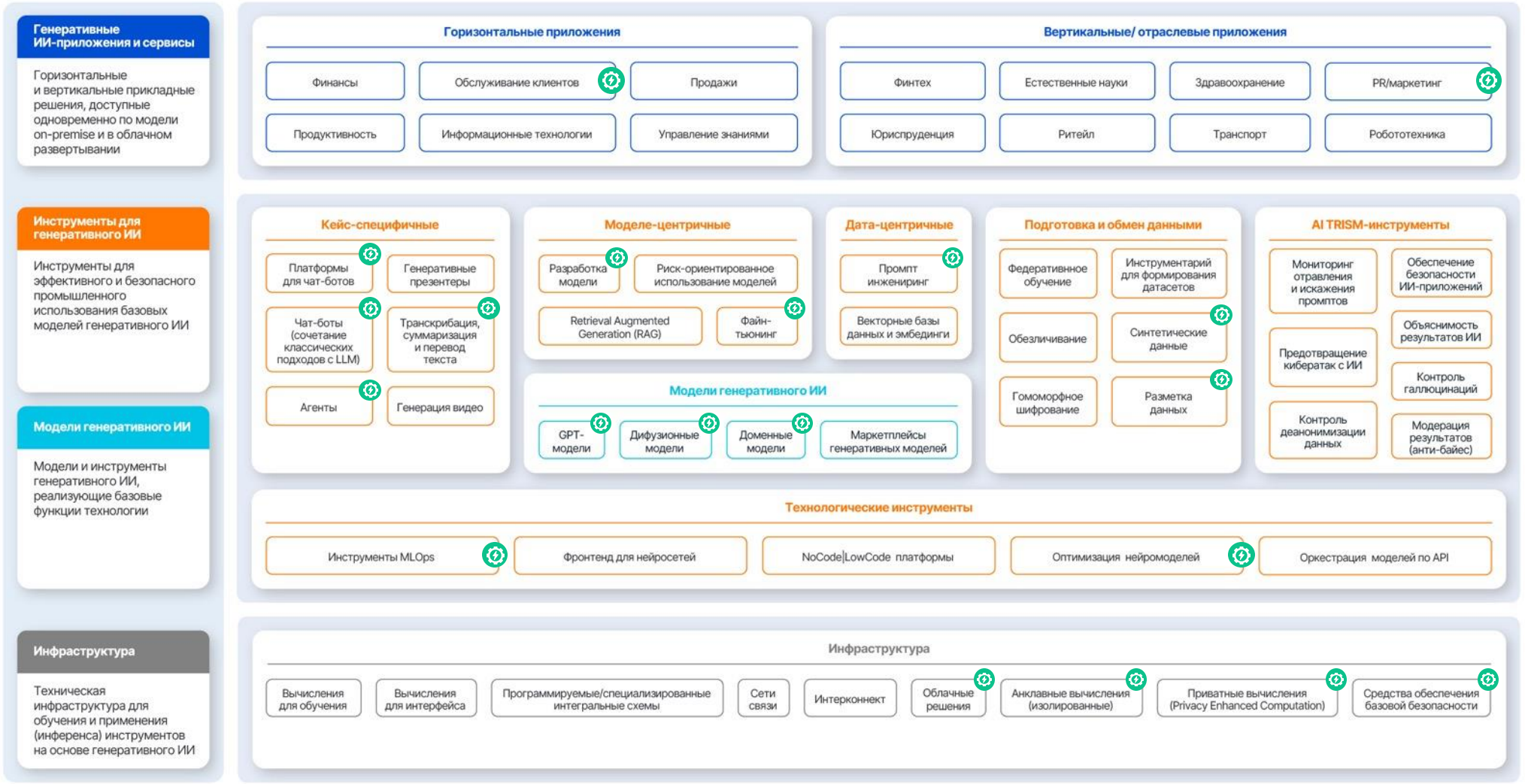
Безопасность ИИ

Установочная встреча РГ

04.09.2024

Архитектура экосистемы генеративного искусственного интеллекта

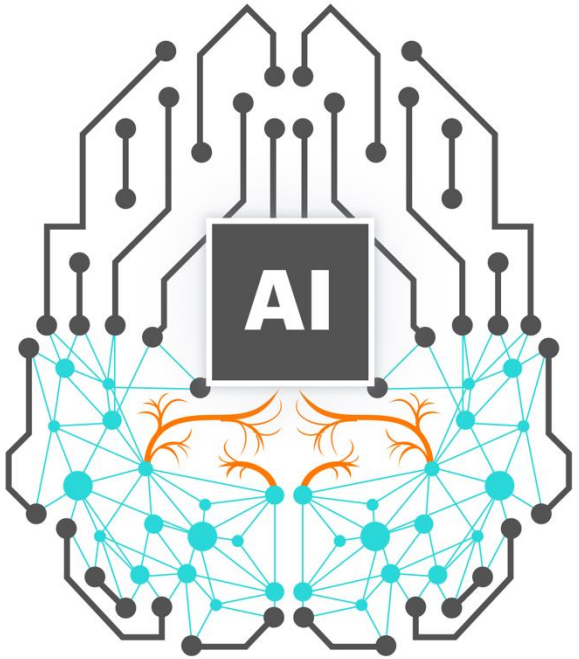
Конкурентноспособные решения



Отсутствуют конкурентоспособные решения в блоке AI TRISM.

ФИНТЕХ АССОЦИАЦИЯ

Жизненный цикл безопасного применения - MLSecOps



Жизненный цикл безопасного применения ИИ - MLSecOps
(включая специфику ГенИИ)

MLSecOps — подход, который объединяет операционные аспекты машинного обучения с вопросами безопасности.
Направлен **на снижение рисков**, которые могут принести модели AI/ML/GenAI в организацию.

Жизненный цикл безопасного применения ИИ - MLSecOps
(включая специфику ГенИИ)

MLSecOps — подход, который объединяет операционные аспекты машинного обучения с вопросами безопасности.
Направлен **на снижение рисков**, которые могут принести модели AI/ML/GenAI в организацию.

Индустриальный фреймворк по ИИ. Для чего?

ФИНТЕХ АССОЦИАЦИЯ

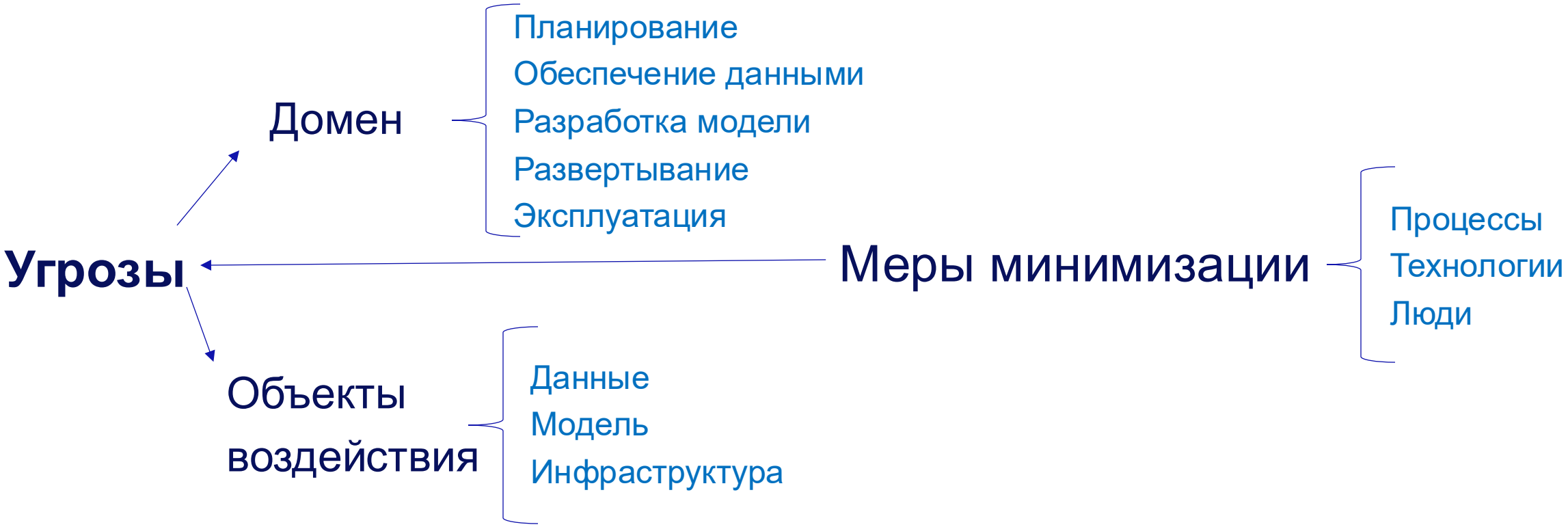
Индустриальный фреймворк безопасного применения ИИ:

- Набор принципов, практик, подходов и методов для минимизации рисков и обеспечения безопасного применения решений на основе ИИ в финтехе.
- Включает в себя **технологические инструменты (Технологии)**, **практики (Процессы)** и **компетенции (Люди)**

Для чего нужен фреймворк:

- консолидация лучшего опыта и практик по минимизации угроз ИИ;
- обеспечение диалога финтеха с технологическими организациями и разработчиками инструментов;
- систематизация подходов к анализу безопасности ИИ-решений.

Индустриальный фреймворк по ИИ. Как выстроен?

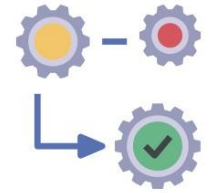


Индустриальный фреймворк по ИИ. Типы мер.

ФИНТЕХ АССОЦИАЦИЯ

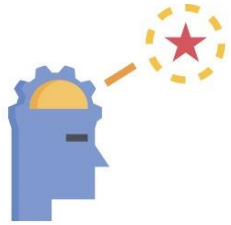
Угрозы

Меры минимизации



ПРОЦЕССЫ

- Внедрение дополнительных процессов проверки качества моделей
- Внедрение процессов контроля за тюнингом модели
- и т.д.



ТЕХНОЛОГИИ

- Внедрение инструментов отслеживания вывода моделей
- Внедрение инструментов выявления галлюцинаций моделей
- и т.д.

ПЕРСОНАЛ

- Повышение экспертизы сотрудника в определенной области
- Внедрение отдельных ролей в ИБ (например, промт-инженер)
- и т.д.

Индустриальный фреймворк по ИИ. Пример

ФИНТЕХ АССОЦИАЦИЯ

Домен	**Угрозы**	**Риски**	**Объект воздействия** (данные, модель, инфраструктура)	**Меры минимизации**		
				Люди	Технологии	Процессы
Планирование						
Обеспечение данными						
Разработка модели						
Развертывание						
Эксплуатация						