ФЕДЕРАЛЬНЫЙ ЦЕНТР
ПРИКЛАДНОГО РАЗВИТИЯ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

ПРОБЛЕМЫ И ВЫЗОВЫ ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ПРИ ВНЕДРЕНИИ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРОМЫШЛЕННОСТИ

Безопасность при внедрении ИИ и высокая стоимость ошибок

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Вызовы внедрения ИИ

Устаревшая инфраструктура предприятия

Данный вызов особенно актуален для предиктивного обслуживания станков и иного оборудования, где ошибка может привести к дорогостоящему ремонту.

Решение:
Модернизация инфраструктуры – является важным фактором, который влияет на качество собираемых данных.

Угрозы кибербезопасности

Уязвимости в оборудовании и цифровая связанность ИИ с сетевой инфраструктурой предприятия, к которым относятся:

- атака на данные;
- взлом моделей ИИ;
- атака на IoT-устройства;
- шифрование данных и блокировка систем

Для минимизации рисков предприятиям приходится проводить самостоятельные тестирования и валидацию ИИ систем, что увеличивает время на внедрение технологий.

Сопротивление персонала

Страх потери рабочих мест и отсутствие востребованных программ и/или подготовительных курсов.

Решение:
Гибридный подход использования технологий ИИ. Включение программ обучения и/или программ переподготовки в ВУЗах. Повышение ценности знаний работы с технологиями ИИ может стимулировать интерес молодых специалистов к специальностям инженерной направленности и как следствие ускорить процесс внедрения технологий ИИ.

Правовое регулирование

Регуляторный режим в Российской Федерации часто полагается на этические принципы, что в свою очередь делает его менее строгим в отличии от иных сфер деятельности.

Решение:
Частичной компенсацией нынешних вызовов является пилотный проект – предприятия начинают с пилотных проектов, постепенно масштабируя успешные решения, поскольку зачастую интеграторы таких решений предлагают предоставление в том числе и оборудования в части компромисса и обучение сотрудников работы с ним.

Уровень информационной безопасности

ГИСП

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Цифровой паспорт промышленного предприятия

Расчет уровня цифровизации посредством анализа блока ИИ осуществляется как в **индивидуальном порядке** (для предприятия промышленности), так и **в призме отрасли** (для регулятора) и ориентирован на определение соотношения используемых технологий ИИ на предприятиях промышленности.

Наиболее популярные технологии ИИ:

- Оптическое распознавание символов – **818** внедрений
- Обработка естественного языка – **221** внедрение
- Компьютерное зрение – **177** внедрений
- Перспективные методы ИИ – **122** внедрений
- Интеллектуальная поддержка принятия решений – **100** внедрений
- Распознавание и синтез речи – **84** внедрения

Общее количество: 1 522 внедрений – индекс 11 %

Средства защиты информации – индекс 58 %

Роль ИБ в такой методологии – убедиться в наличии технологий для защиты данных, включая криптографические, аппаратные, программные средства защиты информации, а также электронную подпись и биометрию.

Совокупность этих показателей, а также регулярность их актуализации предприятиями промышленности способствует выявлению белых пятен цифровизации и обеспечения ИБ как для отраслей, так и для предприятия промышленности.

Риски, связанные с использованием злонамеренно измененных датасетов

Отравление данных – это процесс, при котором злоумышленники внедряют в обучающие данные искаженную или вредоносную информацию, что приводит к снижению точности и надежности моделей ИИ.

Последствия отравления данных:

1. Некорректные прогнозы в предиктивном обслуживании
2. Утечка конфиденциальной информации

Угроза безопасности

Снижение точности и надежность систем ИИ

Данный вид атак на датасеты и ML-модели опасен, для защиты от него необходима разработка специальных стратегий, которые включают в себя:

- ✓ строгий контроль за источниками информации
- ✓ фильтрацию данных для обучения
- ✓ проверку моделей на регулярной основе.

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Для классификации определения атаки путем отравления данных, в настоящее время выявлен шаблонный сценарий:

Определение цели атаки

↓

Проведение исследования уязвимостей системы и создание ядовитых данных

↓

Внедрение отравленных данных

↓

Обучение модели на отравленных данных

↓

Модель начинает выдавать некорректные результаты
и в худшем случае активировать backdoor-функции

Особенности работы в условиях повышенных требований к ИБ

Регулирование ИБ осуществляется ФСБ России и ФСТЭК России.

Строгие требования к ИБ в закрытом сегменте обусловлены необходимостью обеспечить защиту критически важных данных и систем от кибератак, однако такие требования часто выступают в противоречие с потребностями внедрения современных технологий ИИ, которые в свою очередь требуют гибкости, масштабируемости и доступа к большим объемам данных..

Внедрение ИИ в закрытом контуре требует комплексного подхода, включающего:

- ✓ создание защищенной инфраструктуры;
- ✓ использование отечественных решений;
- ✓ обучение персонала и сотрудничество с регуляторами.

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Основные сложности:

- фрагментированные данные, которые хранятся в изолированных системах;
- специфичность систем, не позволяющая обеспечить поддержку технологий ИИ;
- запрет на использование публичных облаков, что ограничивает возможности масштабирования ИИ.

Пути адаптации таких предприятий к рынку:

- ✓ внедрение отечественных сертифицированных решений и использование криптографических методов для защиты данных;
- ✓ разработка локальных ИИ-платформ;
- ✓ обучение моделей на синтетических данных и применение аугментации данных для повышения качества моделей.

Использование локальных решений для обеспечения безопасности и эффективности работы ИИ в промышленных контурах

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Локальные решения позволяют развертывать технологии ИИ в изолированной среде, что минимизирует риск утечки данных и обеспечивает соответствие требований регуляторов, поскольку обеспечивают безопасность.

Примеры локальных решений:

Интеллектуальный автоматизированный прогностический комплекс оборудования
Разработчик: **РТ-Техприемка**

Коробочное решения для NLP и CV, адаптированные для промышленных предприятий
Разработчик: **Sber AI**

Вызовы внедрения локальных решений:

- **Оценка и при необходимости, модернизация локальных вычислительных ресурсов и их развертывания;**
- **интеграция ИИ-решений с существующими системами (ERP, SCADA) увеличивает сложность и сроки внедрения;**
- **Нехватка квалифицированных специалистов.**

Оптимальные пути решения

- ✓ **Поэтапная модернизация вычислительных ресурсов;**
- ✓ **использование открытых стандартов упрощает интеграцию технологий ИИ с существующими системами;**
- ✓ **обучение персонала**
- ✓ **сотрудничество с интеграторами.**

Вопросы безопасности при работе с промданными и их интеграцией в системы ИИ

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Необходима разработка единых стандартов и нормативно-правовой базы, регулирующих обработку, унификацию, стандартизацию и обмен промышленными данными.

В настоящее время на платформе ФГАУ «ФЦПР ИИ» уже существует инструмент, ориентированный на обмен такой информации, однако в отсутствие нормативно-правовой регуляторики – решение всех проблем полной мере пока невозможно.

Решение: простые ML-модели, для оптимизации рутинных задач.

Внедрение нормативно-правовой базы, регулирующей промданные позволит:

- Ускорять разработку или адаптацию Ml-моделей при соблюдении требований предъявляемых к защите данных
- Стимулировать развитие инноваций и доступ к ним, в том числе и для предприятий, работающих в закрытом контуре

Отечественные решений в сфере информационной безопасности для работы с ИИ-системами

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

№ п/п	Компания	Тип решения	Использование технологий ИИ	Защита от атак на ИИ	Примечание
1	**СерчИнформ**	Продукты для защиты конфиденциальной информации и предотвращения инсайдерских угроз	Анализ больших данных для выявления аномалий, использование водяных знаков, контроль монитора АРМ, в том числе от фотографий на телефон	Защита от атак на данные и модели ИИ	Промышленные предприятия, госорганизации
2	**Ростех**	Системы мониторинга и предотвращения атак	Анализ сетевого трафика, выявление киберугроз в реальном времени	Защита IoT-устройств и моделей ИИ, предотвращение атак на предиктивные модели	Оборонная промышленность, КИИ, промышленные предприятия
3	**Лаборатория Касперского**	Комплексные решения для защиты от киберугроз (антивирусы, системы предотвращения атак)	Обнаружение новых видов malware, анализ поведения вредоносных программ	Обнаружение и блокировка атак на данные и модели ИИ, защита от отравления данных и атак на IoT-устройства	Корпоративный сектор, госучреждения, частные пользователи, предприятия промышленности, КИИ, оборонная промышленность

решения с использованием ИИ vs Решения для защиты ИИ-систем

ФЕДЕРАЛЬНЫЙ ЦЕНТР ПРИКЛАДНОГО РАЗВИТИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Критерий	Решения с использованием ИИ	Решения для защиты ИИ-систем
Цель решения	Повышение уровня безопасности за счет анализа данных и выявления угроз с помощью ИИ	Защита ИИ-систем от атак, таких как data poisoning, взлом моделей и атаки на IoT-устройства
Примеры решений	▪ «Лаборатория Касперского» - анализ вредоносного ПО ▪ «Код Безопасности» - анализ поведения пользователей	▪ «Ростех» - защита IoT и моделей ИИ
Технологии	▪ Машинное обучение ▪ Анализ больших данных ▪ Поведенческий анализ	▪ Криптография ▪ Контроль целостности данных ▪ Обнаружение аномалий в моделях ИИ
Преимущества	▪ Автоматизация анализа угроз ▪ Высокая точность обнаружения аномалий	▪ Защита от современных угроз ▪ Обеспечение целостности и конфиденциальности данных

ФЕДЕРАЛЬНЫЙ ЦЕНТР
ПРИКЛАДНОГО РАЗВИТИЯ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

БЛАГОДАРЮ ЗА ВНИМАНИЕ!

Генеральный директор
ФГАУ «Федеральный центр прикладного развития искусственного интеллекта» Министерства промышленности и торговли Российской Федерации

ФЦПРИИ.РФ

t.me/fcprii