

**Программа исследований и разработок по теме
“Кентавр: совместное обучение человека и искусственного интеллекта”**

Цель программы: определить, при каких условиях система, состоящая из человека и ИИ (“кентавр”) может быть эффективнее в решении задачи, чем ИИ, созданный для решения этой задачи; разработать пакет решений по совместному обучению человека и ИИ для создания таких систем; определить практические кейсы, в которых данная системы может применяться

Направления программы

	Субтехнология	Задачи	Связь с продуктовыми решениями
1.	Explainable AI (извлечение кодированных не человеком правил, генерируемых непрозрачными обученными нейронными сетями)	Черный ящик (неинтерпретируемые механизмы принятия решения)	Включение врача на промежуточных этапах при постановке диагноза. Частичное делегирование задач ИИ (анализ рентгеновского снимка), использование совокупности человеческого опыта и рекомендаций ИИ
		Белый ящик (возможно, интерпретируемые механизмы принятия решения)	
2.	Общие знания (совокупность знаний о мире)	Интеграция свойств, используемых человеком при “оценке” объекта	Ускорение процессов обучения ИИ. Возможность обучения на малых данных. Использование накопленного при решении одной задачи опыта для решения другой, аналогичной проблемы
		Внедрение знаний человека в ИИ	
		Трансферное обучение	

3.	Причинно-следственные связи	Выявление связей и факторов, которые влияют на итоговое решение, выдаваемое ИИ	Помощь в освоении и понимании для человека решений, сформированных ИИ в процессе обучения.
		ИИ, выдающий не итог, а основание для принятия данного решения	Система автоматической генерации знаний
4.	Объединение результатов обработки данных ИИ в смешанные датасеты	Создание способов обработки федеративных (объединенных) данных разных типов	Интеграция в системы поддержки принятия решений.
5.	Определение сильных и слабых сторон обработки датасета. Обучение человека их использовать	Формирование вопросов и запросов к человеку в случае нехватки “знаний” в определение объекта и API для учета общих знаний, представленных человеком	Система взаимодействия ИИ и человека
			Помощь в предоставлении машине “общих” знаний
7.	Синтетический интеллект (методология совместной групповой работы человеческого интеллекта и ИИ)	Решение социо-технической проблемы проекта	Развитие и ускорение взаимообучения человека и ИИ
			Социальная адаптивность проекта

8.	Методы сбора размеченных данных	Систематический анализ и сбор данных для дальнейшего внедрения их в СППР	Ускорение процесса обучения ИИ
----	---------------------------------	--	--------------------------------

Область внедрения	Вызов	Задачи	
		Социальная	Техническая
Медицина	Помощь в обнаружении онкологии (редкие случаи)	Преодоление “недоверия” к машинному предсказанию	Выявление причинно-следственной связи редких случаев онкологий, процент точности прогноза ИИ. Создание API для обмена данными между учреждениями здравоохранения
	Определение внутриглазного давления по слепку глаза	Определить ответственность за некорректный прогноз.	Достичь точности определения давления, как минимум, равной точности врачей. Определить случаи, когда отпечаток снят неправильно, и предупредить врача.

	Создание цифрового двойника физического здоровья	Инструмент взаимодействия с психологией человека, как заставить прислушаться к рекомендации “умного” устройства	Создание виртуальной копии, отражающей состояние здоровья. Прогнозирование будущего состояния здоровья при данном образе жизни и создание рекомендаций по улучшению состояния здоровья
Образование	Помощь в понимании используемой технологии	Отработка технологий понимаемого ИИ, объяснение которого воспринимается пользователем	ИИ, выдающий не итог, а основание для принятия данного решения
	Создание личных рекомендаций	Учет достижений по рекомендациям в дальнейших образовательных траекториях (включая учет вузами, школами и доп. образованием результатов движения по индивидуальным траекториям на базе ИИ)	Интеграция разного типа данных, настройка пользовательского выбора на основании чего подбирать рекомендации. Система уточнения через запрос новых пользовательских данных

Вызов	Технологии	Кейсы
Помощь в понимании используемой технологии	Система поддержки принятия решений	Хакатон в рамках чемпионата мира по GO Проведение хакатона с целью выявления причинно-следственных связей при принятии решения
	Interpretable machine learning (Explainable AI white box)	

	Transfer Learning	
Создание личных рекомендаций	Интеграция различных типов данных	Хакатон на лучшее рекомендательное решение
	Transfer Learning	
	Методы сбора неразмеченных данных	
Помощь в обнаружении (лечении) онкологии (редкие случаи)	Transfer Learning	Создание API для обмена данными между медицинскими организациями Хакатон, целью которого является выявление ключевых факторов в редких случаях и интеграция в API запроса на получение данных по всем похожим случаям
	Система поддержки принятия решений	
	Explainable AI	
Создание цифрового двойника физического здоровья	Система поддержки принятия решений	Результаты клинического исследования + поставленный диагноз => вероятно наличие заболевания* Хакатон для обнаружения заболевания*
	Методы сбора неразмеченных данных	