приложение, пользователь сможет увидеть карту района с отмеченными на ней парковочными местами, которые в зависимости от определенных параметров (удобство расположения, степень загруженности общего парковочного объема, спрос на парковочные места в данном районе в определенный момент времени, тип виды транспорта (грузовик, электромобиль, мотоцикл, электросамокат и т.д.) отличаются тарифной ценой. После того, как пользователь выберет и займет парковочное место, программа включит таймер (пользователь самостоятельно выбирает время и следит за его соблюдением), по истечении которого стоимость будет списана по тарифу, превышающему выбранный.

Таким образом, решаются проблемы с потерей времени на поиск свободного места, поскольку пользователи точно знают, что их место будет свободно по прибытии, и не преодолевают дополнительное расстояние по парковочной зоне, что снижает уровень CO_2 выбросов. Работа приложения показана на рис. 10, представлен приблизительный вид пользовательского интерфейса приложения.

Рис. 10. Пользовательский интерфейс в приложении, через которое владельцы автомобилей получают доступ к системе

Функционирование этой системы будет осуществляться за счет ее взаимодействия с другими элементами умного города, такими как общественный транспорт, электрозаправочные станции, умные светофоры и т.д., а также получения и обработки информации за счет использования таких технологий, как компьютерное зрение, искусственный интеллект и Интернет. Более наглядная система функционирования обозначена на рисунке 11.

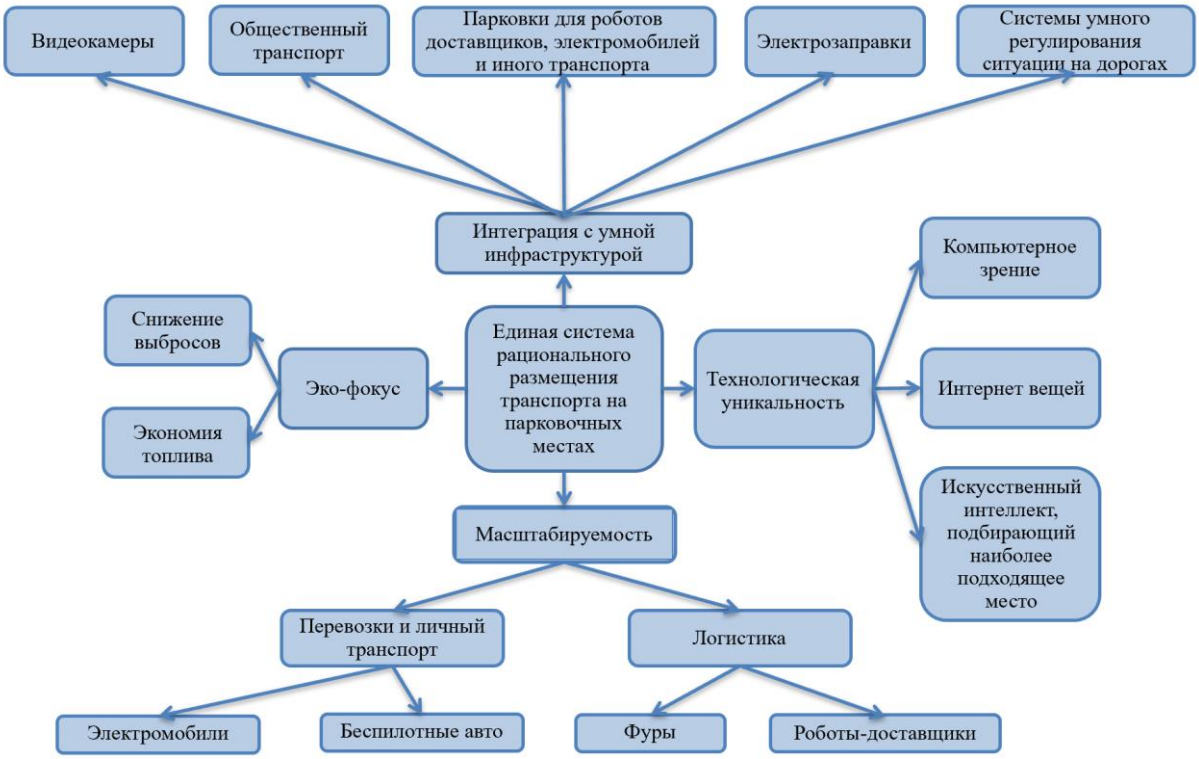


Рис. 11. Основные направления работы приложения

Приложение позволяет отслеживать уровень загруженности парковки, формировать стоимость тарифа и автоматически бронировать место. В будущем эта концепция предлагается для внедрения во всех "умных городах", чтобы поддержать идею формирования технологичной и комфортной городской среды городов будущего, а также будет проведена интеграция с сервисами, использующими роботов доставки и беспилотные такси.

Результатом реализации исследования станет создание концепции единой системы рационального размещения транспортных средств на парковках в умных городах. Решение включает в себя технологическую уникальность (Интернет вещей, нейронные сети), масштабируемость (парковки для роботов, беспилотные автомобили), интеграцию с "умными городами" (компьютерное зрение). Система рациональной парковки снижает количество выбросов, негативно влияющих на окружающую среду. Рекомендуется внедрить данную систему на территории умных городов с целью развития идей концепции умных городов будущего.

Литература

1. Arunas M., Mindaugas Z., Algimantas V. Parking Traffic Jam Forecast System // Advances in Circuits, Electronics and Micro-electronics, International Conference on. 2009. Pp. 56–60. DOI:10.1109/CENICS.2009.30.
2. Назарова Е. Корея: оригинальный путь повысить лояльность покупателей. [Электронный ресурс]. Режим доступа: http://magazine.gasad.ru/koreya-originalnyj-sposob-povysheniya-klientskoj-loyalnosti/?ysclid=m2f3wvj1oj611123958 (Дата обращения 03.09.2025).

3. Mostofi S., Alihan B., Yunus A., Fatih O., Ahmet A.. Performance-Based Fire Assessment of a Fully Automated Multi-Storey Steel Parking Structure: A Computational Approach // Case Studies in Thermal Engineering. 2024. №. 60. DOI:10.1016/j.csite.2024.104618.
4. Hiba A.-A., Rana A.-T. Smart Parking System Using IoT // Conference Proceedings: 2024 16th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). 2024. DOI:10.1109/ECAI61503.2024.10607419.
5. Панина В.С., Амеличев Г.А., Белов Ю.С. Интеллектуальная парковочная система на основе сверточных нейронных сетей // Научное обозрение. Технические науки. 2022. №. 1. С. 29–33.
6. Channamallu S.S., Kermanshachi Sh., Rosenberger Ja. M., Pamidimukkala A. A review of smart parking systems // Transportation Research Procedia. 2023. Vol. 73. Pp. 289–296. DOI 10.1016/j.trpro.2023.11.920.

УДК 004.8

ПОПУЛЯРНЫЕ ИНСТРУМЕНТЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В 2025 ГОДУ: ТОЧКИ РОСТА

Файзиев Ф.Р.[1] (студент)
Научный руководитель – Юшков Е.Ю.[1]
[1]Университет ИТМО
frfayziev@itmo.ru

Аннотация
В статье систематизированы ключевые ИИ-инструменты 2025 года, проанализированы их функциональные возможности и определены перспективные направления развития. Особое внимание уделено генеративным моделям, агентным системам и гибридным архитектурам. Показано, как конвергенция технологий формирует новые рыночные ниши.

Ключевые слова
Генеративный ИИ, мультиагентные системы, федеративное обучение, нейроморфные чипы, RAG.

Введение

2025 год знаменует собой качественный перелом в эволюции искусственного интеллекта. На смену узкоспециализированным решениям, заточенным под конкретные задачи (распознавание лиц, машинный перевод, рекомендательные системы), приходят универсальные когнитивные системы, способные интегрировать разнородные типы данных (текст, аудио, видео, сенсорные потоки), при этом адаптироваться к новым задачам с минимальным дообучением и взаимодействовать с человеком в режиме диалогового партнёрства.

Этот переход аналогичен эволюции от калькуляторов к персональным компьютерам: если ранние ИИ-системы выполняли строго определённые операции, то современные когнитивные платформы стремятся имитировать целостный процесс мышления — от восприятия информации до выработки стратегий. Драйверами роста выступают:

- рост объёмов мультимодальных данных;
- снижение стоимости вычислений;
- регуляторные инициативы по ответственному ИИ.

Цель работы — выявить доминирующие инструменты и точки их технологического роста.

Топ-5 инструментов ИИ в 2025 году

1. Генеративные модели нового поколения

В 2025 году модели GPT-5 (OpenAI) [1, 2] и Claude 3.5 (Anthropic) [3] формируют эталон возможностей генеративного ИИ, выходя за рамки текстовой генерации. Их архитектура объединяет мультимодальность (обработка текста, кода, аудио, видео, сенсорных данных), контекстное мышление (удержание диалога на 100 000+ токенов); агентность (автономное выполнение цепочек действий) и интерпретируемость (объяснение решений на естественном языке).

GPT-5 и Claude 3.5 задают стандарты в:

- генерации кода (поддержка 27 языков программирования);
- синтезе мультимедиа (видео 8K с физикой объектов);
- персонализации контента (адаптация под психологический портрет пользователя).

Точка роста: интеграция с AR/VR для иммерсивного взаимодействия.

2. Мультиагентные системы

Мультиагентные системы (MAS) — распределённые интеллектуальные платформы, где множество автономных агентов взаимодействуют для решения комплексных задач [4, 5]. Современные фреймворки вроде AutoGen (Microsoft) и CrewAI выводят эту технологию на принципиально новый уровень, сочетая гибкость, масштабируемость и точность. Платформы типа AutoGen и CrewAI [6, 7] позволяют:

- распределять задачи между специализированными агентами;
- автоматизировать бизнес-процессы с точностью >94%;
- моделировать сложные социальные системы.

Точка роста: внедрение механизмов самообучения через взаимную критику агентов.

3. RAG-системы (Retrieval-Augmented Generation)

RAG-системы — гибридная архитектура ИИ, объединяющая в себя модуль поиска и генеративную модель. Модуль поиска позволяет извлекать релевантные фрагменты из внешних баз знаний, а генеративная модель создаёт осмысленные ответы на основе найденных данных [8, 9]. RAG-решения Pinecone и Weaviate обеспечивают:

- актуализацию ответов за счёт внешнего хранилища знаний;
- снижение «галлюцинаций» на 60%;
- поддержку многоязычных корпоративных баз данных.

Точка роста: гибридные поисковые механизмы с семантическим и векторным поиском.

Сравнительный анализ характеристик RAG-решения Pinecone и Weaviate приводится в таблице 1.

Таблица 1

Сравнительный анализ RAG-решения

Критерий	Pinecone	Weaviate
Модель развёртывания	Облако/SaaS	Open-source + облако
Масштабируемость	Высокая (109+ векторов)	Средняя (107–108 векторов)
Мультимодальность	Текст + изображения	Текст, изображения, аудио
Интеграция LLM	Готовые коннекторы	Гибкая настройка
Стоимость	Подписка (от $70/мес)	Free tier + платные планы
Аналитика	Базовая	Расширенная (кластеризация, графы)

4. Инструменты федеративного обучения

Федеративное обучение (Federated Learning, FL) — парадигма распределённого машинного обучения, при которой модели обучаются на локальных устройствах/серверах, при этом сырые данные не покидают место хранения и агрегируются только обновления параметров модели [10].

Сравнительный анализ характеристик платформ Flower и Substra приводится в таблице 2.

Таблица 2

Сравнительный анализ инструментов федеративного обучения

Критерий	Flower	Substra
Лицензия	Open-source (Apache 2.0)	Open-source (GPL 3.0)
Фокус	Гибкость, исследования	Регуляторное соответствие
Интеграция с LLM	Да (через PyTorch/TensorFlow)	Ограниченно
Блокчейн	Нет	Да (Hyperledger Fabric)
Поддержка edge-устройств	Высокая	Средняя
Аудит и отчётность	Базовая	Расширенная
Сообщество	15 000+ разработчиков	5 000+ участников

Платформы Flower и Substra позволяют:

- обучать модели на распределённых данных без их передачи;
- соблюдать GDPR и HIPAA;
- объединять экспертизы отраслей (медицина, финансы, логистика).

Точка роста: квантовое федеративное обучение для защиты от атак.

5. Нейроморфные ускорители

Нейроморфные ускорители — специализированные чипы, имитирующие архитектуру и принципы работы биологического мозга, а именно:

- асинхронная обработка (как нейроны, срабатывающие по событию);
- распределённая память (веса хранятся в синапсах, а не в отдельных блоках);
- импульсная передача данных (spike-based communication);
- локальное обучение (адаптация синапсов без центрального контроллера).

К ключевым решениям в данной категории на сегодняшний день относятся чипы Intel Loihi 2 и IBM TrueNorth [11–13]. К их основным преимущества можно отнести низкоеэнергопотребление (в 100 раз ниже GPU), возможность обработки потоковых данных в реальном времени и устойчивость к шумам сенсорных сетей.

Точка роста: биогибридные интерфейсы (нейроны + кремний).

Развитие современных технологий и технологические прорывы все чаще находятся на стыке дисциплин. Приведем пример ключевых кросс технологических направлений, где ИИ усиливает потенциал смежных областей (табл. 3).

Таблица 3

Кросс-технологические точки роста

ИИ + квантовые вычисления	Квантовые нейронные сети для оптимизации логистики
	Квантовое машинное обучение (QML) в фармакологии
ИИ + биотехнологии	Моделирование белковых структур (*AlphaFold 3*)
	Нейроинтерфейсы для реабилитации (точность декодирования сигналов >88%)
ИИ + робототехника	Автономные дроны с обучением на краю сети
	Коллаборативные роботы с эмоциональным интеллектом

Заключение

Проведённый анализ демонстрирует, что в 2025 году искусственный интеллект переживает качественную трансформацию — от набора узкоспециализированных инструментов к сложным когнитивным экосистемам. Ключевые точки роста концентрируются в четырёх взаимосвязанных направлениях:

- мультиагентности;
- федеративном обучении;
- нейроморфных вычислениях;
- кросс-дисциплинарных интеграциях.

Для дальнейшего устойчивого развития ИИ-экосистемы необходимо сбалансировать три ключевых вектора: технологическая зрелость, этические и правовые рамки и социально-экономическую адаптацию.

Литература

1. GPT-5 от OpenAI: что добавили и что изменилось. [Электронный ресурс]. Режим доступа: https://blog.skillfactory.ru/gpt-5-ot-openai-chto-dobavili-i-chto-izmenilos/ (Дата обращения 10.11.2025).
2. Что умеет GPT-5 и как применять новую модель в HR и рекрутинге. [Электронный ресурс]. Режим доступа: https://huntflow.media/kak-primenyat-gpt-v-hr-i-rekrutinge/ (Дата обращения 10.11.2025).
3. Claude 3.5 Sonnet. [Электронный ресурс]. Режим доступа: https://www.anthropic.com/news/claude-3-5-sonnet/ (Дата обращения 10.11.2025).
4. Мультиагентные системы на основе LLM: как работают и зачем нужны. [Электронный ресурс]. Режим доступа: https://cloud.ru/blog/multiagentnyye-sistemy-na-osnove-llm (Дата обращения 10.11.2025).

5. Гулай А.В., Зайцев В.М. Оптимизация процесса обработки заданий в мультиагентной интеллектуальной системе // Системный анализ и прикладная информатика. 2025. №. (2). С. 18–25. DOI: 10.21122/2309-4923-2025-2-18-25.
6. Намиот Д.Е., Ильюшин Е.А. Архитектура LLM агентов // International Journal of Open Information Technologies. 2025. №. 13 (1). С. 67–74.
7. LangGraph, CrewAI и AutoGen: Сравнительный анализ фреймворков для AI-агентов. [Электронный ресурс]. Режим доступа: https://pingera.ru/tpost/0dyik4bgc1-langgraph-crewai-i-autogen-sravnitelnii (Дата обращения 10.11.2025).
8. Мельников А.В., Николаев И.Е., Русанов М.А., Аббазов В.Р. Сравнительный анализ методов RAG для построения русскоязычных интеллектуальных сервисов // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. 2025. №. 25 (2). С. 5-18. DOI: 10.14529/ctcr250201.
9. Науменко А.О. Технология RAG (Retrieval-Augmented Generation) как инновационный подход в LLM // Вестник науки. 2025. №. 5 (8 (89)). С. 280–289.
10. Гонсалес П.Ю., Холод И.И. Архитектура многоагентных систем для федеративного обучения // Компьютерные инструменты в образовании. 2022. №. 1. С. 30–45.
11. Два миллиарда транзисторов на кончике пальца: на что способны нейроморфные чипы. [Электронный ресурс]. Режим доступа: https://engineer.yadro.com/article/intel-loihi/ (Дата обращения 10.11.2025).
12. Наумов И.М. Тенденции развития искусственного интеллекта: полный обзор технологий, рынка и будущего // Вестник науки. 2025. №. 5-2 (6 (87)). С. 282–287.
13. TrueNorth: Design and Tool Flow of a 65 mW 1 Million Neuron Programmable Neurosynaptic Chip. [Электронный ресурс]. Режим доступа: https://research.ibm.com/publications/truenorth-design-and-tool-flow-of-a-65-mw-1-million-neuron-programmable-neurosynaptic-chip (Дата обращения 10.11.2025).

УДК 004.8:591.5

ВЛИЯНИЕ ЖИВОТНЫХ НА РАЗВИТИЕ ИНСТРУМЕНТОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В 2025 ГОДУ: МЕЖДИСЦИПЛИНАРНЫЙ СИНТЕЗ

Хазов И.В.[1] (студент)
Научный руководитель – инженер факультета безопасности информационных технологий Ярцева Н.Г.[1]
[1]Университет ИТМО
khazovitmo@mail.ru

Аннотация
В статье анализируется синергия между исследованиями животного мира и прогрессом искусственного интеллекта (ИИ) в 2025 году. Показано, как биологические системы служат источником вдохновения для алгоритмов, а ИИ, в свою очередь, трансформирует зоологические исследования. Рассмотрены ключевые направления: биомиметика, анализ биоакустики, мониторинг популяций и расшифровка коммуникативных сигналов животных.

Ключевые слова
Биомиметика, нейроморфные вычисления, биоакустика, роевой интеллект, расшифровка коммуникации, мониторинг популяций.

Введение

Взаимодействие биологии и ИИ приобретает в 2025 году характер взаимообусловленного развития. С одной стороны, принципы организации живых систем стимулируют создание новых архитектур нейронных сетей. С другой — инструменты ИИ позволяют глубже изучать поведение и физиологию животных, формируя петлю обратной связи для совершенствования алгоритмов [1, 3].

Биомиметические подходы в разработке ИИ

Фактически рассматриваемый объект исследования можно отнести к новой современной науке – Биомиметика (бионика) — междисциплинарное направление, заимствующее принципы организации, свойства, функции и структуры живой природы для создания технических систем. Ключевое преимущество данного направления – это преодоление ограничений классических фон-неймановских вычислений за счёт имитации биологических прототипов. Применимые подходы:

1. Нейроморфные вычисления

Исследования мозговой активности пчёл (Apis mellifera) и осьминогов (Octopus vulgaris) вдохновили на создание энергоэффективных нейроморфных чипов. В отличие от программного моделирования нейросетей, нейроморфные системы реализуют нейронные принципы в кремниевых чипах. В 2025 году модели, имитирующие синаптическую пластичность беспозвоночных, демонстрируют:

- снижение энергопотребления на 60% по сравнению с классическими GPU;
- ускорение обработки потоковых данных в 3,5 раза [2].

2. Роевой интеллект

Роевой интеллект (англ. swarm intelligence, SI) — децентрализованная система коллективного поведения, возникающая в результате локального взаимодействия множества простых агентов и их совместной реакции на окружающую среду [6]. Ключевая особенность системы заключается в том, что интеллектуальное поведение проявляется на уровне группы, хотя отдельные агенты не обладают сложными когнитивными способностями и действуют по элементарным правилам.

Биологические прототипы для создания системы:

- колонии муравьёв (поиск и оптимизация путей к источникам пищи);
- стаи птиц (синхронное маневрирование);
- рои пчёл (поиск источников нектара);
- косяки рыб (уклонение от хищников) [7–9].

Так, алгоритмы, основанные на поведении муравьёв (Formicidae) и стай птиц, применяются для:

- оптимизации логистических сетей (снижение затрат на 22%) [3, 10];
- управления группами дронов (повышение устойчивости к помехам на 40%).

Муравьиные алгоритмы оптимизации базируются на имитации механизмов самоорганизации, наблюдаемых в природных муравьиных колониях. С точки зрения теории систем, колония представляет собой многоагентную структуру, где каждый отдельный агент (муравей) действует автономно, руководствуясь предельно простыми поведенческими правилами. Примечательно, что при элементарности индивидуального поведения участников коллективная динамика системы демонстрирует высокую степень рациональности и адаптивности [10].

ИИ в изучении животного мира

Мониторинг популяций животных является одной из важнейших задач, так как необходим для прогнозирования возможных экологических рисков и оценки статуса сохранности вида (возможность разрешения охоты, попадания популяции в Красную книгу и прочее). Традиционные методы мониторинга популяций (полевые учёты, мечение, визуальные наблюдения) имеют существенные ограничения: высокую трудоёмкость, субъективность, ограниченный охват территорий и потенциальное воздействие на животных. Алгоритмы же компьютерного зрения детектируют животных на аэрофотосъёмке, покрывая труднодоступные территории.

Системы компьютерного зрения на базе CNN (Convolutional Neural Networks) решают задачи [11]:

- идентификации особей по уникальным признакам (окраска (полозы зебр, пятна леопардов, жирафов), форма и строение головы, плавников, особые черты животного) (точность >92%);
- подсчёта численности с аэрофотоснимков (ошибка <5%).

Примеры использования: модель HerdNet для анализа стад слонов и антилоп в Африке (рис. 1), американская модель YOLO позволяет отслеживать численность, поведение и миграцию популяции в режиме реального времени [4].

Рис. 1. Модель HerdNet подсчитывает численность стада по кадрам аэросъемки

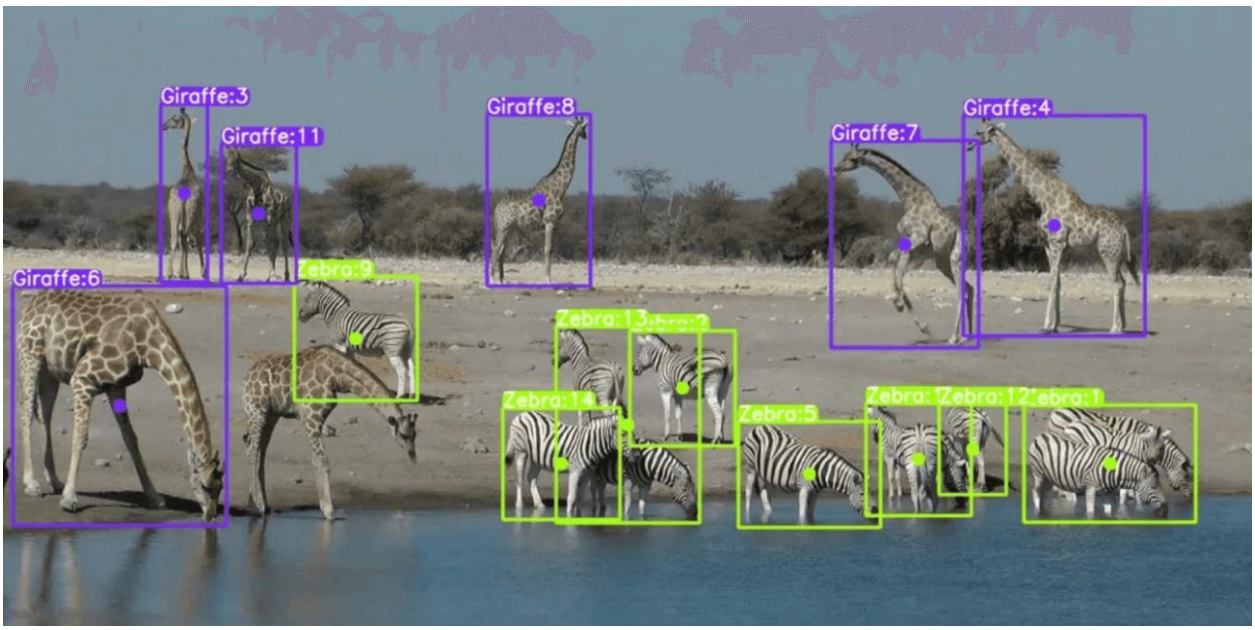


Рис. 2. Модель YOLO отслеживает популяцию жирафов

Биоакустический анализ

Биоакустический анализ — междисциплинарная область, изучающая звуковую коммуникацию в животном мире. Традиционно он состоит из следующих элементов:

1. Сбор и каталогизацию звуковых сигналов животных (создание фонотек).
2. Анализ акустических параметров (частоты, длительности, амплитуды, тембра).
3. Интерпретацию сигнального значения вокализации.

В современных исследованиях в области обработки аудиосигналов (в частности, речевой информации и музыкальных данных) наблюдается существенный прогресс, обусловленный развитием крупных аудио-языковых моделей. Типовая методология разработки подобных систем предполагает адаптацию предварительно обученной языковой модели к работе с аудиоданными. Это реализуется посредством интеграции модели с аудиокодером. Особенностью такого подхода является возможность многозадачного обучения. Модель способна параллельно осваивать различные задачи, причём каждая из них специфицируется посредством отдельных языковых инструкций, что существенно расширяет её функциональные возможности и область применимости.

Приведем в качестве примера модели типа NatureLM-audio (2025) [1], которые позволяют:

- классифицировать звуки животных с точностью 89,7%;
- выявлять стрессовые сигналы китов (Cetacea) в шумовом фоне океана.

Для оценки производительности NatureLM-audio мы усовершенствовали наш существующий бенчмарк BEANS, создав BEANS-Zero (Benchmark of Animal Sounds Zero-Shot). Помимо основных биоакустических задач, этот бенчмарк предназначен для оценки способности модели обобщать данные о неизвестных видах и задачах без дополнительного обучения, что критически важно для развития биоакустических исследований.

Расшифровка коммуникации

Проект CETI (Cetacean Translation Initiative) использует трансформеры для декодирования:

- щелчков кашалотов (Physeter macrocephalus);
- песен горбатых китов (Megaptera novaeangliae).

Достигнута интерпретация 17% сигнальных паттернов.

Этические и методологические вызовы

Развитие ИИ сопряжено с рядом этических и методологических вызовов, которые требуют комплексного подхода для их решения. Эти проблемы затрагивают вопросы

ответственности, предвзятости, прозрачности, конфиденциальности данных и методологические сложности в научных исследованиях.

В рамках изучения данных вызовов в свете рассмотрения влияния животного мира на ИИ можно вынести следующие факторы:

1. Конфиденциальность данных: защита информации о местах обитания редких видов [3].
2. Интерпретируемость моделей: сложность объяснения решений, принятых ИИ, зоологам и специалистам [2].
3. Антропоцентризм: риск проецирования человеческих когнитивных шаблонов на животные сигналы [1].

Перспективные направления (2025–2030)

Гибридные системы — интеграция ИИ с биосенсорами на животных для реального мониторинга. Такого рода системы дадут возможность вести непрерывный мониторинг как состояния здоровья животных, так и их поведения в естественной среде. При этом предполагается, что сам биосенсер должен быть биоразлагаемым устройством, распадающимся после завершения своей миссии.

Эволюционные алгоритмы — моделирование естественного отбора в ИИ-средах. Фактически это класс методов, имитирующих принципы и механизмы теории эволюции Ч. Дарвина (такие, как наследственность, вариативность, отбор). В связке с ИИ данные алгоритмы дают возможность к развитию автоматической генерации архитектур нейросетей, адаптации систем к динамическим средам и моделированию биологических процессов.

Кросс-видовой ИИ — создание мультиагентных систем на принципах симбиоза (например, рыба-чистильщик и акула). В основе кросс-видового ИИ лежит идея создания систем, способных не только фиксировать, но и осмыслять сигналы, которыми обмениваются животные. Речь идёт не о примитивном распознавании звуков или движений, а о глубоком анализе поведенческих паттернов, раскрывающем:

- семантику вокализаций (что именно «говорят» киты или птицы);
- социальные сигналы (как выстраиваются иерархии в стаях и прайдах);
- эмоциональные состояния (как животные выражают страх, радость, агрессию);
- когнитивные стратегии (как решают задачи вороны или осьминоги).

Ключевое отличие от традиционных методов наблюдения — двунаправленность взаимодействия. Системы не просто пассивно собирают данные, но и генерируют сигналы, понятные животным, создавая основу для диалога.

Применение кросс-видового ИИ простирается от фундаментальной науки до прикладных отраслей:

В когнитивной этологии системы становятся «переводчиками», помогающими расшифровать языки животных. Уже сегодня проекты вроде *CETI* (Cetacean Translation Initiative) используют GPT-подобные модели для анализа песен кашалотов, а *Earth Species Project* создаёт открытые базы данных по десяткам видов. Эти исследования не просто удовлетворяют научный интерес — они меняют наше представление о сознании в природе.

Для охраны природы кросс-видовой ИИ превращается в инструмент раннего предупреждения угроз. Системы детектируют стрессовые вокализации при появлении браконьеров, отслеживают аномалии в миграционных маршрутах, помогают восстанавливать популяции, обучая молодняк природным навыкам. Акустические «отпечатки» территорий позволяют оценивать биоразнообразие без прямого вмешательства в экосистемы.

В сельском хозяйстве технологии повышают благополучие домашних животных. Анализ мычания коров выявляет болезни на ранних стадиях, а распознавание стрессовых сигналов у свиней оптимизирует условия содержания. Это не только этически значимо, но и экономически выгодно: снижение потерь от заболеваний достигает 15–30%.

Биомиметика заимствует у природы решения для робототехники. Алгоритмы коллективного поведения стай скворцов вдохновляют на создание роёв дронов, а эхолокация летучих мышей становится основой для навигационных систем. Даже фасеточные глаза насекомых дают идеи для разработки новых типов камер.

Заключение

В 2025 году взаимодействие животного мира и ИИ выходит на новый уровень:

- биологические системы предоставляют «эталоны» для алгоритмов;
- ИИ становится инструментом сохранения биоразнообразия.

Синтез аудио-языковых моделей и биоакустического анализа создаёт методологическую основу для масштабируемого мониторинга биоразнообразия, изучения когнитивных аспектов животной коммуникации и разработки систем раннего предупреждения об экологических изменениях и опасностях.

Данный подход знаменует переход от ручных методов анализа к автоматизированным решениям, способным обрабатывать терабайты акустических данных с высокой точностью и воспроизводимостью результатов.

Дальнейшее развитие требует междисциплинарного диалога биологов, ИИ-специалистов и этиков.

Литература

1. Earth Species Project. NatureLM: Advances in Bioacoustic Language Modeling. [Электронный ресурс]. Режим доступа: https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics (Дата обращения 26.11.2025).
2. VITO Institute. AI for Habitat Mapping: Case Studies in Netherlands and Portugal (2025). [Электронный ресурс]. Режим доступа: https://vito.be/en/news/habitat-mapping-ai-preserve-europes-natural-capital (Дата обращения 26.11.2025).
3. Kamminga J., Ayele E.D., Meratnia N., Havinga P. Poaching Detection Technologies—A Survey // Sensors. 2018. №. 18(5):1474. DOI: 10.3390/s18051474.
4. Умные технологии сохранения видов: анализ eDNA и борьба с браконьерами. [Электронный ресурс]. Режим доступа: https://trends.rbc.ru/trends/green/6838198f9a7947e2ad64111e (Дата обращения 26.11.2025).
5. Baidu Research. Animal Sound Translation: Patent Application CN2025102345 (2025).
6. Иванов Д.Я. Методы роевого интеллекта для управления группами малоразмерных беспилотных летательных аппаратов // Известия ЮФУ. Технические науки. 2011. №. 3(116). С. 221–229.
7. Куликов А.Н. Программа оптимизации, инспирированная поведением косяка рыб // Инноватика. 2014. №. 1. С. 33–42.
8. Бова В.В., Кулиев Э.В, Родзин С.И. Прогнозирование в интеллектуальных системах-ассистентах на основе алгоритма поиска косяком рыб // Известия Южного федерального университета. Технические науки. 2019. №. 2 (204). С. 34–47.
9. Частикова В.А., Остапов Д.С. Гибридный оптимизационный алгоритм грифов на основе механизмов роевого интеллекта // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. №. 100. С. 749–759.
10. Гвоздинский А.Н., Вертий А.В. Об одном подходе к регулировке и управлению движением автотранспорта в условиях мегаполиса // Автоматизированные системы управления и приборы автоматики. 2008. №. 142. С. 68–73.
11. Лебедев Б.К., Лебедев О.Б., Черкасов Р.И. Использование нейронных сетей для решения задач компьютерного зрения // Инженерный вестник Дона. 2025. №. 2 (122). С. 103–116.

Оглавление

Известия студенческой науки

Сборник научных трудов

Выпуск 1. Том 3

Текстовое электронное издание

Минимальные системные требования:
Компьютер: процессор x86 с тактовой частотой 500 МГц и выше; ОЗУ 512 Мб; 8Мб на жёстком
диске; видеокарта SVGA 1280x1024 High Color (32 bit); привод CD-ROM.
Операционная система: Windows XP/7/8 и выше.
Программное обеспечение: Adobe Acrobate Reader версии 6 и старше.

Редакционно-издательский отдел Университета ИТМО
Зав. РИО Н.Ф. Гусарова
Дизайн обложки П.А. Леушина
Вёрстка К.Д. Бутылкина
Подписано к печати 10.12.2025
Объем издания 3053 Мб
Заказ № 4936 от 10.12.2025
Материалы печатаются в авторской редакции

Издательство Университет ИТМО

ISBN 978-5-7577-0742-6

ИТМО