Введение

* познакомить с темой
* актуальность темы
* что определило выбор темы
* чем эта тема интересна именно для Вас
* какова идея исследования
* в чём практическая значимость исследования

литературный обзор = анализ научных трудов

- самостоятельное осмысление материала

- что было сделано исследователями до вас

- что предстоит сделать вам

- почему проблему нужно изучать

Методы

* описательный
* биографический
* культурно-исторический
* сравнительно-исторический
* сравнительно-сопоставительный
* историко-типологический
* историко-генетический
* мифо-поэтический
* социологический
* психологический
* психоаналитический
* структурализм

Цитаты + своё мнение

Заголовок

Тезисный план

Тема: Деградация мышления из-за использования нейронных сетей

- в чём проблема? Почему LLM заставляет деградировать?

История проблемы

- создаёт ли ИИ что-то новое? Создаёт ли человек что-то новое?

Стадии развития ноу-хау: хайп, спад, плато

- угрозы ИИ для образования и социальной сферы

- замена рабочих

Социальный аспект – новинка:

- телефоны

- компьютеры

- печатные машинки

- шариковые ручки

Развитие личности через мелкую моторику – гусиное перо

Как человек думает и когда рождаются идеи?

Что если расслабиться и отпустить угрозу ИИ? Захватит, ну и пусть захватит

Деградация ИИ – коллапс модели

Генеративный – дегенеративный = шутка

Аппокалипсис: ИИ = глупеющий день ото дня слуга

Письмо – сложная мыслительная операция

Чем больше человек пишет собственных текстов, тем яснее он мыслит

Мозг надо тренировать так же, как мышцы

Чем чаще вы нагружаете мозг, тем лучше он работает

Пишущий/непишущий = думающий/недумающий

Умная элита и глупые, косноязычные обыватели

Люди разучатся составлять самостоятельные тексты (ясно мыслить)

- подбор правильных слов

- логичное оформление мысли

- подбор аргументов

- выделение тезисов

- лаконичность

Курсы по ИИ – способ сохранения своего места в профессии

Эффективность = симбиоз человека и ИИ

Ответственность за решение

MES = Manufacturing Execution System

APS = Advanced Planing and Scheduling

Интернет наполняется ерундой

- качество данных для обучения моделей

Огромный объём работы для бенчмарков датасетов

Это проблема, но не крах

AI-тренер

Загрязнение водоёмов – повышаются продажи бутилированной воды

Загрязнение интернета – рост цен на чистые данные

Пропустить через себя

Новый инструмент – для чего подходит, как использовать?

ИИ поедает сам себя

57% контента генерируется ИИ

Pre-training

Утечка смысла

ИИ запоминает наиболее частотные паттерны

ИИ грамотность

Поисковые системы – поиск достоверной информации

Набор данных:  
- научные журналы

- проверенные книги

- авторитетные медиа

=> гранты для писателей, журналистов, художников

Когда AI – основной инструмент в обучении => процесс усвоения знаний упрощается.  
Но:

- хуже запоминают информацию

- трудность к формированию новых навыков

Потеря способности к глубокой концентрации

== Вот здесь можно сделать отсылку к видео, где Клим Жуков рассказывает про хорошие книги 20х годов

Способ использования ИИ: самостоятельно разобраться в теме, потом подключить нейросеть, которая будет выполнять роль доп источника или аналитического инструмента

В этом случае негативного эффекта не возникает, а когнитивная вовлечённость усиливается

Тотальное использование ИИ влиет на способность человека мыслить самостоятельно

Обучение на данных другой модели приводит к коллапсу модели

История развития HomoSapiens – это история эволюции мозга

Как безопасно использовать ИИ, не теряя себя?

Эксперимент:

Человек vs машина

Эссе: собственные знания vs поисковые системы vs человек

Когнитивные функции – методика dDTF

Когнитивные функции:

- восприятие

- обработка Информации

- хранение И

- извлечение И

- память

- внимание

- мышление

- речь

- восприятие

- планирование и организация действий

Перераспределение ментальной нагрузки

ИИ берёт на себя часть задач => мозг адаптируется, работает меньше и теряет навык

Этот эффект накапливается

Первый доступ к ИИ – всплеск мозговой активности

Разделение на думающую элиту и обывателей

- это извечная проблема, которую ИИ может усугубить

Если не пользоваться какими-то функциями, то они отмирают – это и есть деградация

Это избавление от ненужного

Как тогда не деградировать? – Пользоваться этими функциями в другом

Один из способов – выстраивание рабочей деятельности

«Время всегда хорошее» - Евгения Пастернак, Андрей Жвалевский

Мысль: появление роскоши – авторские тексты

Мысль: использовать в работе принцип «составление из частей». Яркая аналогия – даже великие художники делегировали написание своих шедевров ученикам. Не яркая аналогия – лазить по StackOverflow и набирать себе снипеты, из которых потом составить готовый код. Здесь в роли такого умного поиска может выступить ИИ

Как использовать LLM и при этом не деградировать

Структура

Сочинение-рассуждение = вступление, аргументы, контраргументы, заключение

Вступление – знакомство с темой / знакомство со структурой доклада

Рассмотреть проблему с разных сторон

Цитаты – своё отношение

Лингвистика – структура и функция текста

Лингвисты – технический подход

Филологи – интерпретация и культурное наследие

Языкознание

Исследование:

Объект = источники

Предмет = что исследуется

Задачи = исследовательский характер

Угроза для образования

Как и когда человек образовывается или получает образование

Картинки = воображение

Индивидуальность, эксклюзивность, собственный стиль

Как сузить AI -> LLM?

В чём отличие тезисно

Умственное развитие

- мышление клиповое

- память

Как не деградировать, используя LLM в работе

- почему LLM заставляет деградировать

- описание исследования про эссе

- история социального аспекта появления новых технологий:

Телефоны, компьютеры, печатные машинки, шариковые ручки

- деградация ИИ и деградация человека (см стр 3)

Мысль: если человек не деградирует, то и ИИ, который учится у человека, тоже не будет деградировать  
 Вообще, ярко подсветить мысль, что ИИ не самостоятельный – он продукт человеческого труда и учится на другом человеческом труде. Отсюда предположение – почему сверхразум будет уничтожать человека, а не будет ему благодарен? Даже если он будет полностью автономен?

- симбиоз человека и ИИ = эффективность

- ценность поиска чистых данных; поиска достоверной информации

- ИИ – новый инструмент, как его использовать

- как использовать ИИ – стр 6

- деградирует способность мыслить самостоятельно

- когнитивные функции и перераспределение ментальной нагрузки

- выстраивание рабочей деятельности

- угроза для образования. Здесь надо понимать, что разработчик учится постоянно. Или стагнирует

- создаёт ли ИИ/человек что-то новое? Контекст: работа – какой процент творчества?

- как человек думает и когда рождаются идеи

- страх из-за угрозы остаться без работы – что если отпустить страх

Проработать:

- пост Пола Грэмма

- исследование про эссе

План:

1. Актуальность темы:
   1. Потеря работы
   2. Возможности
   3. Деградация/развитие
   4. Переход к актуальности организации своей работы
2. Что такое когнитивные способности?
3. Как когнитивные способности развиваются?
4. Появление новых технологий и реакция на них
5. Проблемы ИИ
6. Выстраивание рабочей деятельности
7. Симбиоз человека и ИИ
8. Заключение? ИИ – это инструмент
   1. Искусство владения инструментом – это будущее
   2. Курсы по ИИ
   3. LLM – максимально удобный для человека инструмент, который интегрируется бесшовно, в отличие от тех же ткацких станков (Луддиты)

Что на слайдах?

Социально-философские проблемы и принципы применения ИИ систем

Социальная и политическая философия

- возможности

- улучшение качества жизни

- горизонты самореализации

Опасности:

- человек

- общество

- окружающая среда

- трансформация базовых составляющих культуры

- философия – рациональный тип мировоззрения

- аргументированное и взвешенное отношение к ИИ

- миропонимание современного индивида, его отношение к окружающему миру

- место и роль человека в мире

- цели современного социума

- чрезмерное делегирование – экзистенциальные основы бытия

- смысл существования

- жизненное предназначение

- как для программистов на производстве объяснить, в чём ценность научных статей/диссертаций

- теоретико-методологические и социально-философские основы планирования развития и внедрения систем ИИ

- требования к разработчикам и пользователям ИИ систем на любых стадиях его развития и совершенствования

-- мб здесь привязать идею гигиены (мыть руки перед едой), провести аналогию с Ветхим Заветом?

- границы применения ИИ

- гос структуры

- научные организации

- бизнес

- гражданское общество

Рефлексия на тему: влияние ИИ на человека и общество

Степень научной разработанности проблемы

Исследования: технические, юридические, социологические, философские

Влияние техники на человека – филососфия техники

- критика технологизации без осмысления

- опасность для человеческой идентичности

- возможность изменения человеческой природы

- социокультурные последствия информационных технологий

М: блок в целом о влиянии технологий на человека + частный блок о влиянии технологий на прикладных разработчиков. При этом лаконично отобразить место целевой аудитории в социуме и в сфере разработки ПО

Нано этика

- Исследования, посвящённые влиянию техники на человека

- Исследования, посвященные определению и классификации ИИ

- Тьюринг «Вычислительные машины и разум» 1950г – впервые высказана идея «мыслящих машин»

- Идея иммитации человека с помощью компьютера – Маккарти, 56 год

- понятие «Сильный ИИ» - философ Серлем

«Такая программа будет не просто моделью разума, она будет буквально разумом»

- «слабый (узкий) ИИ» - Синельникова, Ревинский

Программа, обладающая способностью в соответствии с заложенной архитектурой создавать новую информацию

- представители аналитической философии – сильный и слабый ИИ

- Исследование Легг, Хаттер – «Коллекция определений ИИ»

Исследования, посвящённые проблемам и принципам ИИ

Проблемы применения ИИ – специфика в рамках биоэтики, генной инженерии и информатики

- потенциал изменения человечества в лучшую сторону

- риски, угрожающие основным правам и свободам человека

- какие этические нормы должны быть заложены в ИИ на этапах его разработки

Исследования: применение ИИ в конкретных сферах жизни общества -> конфиденциальность личных данных в наборах больших данных

М: чередовать в докладе блоки с логическими рассуждениями с блоками обзора информации

М: юмор – тонкий и толстый

- Гринбаум «Машина-доносчица» - нравственность узкого ИИ

- ответственность за совершённые действия и принятые решения

- проблема нравственности существует со времён Тьюринга

- возможность причинения вреда человеку со стороны сильного ИИ

- правовая регламентация процессов жизненного цикла ИИ – образовательные риски

- базовые принципы взаимодействия ИИ с человеком

- отсутствие базовых исследований на уровне социально-философского познания

- безопасная стратегия взаимодействия общества с искуственным разумом

**Объект исследования**: системы слабого ИИ для решения практических задач

**Предмет исследования**: проблемы применения систем ИИ, способы и путь их решения в современном обществе

**Цель**: выявить и раскрыть содержание философских принципов взаимодействия ИИ с человеком

- Теоретические подходы к пониманию сущности ИИ

- классификация ИИ в контексте его применения в соц реальности

- наиболее значимые области ИИ в социальной практике

- систематизация документов, регулирующих правовые и этические апекты применения ИИ

- риски для безопасного и эффективного использования

М: безопасность VS эффективность

?? принятие менеджерских решений

- проблема применения ИИ с позиций социальной философии

- в чём негативное влияние ИИ на человека и общество

Новизна:

- доказано, что современное использование ИИ сводится к понятию слабого ИИ

- встраивание ИИ в основы современного социума

- экзистенциальная угроза – утрата человеком привычного для него места в мире

- документы, регулирующие разработку, внедрение и применении ИИ

- необходимость социально-философского исследования ИИ

- проблемы:

- социальной несправедливости

- причинения вреда

- нарушение автономии человека

- нежелательные последствия:

- непрозрачность

- отсутствие ответственности

- нарушение конфиденциальности

Теоретическая база и методология исследования:

- труды отечественных и зарубежных учёных

Понятийный аппарат

- взаимодействие человека и техники

- необходимо определение сущности техники

- техника перестраивает самого человека, изменяет его сущность

- риск потери онтологического статуса

- Целеполагающая функция человека

- Техника = средство достижения целей

М: чтобы знать о реализуемости целей, нужно знать, как они реализуются. Это одна из причин, почему стоит разбираться в теме перед тем, как использовать ИИ

Онтологический статус – вопрос о природе и сущности его бытия, что значит быть человеком, к какой категории сущего он относится

- утрата свободы и ценности личности

- беспрецедентное влияние технологий на социальный порядок и уклад жизни

- возможно ли преобразование человека?

- изменение черт, которые делают человека человеком

- рост безработицы

- профессиональная поляризация

- перестройка структуры управления

Методологическая база

Подходы:

Герменевтический метод исследования – толкование и понимание текстов, их интерпретация

Движение от части к целому и обратно

Теоретическое и прикладное значение

Основное содержание работы

М введение == автореферат?

Глава 1. ИИ: сущность, области применения, особенности исследования

* 1. Определение и классификация ИИ

Термин ИИ – семантически расплывчатый

ИИ = аппаратные комплексы, ПО, наборы данных

* 1. Основные сферы применения и особенности исследования ИИ
  2. Этическое и правовое регулирование ИИ

- выработка стратегии безопасного использования ИИ на уровне академической науки

М: где есть классификация сфер деятельности человека и кто их формирует?

М: правовое регулирование VS социальная реальность

Глава 2. Социально-философские проблемы применения ИИ в современном обществе

- факт причинения вреда – осмысление

- осмысление природы возникновения негативных последствий

2.1. Разнообразие и комплексный, социально-философский характер проблем применения ИИ

- иные технологии – причина проблем, а ИИ – усугубляющий фактор

- человек по отношению к ИИ – оригинал, творец, создатель действительно нового

М: ИИ сносит крышу, если выкрутить креативность на максимум, но ведь именно это и происходит с некоторыми «творцами». Значит ли это, что даже узкий ИИ повторяет поведение «человека» (какого человека?). Можно ли представить работу ИИ как работу каждый раз нового человека? – вот тут новизна и свежесть, можно сделать отдельную статью

- у человека непостижимая с позиций рациональности природа

М: сделать квиз на тему индивидуальности: когда-нибудь шили одежду под себя?

М: делать «философские лирические отступления» о смысле бытия?

2.2. Основные проявления отрицательного влияния ИИ на человека

- отдельные аспекты: необъяснимость и неуправляемость систем ИИ

Таксономия – наука, теория, практика классификации и систематизации объектов в иерархические структуры

- риск отчуждения человека от самого себя

М: типы взаимодействия человека с другими людьми, внутренний диалог, состояние неопределённости и тд – что из этого заменяется ИИ?

К чему это может привести?

- отчуждение человека от собственной сущности и предназначения

- утрата (искажение) статуса единственного существа, способного к интеллектуальной деятельности

- изменение, преобразование окружающей реальности

- некорректное внедрение, применение систем ИИ

2.3. Причины формирования негативных последствий взаимодействия человека и ИИ

- непрозрачность = непрослеживаемость, необъяснимость, неинтерпретируемость

- отсутствие ответственности – сложно для ца IT

- постепенное и непродуманное внедрение = незаметность, бесшовность взаимодействия

Аксиология – общая теория ценностей

Ценности: результат VS процесс

М: угроза – это путь к развитию = переосмысление, переоценка, расстановка приоритетов

Что может и чего не может ИИ

ИИ – инструмент

Пилой нельзя забивать гвозди

ИИ (как эмуляция человека) может всё

Всё ли? Ролик МТС

Вот эту мысль подводить в конец и плавно ей подводить к развязке. Или она и есть развязка? => заключение

QR-код = притча о хлебе (послесловие)

Пламенев = Когда-то давно

- какие результатов мы ожидаем от технологий, а не наооборот, технологии ожидают от нас

- разработка, внедрение ИИ – междисциплинарная задача

- иные формы проектной работы с внедрением ИИ

- развитие социума для выявления и роли человека в мире

- реализация человеком творческого, созидательного потенциала

Глава 3. Философские принципы разработки, внедрения и применения систем ИИ

3.1. Условия эффективного и безопасного применения систем ИИ

- подотчётность систем ИИ человеку

- запрет на нарушение личных границ

Прозрачность – это то, чего не хватает

- полнота информации о системе ИИ

- принцип и методы её работы

- используемые данные

- ограничения

М: здесь принцип работы нейронных сетей не подразумевает прослеживаемости в первом приближении

М: техническая задача – сделать работу нейросети прослеживаемой. В несколько итераций – основные узлы цепочки принятия решений. Возможно, лог действий в первом приближении

Тренд: построение из слабого ИИ инструментов для специфических задач

Модель, в которую будет скармливаться база знаний конкретного разработчика

Возможно: стоит относится к слабой ИИ как к человеку – обучать и накапливать знания по узкоспецифичной области

М: восприятие узкого ИИ как человека поможет нивелировать негативный эффект от его использования?

М: ИИ не всесилен, может ошибаться, у него есть плюсы и минусы

Опасность замены людей системами ИИ: потеря первоисточника знаний и компетенций

М: проблема взаимодействия человека с ИИ – коммуникация с другим человеком (общение).

ИИ генерируют простыню декларативного текста

ИИ не может отвечать вопросом на вопрос – очень синтетическая штука

Правовая ответственность при использовании ИИ?

М: ИИ как инструмент VS ИИ как актор

Ожидания от ИИ, что он может делать всё идеально

3.2. Принцип справедливости, автономии и непричинения вреда человеку в контексте использования ИИ

- независимость человека в принятии решений

- самостоятельность

- системы ИИ не должны усугублять существующие в обществе предрассудки и заблуждения

- решение человеком собственных проблем

- достижение самостоятельно определённых целей развития

- разработка стандартов?

- человек в цифровой действительности

Как использовать ИИ и не деградировать

- ИИ – не мозг. Молоток вбивает гвозди, но дом строишь ты сам

- ИИ даёт ответы, но не думает за человека. Необходимо перепроверять – аналогия со StackOverflow

Мб сделать ИИ Overflow, где ответы будут верифицироваться людьми и улучшаться – это позволит прослеживать

- Попробовать сначала разобрать самому. Любопытство: «скажи мне всё»? В докладе сделать интерактив: а расскажите, какая мысль возникает у вас в голове, когда задаёте вопрос ИИ

- Настройка промпта – настройка себя. Полная аналогия с поиском. Генеративный ИИ == продвинутый поиск

- если перестаешь вспоминать, связывать, анализировать – тупеешь.

Вот тут раскрыть, что включение ИИ в цепочки мышления – это и есть пагубность

- тренировка памяти – естественная, искусственная

М: дать в докладе не раскрытые темы, которые остануться на подумать

М: взаимодействие зрителя с докладом vs сухая информация – подумать над этим

- делать через ИИ, потому что неохота – путь вниз

Автоматизация – это значит освободить время для важного и ценного. А что является важным и ценным?

М: мб ИИ и заставит сейчас задумать о том, что важно, а что не важно?

Самое губительное: отдавать вопрос, проблему ИИ на автомате, без размышления. Как искать что-то без вспоминания?

- ИИ = усилить себя, а не спрятать

- кто пользуется результатом работы ИИ?

- ИИ не заменяет переживания, которые возникают в процессе поиска информации, её интерпретации

- обезьяна с калькукулятором

- человек: чувствуй, ошибайся, удивляйся, восхищайся

- Обучение основам ИИ и основам взаимодействия с ИИ

Вечные проблемы ИИ

Генеративный ИИ на основе нейронных сетей

2023 год. Письмо: Илон Маск, Стив Возняк – приостановить разработку и обучение нейросетей

- взвешенное отношение к ИИ

- ИИ и проблема первоисточников информации

Гигиена работы с информацией

- ИИ = технология автоматизации человеческого труда

- имитация некоторых интеллектуальных способностей человека

- принципиальная неспособность к осмысленному творчеству

Гипостазирование – логическая ошибка мышления, при которой абстрактные идеи, понятия или свойства рассматриваются как реально существующие, самостоятельные объекты или вещи

Технология = гипостазирование человеческих представлений и ценностей

- люди воплощают свои представления и ценности в технических решениях

- концепции и технологии IT = глубокие философские корни

Эксплицитно – явно, открыто, в противоположность скрытому или подразумеваемому

ООП = Платон, ФП = Декарт, двоичная система счисления = Лейбниц

- философская основа ИИ проблематична по своей природе

- философские принципы были заложены в античный период

- философские концепции оказали влияние на формирование кибернетики, информатики, ИИ

- древнегреческая философия – пифагорейцы, Платон, Аристотель

- рационализм – Декарт, Лейбниц, Гегель

- детерминизм – Демокрит, Спиноза, Лаплас

Рационализм: мир познаваем через математику; идея превосходства рационального познания перед другими его видами

Детерминизм: всё подчинено законам, включая мышление

Формализм: реальность можно описать алгоритмически

Пресуппозиция (презумпция) – скрытое, подразумеваемое утверждение, которое предполагается истинным, чтобы основное утверждение имело смысл

Математика, естественные науки, техника – выросли из философии

Христианство: аспект человеческого бытия непостижим рационально

- современные интеллектуальные системы ограничены в своих возможностях, потому что базируются на античных философских концепциях

- персоналистические, экзистенциальные концепты игнорируются

Кто такой человек?

- развитие ИИ оценивают на основе сопоставления его возможностей с возможностями человека

- не существует единой, общепризнанной антропологической концепции

- достижение человеческого уровня ИИ всегда может быть оспорено

- тест Тьюринга критикуется. По своей сути – это тест

- человеческая рациональность = гипертрофированная до нечеловеческого уровня человеческая рациональность

- почему машина должна быть человекоподобной?

Интеллект vs мышление

- интеллект – способность к решению задач

- что насчёт способности к постановке задач?

- машины не могут задавать вопросы

- машины не могут самостоятельно ставить себе проблемы

- проблема = несоответствие между желаемым и действительным

- у машин нет желаний => нет проблем => нет вопросов

- исследовательская деятельность на академическом vs на бытовом уровнях

- мышление – более широкое понятие по сравнению с интеллектом

- мышление = интеллект + постановка задач

- Artificial Intelligence -> Artificial Thinking -> Artificial Philosophy

Может ли машина желать?

Св-ва человека: быть, знать, хотеть

- желание – то, что определяет человека совокупно с мыслью

- животные не способны мыслить, а машины желать

Волюнтаризм – решающая роль в воле человека, решения принимаются субъективно, игнорируя объективные обстоятельства и законы

- ? желание первично по отношению к мысли

- цели определяются ценностями

Свобода воли

- интеллектуальная система не обладает иным руководящим началом, кроме алгоритма, который внёс в неё разработчик

=> она не способна кардинально менять рамки своего поведения

- мета-алгоритм, по которому формируются алгоритмы

- поведение ИИ детерминированно => предсказуемо

Методологические проблемы современной кибернетики

- разработчики ИИ не учитывают ряд особенностей, присущих мыслительной способности человека

- кибернетика как наука имеет методологические проблемы

- технико-математическая направленность

Рецепция (гуманитарное) – процесс заимствования и адаптации чужих норм, идей, культурных форм и их интеграция в собственную систему

- теория организационных систем Новикова

- нечёткая логика Заде

- темпоральные логики – логика, в высказываниях которой учитывается временной аспект

- диалог между естественно-научным и социогуманитарным

- теория принятия решений, теория игр, ИИ – специалисты в сфере кибернетики ставят и решают задачи в различных предметных областях по своему усмотрению

Проблемы ИИ: этика данных, экология, стереотипы и ошибки в ИИ

- нейросети вводят нас в заблуждение – детерминировать источник информации. Но вводить в заблуждение могут и некомпетентные люди в жизни и в комментариях в интернете

Александр Немальцев – руководитель группы машинного обучения в VK

Предвзятость алгоритмов

- предубеждения и стереотипы у нейросетей

- работа ИИ – это результат данных, которые ему скормили

- основная задача LLM – максимизировать точность предсказаний

- публичный датасет – путь к контролю и прозрачности LLM?

Eleuther AI = датасет The Pile

Ошибки и дезинформация

- нейросеть не думает, что правда, а что ложь. Она ищет статистически наиболее вероятный ответ = умный, модифицированный поиск

Вопрос: «сколько камней в день стоит есть?»

- мониторинг и исправление фейков – не быстрый процесс

Решается надстройками в виде фильтров

ИИ и экология

- обучение LLM – энергии столько же, сколько потребляет город; огромное количество воды для охлаждения

- вопрос экономической эффективности LLM

- Проблема анонимности – привести факты утечек данных

- регламент GDPR, который требует согласия пользователей на любую обработку данных

- гигиена в вопросе персональных данных

ИИ в политике

- политические наклонности чат-ботов

ИИ и искусство

- можно ли считать ИИ контент искусством

- можно ли обучать ИИ на изображениях?

Авторское право

- нейросети могут автоматизировать только 26% задач профессиональных художников и дизайнеров

А сколько у программистов? Составить модель задач, которые выполняют программисты

- ИИ контент очень однотипный и неточный

- «креатива» у ИИ нет – он просто комбинирует уже существующие стили и образы БД

Юридическая ответственность

- правовой вакуум, который не учитывает последствия экспериментов с ИИ

- фатальные случаи 2018г Tesla

- ИИ – нейтральный инструмент, который можно использовать по-разному

- пока мы не научились регулировать ИИ на уровне закона

Проблемы и вызовы ИИ

- предвзятость, основанная на данных

- модели = чёрные ящики

Необходима объяснимость и понятность решений, их прослеживаемость

- рынок труда – временное исчезновение профессий (их переход в роскошь)

- Обучение ИИ – огромное кол-во данных и ресурсов

- проблема универсальности – плохо работают в условиях, к которым они не подготовлены

- проблема надёжности и устойчивости

Концептуальная проблема

- безопасность – манипулирование поведением

- вредоносное использование – автоматизация кибератак

- создание прозрачных и объяснимых алгоритмов

- уменьшение объёма данных и ресурсов

- прямо таки не хватает большого объёмного гайда о том, как вообще работать с этим инструментом

Сейчас это выглядит так: пишет в чат что угодно. ИИ отвечает что угодно

Если это инструмент, к нему нужна инструкция

- глава OpenAI – Сэм Альтман

- заявление об опасностях ИИ

- создание ИИ систем, которые человечество не сможет контролировать

- как разработать системы, которые не нанесут вреда

- отчёт Всемирного Экономического Форума 2024: ИИ вводит в заблуждение

- Сатья Наделла – стандарты

В 2024 году ИИ пользуются 314 млн человек

Ноябрь 2023 года – 28 государств подписали декларацию Биетчли

- вопрос восприятия человеком контента, сгенерированного ИИ

Сдерживание развития ИИ:

- кадры

- оборудование и технологии

- данные

- Развитие ИИ – технический и социальный тренд

- вписывается в череду индустриальных инноваций прошлого

- прежние технологии были направлены на улучшение и ускорение физического труда

- ИИ впервые в истории позволяет оптимизировать интеллектуальный труд

=> наибольшие изменения произойдут в информационно ёмких индустриях

- значительные сокращения из-за ИИ

- прежние технологические революции – рынок труда трансформируется

Люди получают больше благ, прикладывая меньшее количество усилий – экономика становится более эффективной

- аналитика Bloomberg – мировой ВВП вырастет к 2030 году на 15.7 трлн благодаря ИИ

- Билл Гейтс – люди будут работать всего 3 дня в неделю

Нехватка компьютерных чипов и электроэнергии

- кризис на рынке полупроводников

- чипы

- электроэнергия

2023 год – принятие закона об ИИ Европарламентом

Безопасность в значении safety, а не security

? В РФ недостаточно регулирования ИИ

- нужно смириться с тем фактом, что машины будут делать часть работы лучше и эффективнее людей

- запрет правообладателей на использование их материалов в обучении

- прецедент New York Times

- генераторы научного и общественного знания

Ограничения – ИИ = работа на открытых данных, ранжирование их по степени значимости и авторитетности

- искусственно затормозить эволюцию ИИ невозможно

- концепция «технологической сингулярности» - ИИ может начать совершенствовать сам себя

- улучшение производительности креативного класса

- улучшение таргетированных предложений

Проблемы и перспективы современного ИИ

- программы, которые выполняют строго определённые задачи, если человек их этому обучит

- машина не понимает, что она делает (в противовес диалогу со специалистом). Уточнение условий, но не сути проблемы

- определение собак – программа не понимает, что такое собака

- качество и объём данных

- ИИ не способен создавать, способен подражать по своей природе

- возможности – обработка

- ИИ не может функционировать без человека

- ?? ИИ избавит от однообразных задач

Пять проблем, которые пока не может решить ИИ

- человечество стремится к атвоматизации всех процессов

- механическая -> информационная революция

- на данный момент ИИ – это возможности (потеря компетенций)

- нейросети работают на базе накопленного опыта – где его брать, если отказаться от источника этого опыта?

Проблемы ИИ: о чём следует знать и помнить

Модели для генерации: текста, иллюстраций, звука, видео, документов, медиаданных

- проблема поиска данных – ограничения

- синтетические данные

Почему развитие ИИ может остановиться: ключевые проблемы и их решения

- накопление ошибок

Проблема контроля ИИ

- призыв начать исследовать управление ИИ до разработки суперинтеллекта

- внезапное зарождение суперинтеллекта

? система должна рассуждать о намерениях людей

? последствия действий существующего ИИ

- проблема коммуникации с техникой

Цель – рассказать о том, какие действия выполняет разработчик

Как работают когнитивные способности?

В чём причина деградации – какой механизм?

Как строить свою работу, чтобы этого избежать?

Какие навыки страдают при использовании ИИ?

А при делегировании другим людям?

Почему вообще навыки или когнитивные способности могут деградировать?

Какие умственные способности нужны для программиста?

Разработка ПО != написание кода

Какие умственные операции совершает разработчик?

- доработка кода

- поиск и исправление багов

- умение работать с документацией (гайды, обсуждения StackOverflow, чтение и размышление над статьями на разные темы)

- умение обсуждать решения

- умение продумывать структуру системы/кода

- системное мышление – как работает система, состоящая из блоков

- написание документации

- разработка в более широком смысле

- умение разбираться в системах, фреймворках

- умение проектировать системы

Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for Essay Writing task

- 3 groups: LLM, Search Engine, Brain-only

- EEG – cognitive load

- analyzed essays using NLP

- Brain – Search – LLM = cognitive activity scaled down

- LLM offers immediate convenience, our findings highlight potential cognitive costs

- Result: - concerns about the log-term educational implications

- need for deeper inquiry into AI’s role in learning

Влияние ИИ на когнитивную сферу человека в контексте ценностно-мировоззренческого анализа

- воздействие ИИ на когнитивные возможности через призму развития у него ценностно-мировоззренческих качеств личности

- новизна – в неоднозначности влияния

- субъект с развитым мировоззрением применяет программы ИИ как средство для совершенствования своих когнитивных качеств

- недостаточно развитый субъект может не осознавать негативные когнитивные последствия взаимодействия с ИИ

- нет понимания полезностей тренировки памяти, функций абстрактного мышления

- осмысление проблем духовной жизни

- методическая концепция гуманитарного воспитания и образования

- ИИ – комплекс технологических решений, дающих возможность имитировать когнитивные функции человека

- узкий ИИ не обладает универсальными познавательными характеристиками людей

- сильного ИИ пока не существует

- взаимосвязь людей и интеллектуальных программ приобрела в современном социуме характер постоянно усиливающейся тенденции

- свыше 30% россиян регулярно применяют ИИ в работе

- выявление путей и форм влияния ИИ на когнитивные характеристики человека в контексте ценностно-мировоззренческого анализа, феноменов его духовной жизни

- недостаточная изученность темы

Задачи:

- изучить влияние ИИ программ на когнитивную сферу человека

- определить условия реализации воздействия

- охарактеризовать пути влияния на память, творческие возможности и познавательный потенциал людей в контексте развития их мировоззрения

- база = материалы и положения, высказанные авторами в их работах

М: почему научные статьи, а не хабр?

- взаимодействие людей с ИИ – эффективное применение последнего

- положительное и отрицательное влияние ИИ на психику и сознание человека, когнитивные характеристики

- новая психическая реальность цифрового сознания

- сравнительный анализ интеллекта человека и ИИ

- научных исследований на духовный мир человека пока мало

- диалектический метод познания идеальных явлений духовного мира людей

- духовный мир изучается на базе методологических принципов развития и связи системы познавательных структур субъекта и явлений объективированного уровня духовного бытия общества (ИИ как часть)

- объективированный – сплав познаваемого им объекта и субъективного замысла творца

- метод от абстрактного к конкретному

- метод формализации

- метод идеальных моделей

- взаимосвязь познавательных структур человека, мировоззрения и феноменов объективированного уровня духовной сферы

- системный метод исследования – ценностно-мировоззренческого анализа

- мировоззрение (идеальные ценности) = ядро сознания человека – определяющий элемент системы феноменов его духовного мира

- ценностные компоненты мировоззрения влияют на характер когнитивной деятельности субъекта

- сознание и мировоззрение человека постоянно взаимодействует с явлениями объективированного уровня духовной сферы общества

- отсюда связь всех явлений

Пример:

- гипертекст – феномен объективированного уровня духовной сферы

- многозадачность – мировоззренческое качество современных людей

- клиповое мышление – познавательная особенность

- похожесть тенденций влияния интернета и ИИ на духовный мир человека

- знания о факторах, условиях, специфике воздействия ИИ – образовательная концепция цифрового общества

Обсуждение и результаты

- утвердилось представление о важной роли ИИ, который заменяет простой умственный труд человека

- ИИ: логическое мышление, анализ огромных объёмов информации, скорость и точность «мыслительных операций»

- другие характеристики духовного мира человка: чувства, эмоции, привязанности, моральные принципы

- человек часто наделяет ИИ человеческими качествами

- сведения воспринимаются не как информация, подобранная по некоторой проблематике, а как ценные знания, проверенные компетентным специалистом

- рост доверия к ИИ и перекладывание на него ответственности за когнитивные и практические решения

- аварии с Tesla: человек воспринимает автопилот как полноценного участника дорожного движения, но юридическая ответственность возложена на водителя

- сомнение человека в своём разуме

- разочарование ы «собственном уме … когда люди передоверяют свои умственные функции машине, то это говорит о том, что они освободились от всякой претензии на знание»

- система духовной жизни общества – взаимодействие объективированного и субъективированного уровней духовной сферы

- субъективизм – точка зрения, которая учитывает только одного человека

- в цифровом мире часть интеллектуальных функций на субъективированном уровне передаётся ИИ

- ИИ выводит часть объективированного идеального из сферы естественного интеллекта

- усиливается вытеснение интеллектуальной деятельности людей функционированием ИИ

- ИИ – достижение естественного человеческого разума, у которого отсутствуют когнитивные способности

- содержание объективного духовного должно осмысливаться людьми, тк этот уровень имеет значение только в единстве с тем, кто понимает объективный уровень

- из-за того, что ИИ выводит часть объективного идеального, из-за этого не субъективируется часть содержания объективного уровня

- в процессе работы ИИ не происходит полное превращение объективированного уровня духовной сферы в идеальное бытие субъекта

- ИИ слабого типа не субъект, а объект

- дисбаланс между объективированным и субъективированным уровнями духовной сферы

- растёт влияние объективированного уровня духовной сферы

- «Умные программы» - объективированный дух человека

- человек чувствует себя беспомощнее в системе духовной сферы

- «окаменевший дух» - культура, которая без воспринимающего её сознания человека теряет смысл – Гегель

- предстоит определить опасности, возникающие из-за этого

- в цифровом мире всё очевиднее становится разница между информацией и знанием

- Информация = сведения, представленные в той или иной форме

- Информация объективирована и представлена на материальных носителях

- Знания – субъективированная форма информации

- носит личностный характер, это результат понимания и осмысления сведений человеком в ходе его практической и когнитивной деятельности

- Информация = текст, символы, изображения

Знания = индивидуальные понятийные структуры, в сознании человека. Эти индивидуальные понятийные структуры отражают его понимание явлений окружающего мира и их взаимоотношений

- ИИ облегчил познавательную деятельность человека <= упростил поиск и доступность информации

На первый взгляд благодаря этому умственный труд становится эффективнее

- во-первых – это только информация с цифровых ресурсов

Во-вторых – это именно информация, а не знания

В-третьих – ИИ и поиск отбрасывают часть информации, которая может быть полезна

- пассивное использование сведений является «функционально ограниченной» познавательной активностью человека

- технологии, усиливающие интеллектуальные способности человека, превратились в испытание сложившихся способов формирования интеллекта и субъективности как особенности человеческой деятельности

- воздействие ИИ на личность зависит от ценностно-мировоззренческих характеристик личности

- ИИ не относится к информации как к ценности в человеческом смысле

- ИИ не имеет «субъективного образа, отражающего действительность»

- цель познавательного процесса для ИИ – информация, а для человека – знание

- знание субъективно принадлежит человеку, а не пребывает в объективированном виде

- знание – это не просто принятие информации в качестве данности

- знание – это осмысление, активный процесс отражения и субъективной фиксации свойств предмета в человеческом сознании

- информация – это сообщение человеку, либо знание в сознании каких-то других людей

- привычка пользоваться готовой информацией, а не превращать её в долговременные знания

М: информация от ИИ не является по сути знанием, ибо не «пропущено» через сознание человека

- исчезает этап предварительного отбора информации, её обобщения, абстрагирования – всё это переходит к ИИ

- продолжается и углубляется процесс, который начался с развитием интернета

- тенденции многозадачности и клипового мышления продолжают совершенствоваться и углубляться

- формируются иные познавательные ценности – умение быстро найти информацию

- не ценится как раньше память

- чем больший объём знаний задействует человек в процессе мышления, тем более широкими станут его обобщения, продуктивнее и глубже характер познавательного процесса

- ИИ не способствует развитию способностей самому анализировать, широко обобщать информацию

- ИИ полезен при рассмотрении конкретных проблем

- в цифровом мире изменилось отношение к информации и знаниям

- в 17 веке образованный человек мог освоить большую часть знаний, которые были для него субъективно ценными

- в цифровом мире информация является приоритетной ценностью, однако важность отдельной единицы информации постоянно снижается

- даже если человек сделает приоритетом когнитивную активность, то узнает лишь мизерную долю информации из духовной сферы

- То формируется убеждение о ненужности постоянного серьёзного осмысления инфы

- формируется убеждение, что нормально поверхностно смотреть на вещи, тк основная задача – охват большого кол-ва информации

- реальность способствует простой вере в сведения, что не помогает научному мышлению

М: здесь вызов и для научной сферы – стартапы, где продукт создаётся путём проб и ошибок. Пример: разработка ракет на заре космонавтики

- не складывается понимание важности изучения сущности объектов мира , не формируется стремление к глубокому осмыслению вопросов

- информацию можно брать из интернета без извлечения главных идей

М: часто в работе разработчика проще опытным путём «закатать» решение, не тратя уйму времени на погружение в суть вопроса

- теряется понимание различия между информацией и знаниями, собственные познавательные усилия перестают цениться

М: мб здесь важно не только самостоятельное построение разработчиком деятельности, но и построение работы над проектом таким образом, чтобы делить задачи, фазы