УДК 34

Мультиагентные системы в обеспечении безопасности образовательных учреждений от физических угроз

Шаронов Никита Сергеевич

Ведущий специалист отдела разработки,
ООО «Академия Бизнес Решений»,
студент-магистрант,
Пермский государственный гуманитарно-педагогический университет,
614990, Российская Федерация, Пермь, ул. Сибирская, 24;
e-mail: 69224@pspu.ru

Боярко Сергей Александрович

Ведущий специалист отдела разработки,
ООО «ЮНИКОРН»,
студент-магистрант,
Пермский государственный гуманитарно-педагогический университет,
614990, Российская Федерация, Пермь, ул. Сибирская, 24;
e-mail: 69216@pspu.ru

Аннотация

Предложена и обоснована новая мультиагентная архитектура для обеспечения физической безопасности образовательных учреждений, включающая адаптивный графовый оркестратор и динамическую активацию специализированных агентов. Описана многослойная структура, обеспечивающая отказоустойчивость и лёгкость масштабирования. Процесс внедрения ML-агента-аудитора. Практическая значимость работы заключается в возможности прямого применения разработанных решений для защиты школ и вузов. Предложенная архитектура и методики тестирования могут быть использованы преподавателями вузов при подготовке студентов IT-направлений для преподавания дисциплин по проектированию интеллектуальных систем. Материалы исследования можно использовать специалистам учебных заведений, которые занимаются внедрением инновационных технологий в инфраструктуру безопасности.

Для цитирования в научных исследованиях

Шаронов Н.С., Боярко С.А. Мультиагентные системы в обеспечении безопасности образовательных учреждений от физических угроз // Педагогический журнал. 2025. Т. 15. № 6А. С. 14-22.

Ключевые слова

Образовательные учреждения, мультиагентные системы, безопасность, искусственный интеллект, противодействие угрозам.

Введение

В современных условиях разработки программного обеспечения возрастает потребность в гибких и масштабируемых решениях, позволяющих сократить трудозатраты и сроки воплощения идей. Большие языковые модели (БЯМ) становятся незаменимым инструментом для автоматизации рутинных задач. Программные агенты восполняют слабые стороны БЯМ, интегрируя специализированные модули – будь то математический движок, экспертная система или интерфейс к корпоративным сервисам – и обеспечивая высокую точность и надёжность [Guo и др., 2024].

Однако использование единичного агента нередко оказывается неэкономичным: для каждой задачи нужно разворачивать отдельный модуль, что приводит к дублированию усилий и увеличению затрат на поддержку. Мультиагентные системы, согласно классификации из ГОСТ Р 59277-2020: Классификация систем искусственного интеллекта [ГОСТ Р 59277-2020, 2020] определяет как сети взаимодействующих интеллектуальных агентов, способных решать задачи, которые сложно или невозможно решить одним агентом. МАС, предлагают более эффективный подход: несколько специализированных агентов объединяются при помощи оркестратора в единую структуру, а БЯМ выступает связующим звеном, переводя результаты их работы в понятный для пользователя текст. В таком формате агенты обмениваются данными и совместно решают комплексные задачи, сохраняя при этом модульность и расширяемость системы.

В исследовании стоит задача оценить потенциал применения современных мультиагентных архитектур в разработке программных продуктов для обеспечения безопасности образовательных учреждений. Практический подход позволяет выявить реальные трудности внедрения, сформулировать рекомендации по оптимизации процессов и принять обоснованное решение о включении подобной технологии в портфель продуктов. Полученные результаты могут быть использованы в образовательных целях — как основа для учебных курсов, проектной деятельности студентов и повышения квалификации преподавателей в современных архитектурах программных систем. Это делает исследование особенно полезным для преподавателей, студентов IT-направлений, а также для методистов и руководителей образовательных учреждений, стремящихся внедрять инновационные технологические решения в учебный процесс и инфраструктуру.

Обзор литературы

Исследования демонстрируют, что как статистические методы, так и подходы машинного обучения (ML) применяются для решения разнообразных задач, включая предсказание, классификацию и анализ данных. Эти инструменты находят применение в таких сферах, как материаловедение, нейронаука и здравоохранение — способствуя ускорению научных исследований, интерпретации закономерностей и аналитике улучшений. Выбор между статистическими методами и ML зависит от цели: статистика чаще используется для построения обобщений, тогда как ML — для повышения точности предсказаний [Bzdok, Altman и Krzywinski, 2018, с. 233–234].

Мультиагентные системы (МАС) представляют собой сети взаимодействующих автономных агентов, которые могут сотрудничать для решения сложных задач. В контексте безопасности образовательных организаций они могут быть использованы для управления чрезвычайными ситуациями, такими как пожары, теракты или техногенные аварии.

Применение мультиагентных систем в экстренном реагировании уже применяется в ESARS: A Situation-Aware Multi-Agent System for Real-Time Emergency Response Management [EJ-Compute, б.г.], система использует данные из социальных сетей для управления чрезвычайными ситуациями в реальном времени. Она включает агентов для сбора информации, классификации событий и уведомления заинтересованных сторон. Аналогичный подход может быть применен для мониторинга угроз, таких как пожары или теракты, в школах и университетах.

Принципы, изложенные в A Multi Agent Based Approach for Prehospital Emergency Management, могут быть адаптированы для чрезвычайных ситуаций на территории образовательных учреждений. Принципы многоагентного подхода (автономные сенсоры и аналитические модули, обмен сообщениями в реальном времени, классификация инцидентов, оптимальное распределение реагирующих сил и интеграция с другими системами) умело ложатся на задачу обеспечения физической безопасности образовательных учреждений [Safdari и др., 2017, с. 171–178].

Так как инфраструктура образовательных учреждений относится к критической то и требования к безопасности отдельных ее элементов должны быть соответствующие. Статья "Security in Multi-Agent Systems" предоставляет общую модель безопасности для МАС, включая категоризацию угроз (атаки, уязвимости) и требования к безопасности на уровне агентов и системы. Ключевые моменты могут быть использованы для разработки безопасных МАС в образовательной среде, например, для защиты от кибератак на системы безопасности [ScienceDirect, б.г.]. Статья "Designing a safety management system for higher education centers" описывает разработку системы управления безопасностью для высших учебных заведений, включая аспекты пожарной безопасности, здоровья и IT [PMC, б.г.].

Предложенные системы должны соответствовать государственному стандарту ГОСТ Р 70950-2023: Технологии ИИ в образовании [ГОСТ Р 70950-2023, 2023] в котором указаны меры безопасности, такие как защита персональных данных и обеспечение информационной безопасности в государственных информационных системах (в разделах 2.4, 2.6, 2.7). В Федеральный закон № 149-ФЗ: Информационная безопасность установлены общие требования к информационной безопасности, которые должны быть применены к системам ИИ, включая мультиагентные, для защиты данных и предотвращения несанкционированного доступа [Федеральный закон №149-ФЗ, 2006].

Архитектура решения

В рамках исследования была задача представить архитектуру решения, которая помогает обеспечивать безопасность образовательной организации реагируя на события умных устройств, анализируя журналы действий, как в организации, так и внутри системы и автоматизировано выявляя угрозы с помощью моделей искусственного интеллекта и использование данной системы в курсе «Проектирование информационных интеллектуальных систем», использующемся при подготовке студентов бакалавриата в Пермском Педагогическом университете.

В ходе проектирования были выделены следующие требования к системе: модульность, поддержка независимого и параллельного выполнения агентов, интеграция с внешними CRM-системами и физическими компонентами, возможность расширения за счет подключения новых модулей без изменения инфраструктуры или модифицирования основной архитектуры системы.

Мультиагентный подход позволил добиться следующих архитектурных характеристик: распределенность (компоненты системы физически отделены друг от друга), отказоустойчивость, гибкость и масштабируемость (новые агенты и интеграции добавляются без изменения центральной логики, а логика работы конкретного агента или интеграция с внешней системой локализуются).

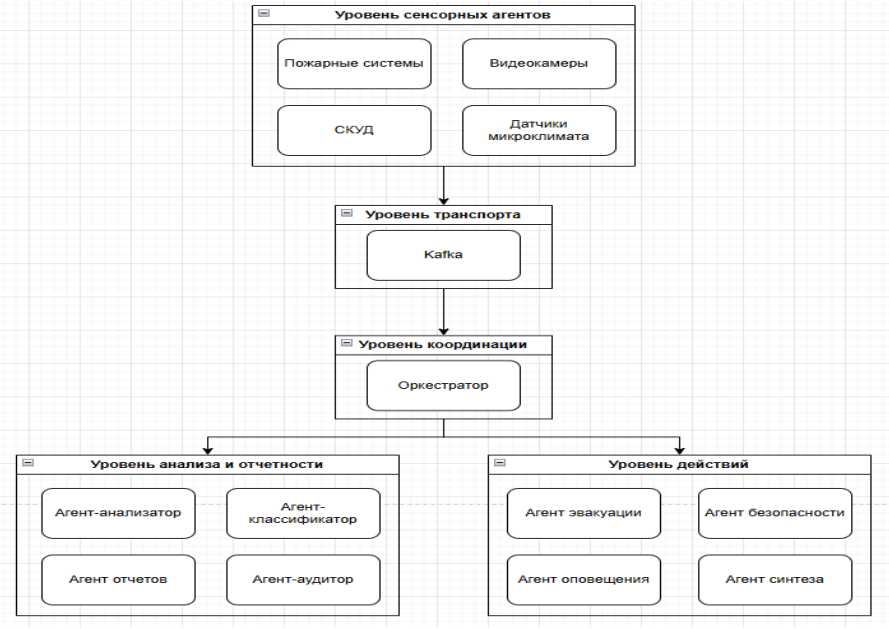

Рисунок 1 - Архитектура системы

Рассмотрим архитектуру компонентов по нескольким уровням:

1. Самый низкий уровень – это уровень сенсорных агентов. Тот уровень, который генерирует события. То есть, это датчики дыма, температуры, микроклимата, движения, видеокамеры.

В случае с физическими датчиками – это простейшие агенты, которые могут делать локальные выводы, например появилось движение или задымление.

В случае с видеокамерами – они передают видеопоток, который обрабатывается сервисом видеоаналитики, откуда уже поступают события. Если это продвинутые сервисы аналитики, они могут выявлять скрытое оружие и настроение людей.

2. Транспортный уровень – это уровень, где происходит коммуникация и маршрутизация. Все агенты обмениваются сообщениями между собой через общую шину данных, в качестве которой выступает решение с открытым исходным кодом Apache Kafka. Kafka передает сообщения следуя общепринятому протоколу. В качестве протокола выступает JSON с строгой типизацией – текстовый формат обмена структурированными данными. Строгое следование формату и типам данных – это обязательная часть функционирования системы. Агенты могут беспрепятственно подключаться к системе на основании этого фактора.
3. Уровень координации — это уровень с централизованным агентом оркестрации. Оркестратор управляет выполнением агентов через загружаемую структуру данных, которая представлена в памяти графом состояний (он же сценарное дерево) для удобства как работы с ним, так и представления данных для пользователей. Граф определяет

порядок вызова агентов и может изменяться в зависимости от условий запроса, что позволяет оптимизировать время и полноту ответа и повышать гибкость обработки.

4. Уровень действий – это уровень с агентами исполнителями. На данный момент выделены следующие агенты: Агент эвакуации (открывает двери, управляет световыми табло), Агент оповещения (генерирует push-уведомления, делает рассылки), Агент синтеза (синтезирует речь для оповещений), Агент безопасности (выявляет угрозы и отправляет их охране на подтверждение инцидента).
5. Уровень анализа и отчетности – это самый верхний уровень, результаты работы которого видят конечные потребители информации. Агент-анализатор получает срез событий и формирует описание ситуации. Агент-классификатор определяет тип и уровень угрозы, Агент отчетов формирует отчет, по которому видно, что произошло и какие действия рекомендованы. Агент-аудитор проверяет на ложное срабатывание.

Граф взаимодействия агентов может изменяться, в зависимости от количества подключенных и доступных агентов, а также в зависимости от уровня и типа угрозы. Далее рассмотрим типовой сценарий работы.

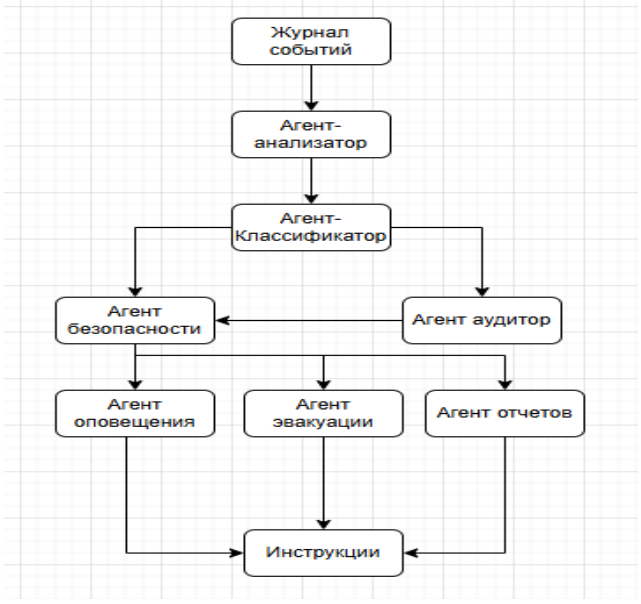

Рисунок 2 **- Граф работы агентов**

Датчики дыма фиксируют превышение порога. Начинают генерировать события о срабатывании. События с датчиков через шину данных попадают в оркестратор. Оркестратор запускает агента-анализатора, который регистрирует угрозу в журнале угроз. Агент-классификатор выявляет локализованный пожар и классифицирует высокий уровень угрозы. Агент безопасности формирует отчет и отправляет охране на подтверждение инцидента.

После подтверждения запускаются агенты оповещения и эвакуации, которые управляют движением потока людей делая рассылку через мессенджеры, открывая двери, обесточивая питание в области пожара, включая указатели и громкоговорители.

Также агент отчетов формирует отчет по инциденту. Пример отчета: «Обнаружено задымление в аудитории А306. Предприняты меры по эвакуации. Угроза классифицирована как средняя. Причина: перегрев оборудования. Рекомендовано отключить питание в западном крыле.»

В ходе тестирование получилось добиться точности порядка 85% с высоким количеством ложных срабатываний. Ложные срабатывания могли происходить как от физических причин: пыль, перегрев датчиков, ошибки в калибровке, потери связи, так и от работы агентов: неправильная трактовка событий, классификация нетипичного поведения как угроз. Данные инциденты могли вести к негативным последствиям: подрыв доверия к системе, неоправданная эвакуация, штрафы за вызовы спецслужб, нарушение учебного процесса, стресс обучающихся.

Добавление агента-аудитора, который обучался на истории тревог и учился фильтровать по признакам ложных срабатываний. Агент проверял обоснованность тревог на основе текстовых логов, сигналов других агентов, датчиков поблизости и исторического контекста. При несоответствии — сигнал классифицируется как возможное ложное срабатывание. С этим агентом получилось избавиться от большей части ложных сигналов и увеличить точность обработки до 95% при нагрузке более ста сигналов в секунду.

Перспективы

В дальнейшем развитие предложенной мультиагентной архитектуры может быть направлено на более глубокую интеграцию с уже существующими в образовательных учреждениях системами видеонаблюдения и контроля доступа. Создание специализированных адаптеров для API популярных VMS- и СКУД-платформ позволит использовать уже развернутую инфраструктуру камер и турникетов, а также получать синхронизированные метаданные о перемещении персонала и посетителей без необходимости дополнительного аппаратного развертывания.

Не менее значимой представляется задача внедреия механизмов обучения моделей классификации, что позволит агентам-аудиторам постепенно повышать свою точность на основе обратной связи операторов и анализа реальных инцидентов, сохраняя при этом конфиденциальность данных при распределённой обработке между несколькими учреждениями.

Вопросы обеспечения информационной безопасности межагентного взаимодействия должны решаться путём внедрения протоколов TLS/mTLS и ролевых моделей доступа с подробным аудитом логов, что позволит минимизировать риски несанкционированного вмешательства и подмены данных [Pronichev и Vitkova, 2019]. С практической точки зрения важным направлением является создание корпоративных веб- и мобильных интерфейсов для диспетчеров и ответственных сотрудников, объединяющих визуализацию карты угроз в реальном времени, статусы реагирования и протоколы эвакуации, а также поддерживающих интеграцию с мессенджерами и системами автоматизированного голосового оповещения. Разработка руководств пользователя и другой сопровождающей документации увеличит темпы внедрения разработки в организациях.

Эти этапы развития обеспечат масштабируемость и адаптивность архитектуры не только в образовательной среде, но и в других критически значимых объектах.

Выводы

Проведённое исследование подтвердило, что разработанная мультиагентная система с адаптивной графовой архитектурой и динамической активацией агентов эффективно решает задачи физической безопасности образовательных учреждений, оперативно реагируя на

события разного уровня сложности — от срабатывания датчиков дыма до анализа видеопотока. Централизованный оркестратор и модульная структура на базе шины данных Apache Kafka обеспечивают отказоустойчивость, гибкость расширения и высокую скорость обработки инцидентов. Введение специализированных агентов-аудиторов, обученных на реальных логах, позволило снизить долю ложных срабатываний до уровня менее 5 % при высокой нагрузке, что повышает доверие к системе и снижает риски неоправданных эвакуаций. Результаты работы имеют существенное практическое и образовательное значение: они могут быть внедрены в реальные программные продукты для защиты вузов и школ, а также использоваться в учебных курсах и методических материалах для подготовки студентов и повышения квалификации преподавателей в области проектирования интеллектуальных систем.

Обсуждение

Распределённая мультиагентная архитектура превосходит единичные решения в задачах обеспечения безопасности, полученные результаты подтверждают, что разделение системы на уровни сенсорных агентов, транспортной шины данных и централизованного оркестратора обеспечивает одновременно модульность, отказоустойчивость и высокую скорость обработки инцидентов. Это согласуется с выводами из статьи: A Situation-Aware Multi-Agent System for Real-Time Emergency Response Management, где подчёркивается важность гибкой маршрутизации сообщений между агентами для быстрого реагирования на угрозы [EJ-Compute, б.г.].

Внедрение специализированного агента-аудитора, обученного на исторических логах, позволит снизить количество ложных срабатываний, по результатам внедрения удалось добиться: точность обработки событий выросла до 95 % при высокой нагрузке, что соотносится с рекомендациями из исследований по применению ML-моделей для фильтрации аномалий в мультиагентных системах. Это иллюстрирует синергетический эффект комбинирования статистических методов и машинного обучения, о котором говорится в обзоре литературы.

Применение брокеров и строгая типизация сообщений упрощают масштабирование и интеграцию с внешними CRM- и физическими системами. Добавление новых агентов и адаптеров к существующей инфраструктуре в ходе тестирования не требовало изменений в ядре платформы. В согласовании с требованиями ГОСТ Р 59277-2020 к модульным системам ИИ [ГОСТ Р 59277-2020, 2020].

Важность образовательного аспекта, подкрепляется описанной архитектурой и методами её оптимизации, которые могут быть использованы в учебных курсах по проектированию интеллектуальных систем, а результаты тестирования — в практических лабораторных работах. Таким образом, проверенные гипотезы демонстрируют эффективность разработанного решения в контексте безопасности, и подтверждают его значимость для подготовки профессионалов в сфере IT и методического сопровождения образовательных организаций.

Библиография

1. Guo T., Chen X., Wang Y., Chang R., Pei S., Chawla N. V., Wiest O., Zhang X. Large Language Model based Multi-Agents: A Survey of Progress and Challenges [Электронный ресурс] // arXiv preprint arXiv:2402.01680v2; revised 19 апреля 2024. – DOI: 10.48550/arXiv.2402.01680. – Режим доступа URL: https://arxiv.org/pdf/2402.01680v2. – Дата обращения: 24.04.2025.
2. ГОСТ Р 59277-2020: Классификация систем искусственного интеллекта, раздел 6, пункт 7.1 [Электронный ресурс] // AllGOSTs.ru. – Режим доступа URL: https://allgosts.ru/35/020/gost_r_59277-2020. – Дата обращения:

19.05.2025.
3. Bzdok D., Altman N., Krzywinski M. Statistics versus Machine Learning [Электронный ресурс] // Nature Methods. 2018. Vol. 15, pp. 233–234. – hal-01723223. – Режим доступа URL: https://hal.science/hal-01723223. – Дата обращения: 20.04.2025.
4. ESARS: A Situation-Aware Multi-Agent System for Real-Time Emergency Response Management [Электронный ресурс] // EJ-Compute. – Режим доступа URL: https://www.ej-compute.org/index.php/compute/article/view/83. – Дата обращения: 19.05.2025.
5. Safdari R., Shoshtarian M., Mohammadzadeh N., Shahraki A. D. A Multi Agent Based Approach for Prehospital Emergency Management [Электронный ресурс] // Bulletin of Emergency & Trauma. 2017. Vol. 5, No 3, pp. 171–178. – PMCID: PMC5547204. – Режим доступа URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC5547204/. – Дата обращения: 19.05.2025.
6. Security in Multi-Agent Systems [Электронный ресурс] // ScienceDirect. – Режим доступа URL: https://www.sciencedirect.com/science/article/pii/S1877050915023972. – Дата обращения: 19.05.2025.
7. Designing a Safety Management System for Higher Education Centers [Электронный ресурс] // PMC. – Режим доступа URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC6745889/. – Дата обращения: 19.05.2025.
8. ГОСТ Р 70950-2023: Технологии ИИ в образовании [Электронный ресурс] // ГАРАНТ. – Режим доступа URL: https://base.garant.ru/408052199/. – Дата обращения: 19.05.2025.
9. Федеральный закон № 149-ФЗ «Об информации, информационных технологиях и о защите информации» [Электронный ресурс] // КонсультантПлюс. – Режим доступа URL: https://www.consultant.ru/document/cons_doc_LAW_61798/. – Дата обращения: 19.05.2025.
10. Pronichev A. P., Vitkova L. A. MODELING OF MULTIAGENT SYSTEM OF CYBER-PHYSICAL DEVICES FOR SOLVING PROBLEM MANAGEMENT AND SECURITY CONTROL PROBLEMS [Электронный ресурс] // International Research Journal. 2019. № 9 (87). – DOI: 10.23670/IRJ.2019.87.9.003. – Режим доступа URL: https://research-journal.org/en/archive/9-87-2019-september/modelirovanie-multiagentnoj-sistemy-kiber-fizicheskix-ustrojstv-dlya-resheniya-problem-upravleniya-i-kontrolya-bezopasnosti-perimetra. – Дата обращения: 19.05.2025.

Multi-Agent Systems in Ensuring Physical Security of Educational Institutions

Nikita S. Sharonov

Lead Development Specialist,
Academiya Biznes Resheniy LLC,
Master's Student,
Perm State Humanitarian Pedagogical University,
614990, 24 Sibirskaya str., Perm, Russian Federation;
e-mail: 69224@pspu.ru

Sergei A. Boyarko

Lead Development Specialist,
UNICORN LLC,
Master's Student,
Perm State Humanitarian Pedagogical University,
614990, 24 Sibirskaya str., Perm, Russian Federation;
e-mail: 69216@pspu.ru

Abstract

A novel multi-agent architecture for ensuring physical security in educational institutions is proposed and substantiated, incorporating an adaptive graph orchestrator and dynamic activation of

specialized agents. A multi-layered structure ensuring fault tolerance and ease of scaling is described. The process of implementing an ML audit agent is outlined. The practical significance of the work lies in the direct applicability of the developed solutions for protecting schools and universities. The proposed architecture and testing methodologies can be used by university lecturers teaching IT students in courses on designing intelligent systems. The research materials can be utilized by educational institution specialists involved in implementing innovative technologies into security infrastructure.

For citation

Sharonov N.S., Boyarko S.A. (2025) Multiagentnyye sistemy v obespechenii bezopasnosti obrazovatel'nykh uchrezhdeniy ot fizicheskikh ugroz [Multi-Agent Systems in Ensuring Physical Security of Educational Institutions]. *Pedagogicheskii zhurnal* [Pedagogical Journal], 15 (6A), pp. 14-22.

Keywords

Educational institutions, multi-agent systems, security, artificial intelligence, threat counteraction.

References

1. Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N.V., Wiest, O. and Zhang, X., 2024. Large Language Model based Multi-Agents: A Survey of Progress and Challenges. arXiv preprint. [online] Available at: https://arxiv.org/pdf/2402.01680v2 [Accessed 24 Apr. 2025]. DOI: 10.48550/arXiv.2402.01680.
2. GOST R 59277-2020, Classification of Artificial Intelligence Systems, Section 6, Clause 7.1. [online] Available at: https://allgosts.ru/35/020/gost_r_59277-2020 [Accessed 19 May 2025].
3. Bzdok, D., Altman, N. and Krzywinski, M., 2018. Statistics versus Machine Learning. Nature Methods, 15, pp.233–234. [online] Available at: https://hal.science/hal-01723223 [Accessed 20 Apr. 2025].
4. EJ-Compute, ESARS: A Situation-Aware Multi-Agent System for Real-Time Emergency Response Management. [online] Available at: https://www.ej-compute.org/index.php/compute/article/view/83 [Accessed 19 May 2025].
5. Safdari, R., Shoshtarian, M., Mohammadzadeh, N. and Shahraki, A.D., 2017. A Multi Agent Based Approach for Prehospital Emergency Management. Bulletin of Emergency & Trauma, 5(3), pp.171–178. [online] Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC5547204/ [Accessed 19 May 2025].
6. ScienceDirect, Security in Multi-Agent Systems. [online] Available at: https://www.sciencedirect.com/science/article/pii/S1877050915023972 [Accessed 19 May 2025].
7. PMC, Designing a Safety Management System for Higher Education Centers. [online] Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC6745889/ [Accessed 19 May 2025].
8. GOST R 70950-2023, AI Technologies in Education. [online] Available at: https://base.garant.ru/408052199/ [Accessed 19 May 2025].
9. Federal Law No. 149-FZ, On Information, Information Technologies and the Protection of Information. [online] Available at: https://www.consultant.ru/document/cons_doc_LAW_61798/ [Accessed 19 May 2025].
10. Pronichev, A.P. and Vitkova, L.A., 2019. Modeling of Multiagent System of Cyber-Physical Devices for Solving Problem Management and Security Control Problems. International Research Journal, (9)87. DOI: 10.23670/IRJ.2019.87.9.003. [online] Available at: https://research-journal.org/en/archive/9-87-2019-september/modelirovanie-multiagentnoj-sistemy-kiber-fizicheskix-ustrojstv-dlya-resheniya-problem-upravleniya-i-kontrolya-bezopasnosti-perimetra [Accessed 19 May 2025].