Т1 Форум

Безопасность — обратная сторона ИИ-силы

T1

ИИ: новый канал утечек и атак

T1 Форум

Взрыв утечек через ИИ

30 кратный рост за год, сотрудники сами раскрывают тайны LLM

Нет правил – нет защиты

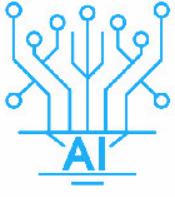
>60% компаний без политики по ИИ, ИИ-сервисы стали «теневым ИТ»

ИИ на вооружении хакеров

Гипер-персонализированный фишинг, дипфейки, эксплойты за минуты

Примеры инцидентов

CISA и инженеры Samsung **случайно** слили данные через чат-бота



Представьте, что ваши данные утекли…
в нейросеть

Как должно быть: ИИ под охраной

T1 Форум

Новые угрозы для моделей

Prompt-инъекции, отравление данных, кража модели, утечки через ответы

Архитектура безопасности ИИ

Изоляция инфраструктуры, Zero Trust доступ, мониторинг запросов

ИИ-шлюз

Фильтрация prompt'ов и ответов (блокирует конфиденциальное, токсичное)

Контроль канала

Шлюз веб-безопасности останавливает загрузку секретов в сеть и фальшивые ответы ИИ

Политики и осведомленность

Правила по ИИ для сотрудников, обучение на кейсах (как не надо делать).

Защитим модели, данные и контекст
по-взрослому

ИИ на страже: умнее, быстрее, проактивнее

T1 Форум

Аналитика и расследования

ИИ разбирает большие логи, находит аномалии и связи за секунды

Детекция угроз 2.0

Поведенческий анализ трафика (ловит новые атаки без сигнатур), ML-антивирусы - реакция на неизвестное

Приоритизация рисков

ИИ оценивает уязвимости и инциденты по критичности, фокус на главном

Автоматизация реагирования

Чат-боты-ассистенты для SOC, плейбуки с ИИ ускоряют ответ (минимум рутины)

Проактивная защита

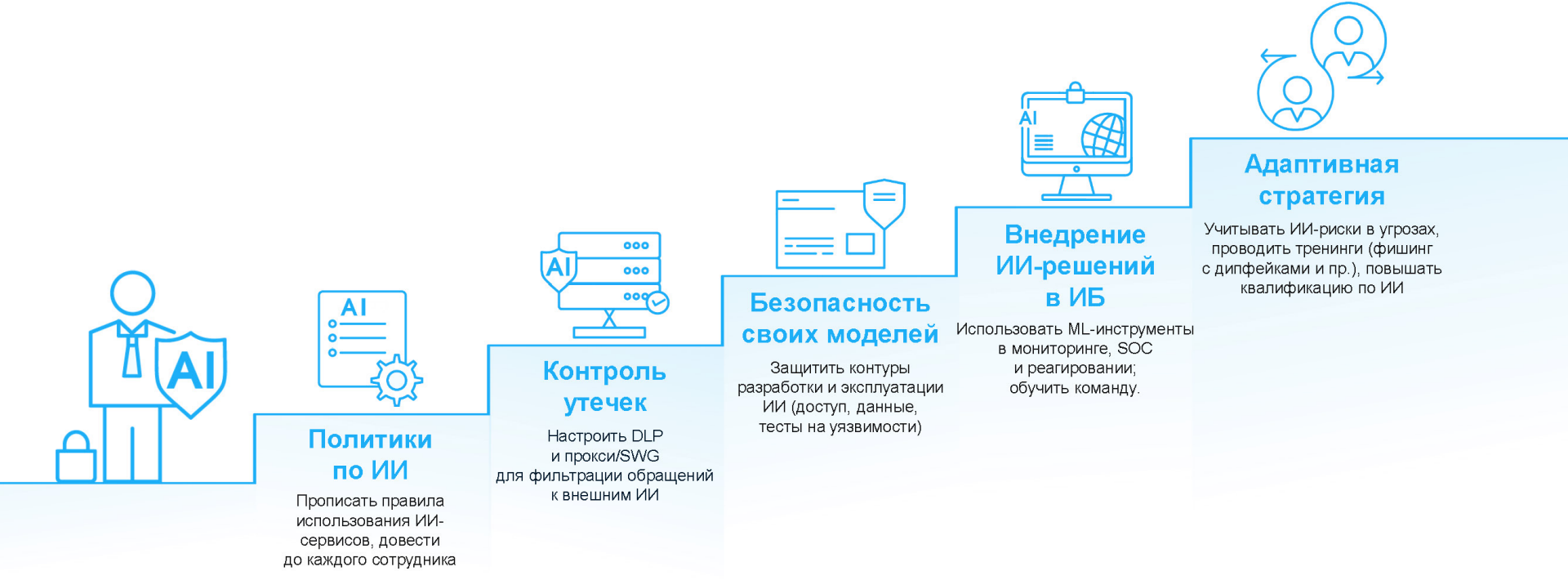
Прогнозирование инцидентов, UEBA выявляет инсайдеров по отклонению от нормы

Применяем ИИ,
чтобы опередить атаки

Безопасный ИИ = Безопасный бизнес

T1 Форум

Мы знаем, что делать



Политики по ИИ

Прописать правила использования ИИ-сервисов, довести до каждого сотрудника

Контроль утечек

Настроить DLP и прокси/SWG для фильтрации обращений к внешним ИИ

Безопасность своих моделей

Защитить контуры разработки и эксплуатации ИИ (доступ, данные, тесты на уязвимости)

Внедрение ИИ-решений в ИБ

Использовать ML-инструменты в мониторинге, SOC и реагировании; обучить команду.

Адаптивная стратегия

Учитывать ИИ-риски в угрозах, проводить тренинги (фишинг с дипфейками и пр.), повышать квалификацию по ИИ

Т1 Форум

Спасибо
за внимание