УДК 004.056

М.М. БАРИНОВА[1,2]
Научный руководитель – доцент Г.П. ГАВДАН[1]
[1]*Национальный исследовательский ядерный университет «МИФИ», Москва*
[2]*Федеральная таможенная служба, Москва*

# ЗАЩИТА ИНФОРМАЦИИ В ГИС С ИСПОЛЬЗОВАНИЕМ СЕРВИСОВ ИИ: НОРМАТИВНО-ТЕХНИЧЕСКИЙ АНАЛИЗ И ПЕРСПЕКТИВЫ

В работе проведен анализ проблем и методов защиты информации в государственных информационных системах (ГИС), интегрирующих сервисы искусственного интеллекта (ИИ). Особое внимание уделено противоречию между стратегической важностью ИИ для цифровой трансформации и рисками, связанными с использованием иностранных программных компонентов. Предложены ключевые направления для развития системы защиты информации.

## Введение

Цифровая трансформация государственного управления, стимулируемая такими документами, как «Национальная стратегия развития искусственного интеллекта до 2030 года», делает внедрение ИИ в ГИС обязательным (imperative). Однако интеграция интеллектуальных систем, особенно с имеющейся зависимостью от иностранного программного обеспечения (ПО), создает новые векторы атак и риски утечек конфиденциальных данных. Это требует разработки специализированных средств и методологий защиты.

## Ключевые проблемы и анализ литературы

Анализ работ [1–3] выявил основные исследовательские фокусы в области безопасности ИИ:

– таксономия ИИ-атак и разработка методов противодействия (adversarial attacks);

– риски использования ПО с открытым исходным кодом в жизненном цикле моделей ИИ;

– пробелы в нормативном регулировании и оценке соответствия ИИ-сервисов.

На примере сервиса автоматического анализа изображений в российских таможенных органах выявлена фундаментальная проблема: недостаточность чисто технических мер защиты без комплексного методологического и нормативного обеспечения на государственном уровне.

**Перспективные направления обеспечения безопасности**

На основе проведенного анализа предлагается сконцентрироваться на следующих ключевых мерах:

– *защита данных*: строгое шифрование данных, как при хранении, так и при передаче (с использованием TLS/SSL);

– *контроль доступа*: внедрение многофакторной аутентификации и ролевой модели доступа для минимизации рисков несанкционированного доступа;

– *защита моделей ИИ*: применение методов обфускации алгоритмов для противодействия реверс-инжинирингу (например, регулярное тестирование моделей на устойчивость к adversarial-атакам);

– *повышение осведомленности*: обязательное обучение персонала лучшим практикам кибербезопасности;

– *обеспечение целостности*: исследование применения блокчейн-технологий для верификации данных и отслеживания изменений в системе.

**Выводы**

Несмотря на сформировавшуюся терминологию, практическое применение безопасных ИИ-сервисов в ГИС остается нетривиальной задачей. Технические меры защиты (шифрование, аутентификация) необходимы, но недостаточны. Ключевой проблемой является отсутствие развитой системы отечественных стандартов и методик оценки безопасности и надежности ИИ. Требуется развитие комплексного подхода, объединяющего технические решения, нормативное регулирование и методологическую поддержку государственных органов. Это позволит обеспечить не только информационную безопасность, но и доверие, надежность и устойчивость критически важных ИИ-сервисов.

*Список литературы*

1. Костогрызов А.И., Нистратов А.А. Анализ угроз злоумышленной модификации модели машинного обучения для систем с искусственным интеллектом // Вопросы кибербезопасности. 2023. № 5 (57). DOI: http://dx.doi.org/10.21681/2311-3456-2023-5-9-24.

2. Марков, А. С. Важная веха в безопасности открытого программного обеспечения / А.С. Марков // Вопросы кибербезопасности. – 2023. – № 1(53). – С. 2–12. – DOI 10.21681/2311-3456-2023-1-2-12. – EDN OHYLTR.

3. Арустамян, Сас С.; Варениса, Виталий В.; Марков, Алексей С. Методические и реализационные аспекты внедрения процессов разработки безопасного программного обеспечения. Безопасность информационных технологий, [S.l.], v. 30, n. 2, p. 23–37, мая 2023. ISSN 2074-7136. Доступно на: <https://bit.spels.ru/index.php/bit/article/view/1499>. (Дата доступа: 24 окт. 2025) DOI: http://dx.doi.org/10.26583/bit.2023.2.01.