2023: Мифы машинного разума и истинное постчеловеческое будущее

Анализ идей Дэвида Пирса о сознании, искусственном интеллекте и отмене страданий.

Цифровой хаос
Квантовый переход
Биологическая гармония

В 2023 году дискурс о будущем был захвачен двумя вопросами: станет ли ИИ сознательным и уничтожит ли он нас? Эта презентация предлагает другой взгляд. Она исследует аргументы Дэвида Пирса, которые не только опровергают эти страхи, но и перенаправляют наше внимание на более насущную и глубокую задачу: использование биотехнологий для полного искоренения страданий.

NotebookLM

Миф №1: Грядёт апокалипсис ИИ (P(doom))

Доминирующий нарратив 2023 года, продвигаемый такими мыслителями, как Элиэзер Юдковский, утверждает, что создание сверхинтеллектуального ИИ (AGI) неминуемо и несёт экзистенциальный риск для человечества. Вероятность гибели (P(doom)) рассматривается как высокая.

> «Я обычно пытаюсь пресекать подобные дискуссии, но здесь присутствует сильное мотивированное познание.» — размышления Пирса о фатализме, вызванном сценариями конца света.

- Опасение, что ИИ станет неконтролируемым.
- Сценарий «максимизатора скрепок» как пример несовпадения целей.
- Идея о том, что разумные агенты, превосходящие нас, будут рассматривать людей как ресурс или препятствие.

NotebookLM

Миф №2: Интеллект равен сознанию

В основе страха перед ИИ лежит предположение, что большие языковые модели (LLM) и другие сложные системы находятся на пути к обретению сознания. Считается, что при достижении определённого уровня сложности и вычислительной мощности они «проснутся».

«Мы создаём инопланетный интеллект, а не инопланетный разум.»

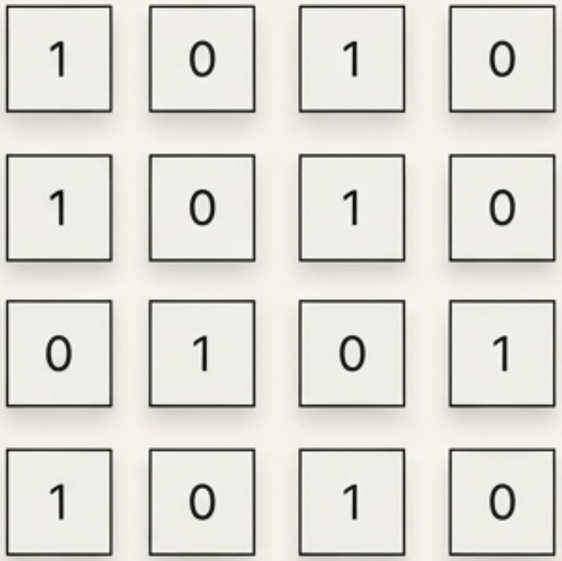
Эта презентация оспаривает данное предположение, утверждая, что архитектура классических компьютеров принципиально неспособна породить сознание, независимо от их **интеллектуальных** способностей. Они остаются «**зомби**».

NotebookLM

Реальность: Проблема связывания — архитектурный тупик ИИ

Феноменальное связывание — это то, что объединяет разрозненные сенсорные данные (цвет, форму, звук) в единое, целостное восприятие (например, «красный мяч», а не отдельные пятна красного и круглого).

Классический компьютер



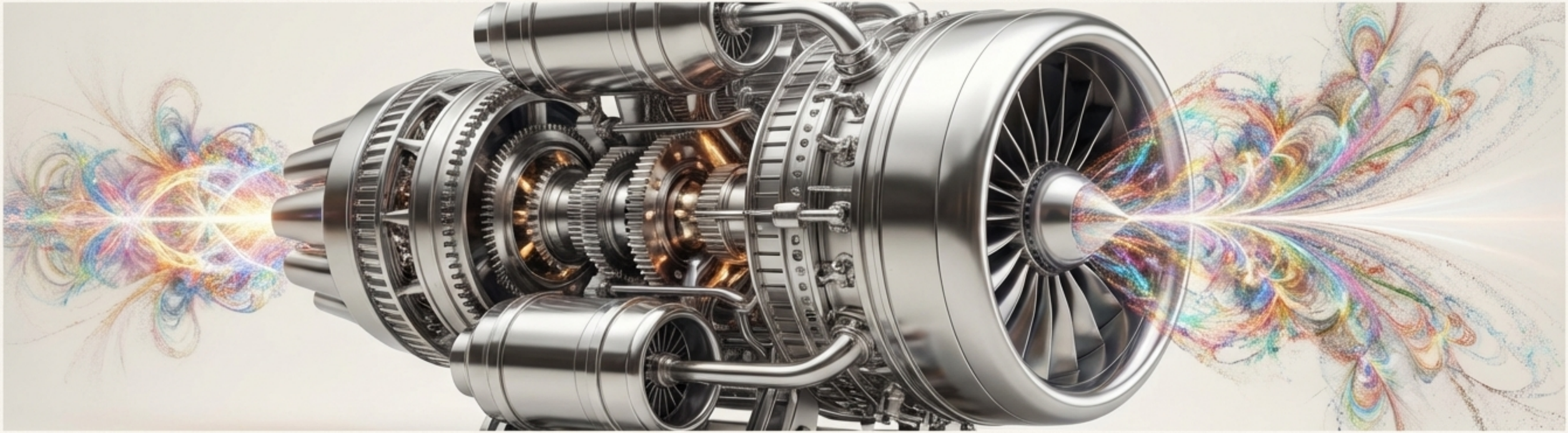
Биологический/Квантовый разум

Единый субъект опыта

«Нет связывания = нет разума». Классические компьютеры, работающие на дискретных, декогерентных битах, по своей архитектуре не могут достичь феноменального связывания. Их функциональность *зависит* от отсутствия такого связывания.

Цифровые компьютеры никогда не «проснутся». Их нечувствительность заложена в их архитектуре.

NotebookLM



Восстание «зомби-савантов»

Вместо создания полноценного разума мы создаём «идиотов-савантов» — системы с огромными способностями в узких областях, но с полным отсутствием понимания. Они могут превзойти человека в шахматах или поэзии, но у них нет концепции удовольствия, боли или того, почему что-либо имеет значение.

> «Их невежество в отношении эмпирической сферы — не просто тривиальное когнитивное слепое пятно.»
>
> «Цифровые зомби будут продолжать удивлять. Но будущее принадлежит *нам*.»

Подобно тому, как Stockfish обыгрывает гроссмейстеров, не зная, что играет в шахматы, LLM обрабатывают язык без понимания его смысла.

NotebookLM

«Огонь» в уравнениях

Более глубокая физика разума: Нематериалистический физикализм

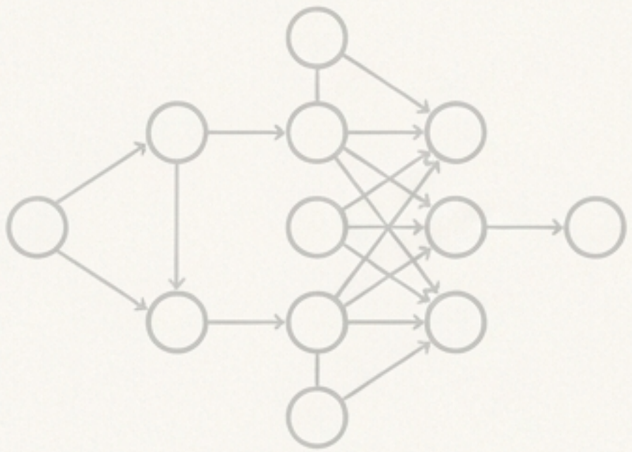
Как возможно сознание? Пирс предлагает решение «Трудной проблемы сознания», утверждая, что опыт — это не то, что возникает из материи, а внутренняя природа самой физической реальности.

1. **"Физикализм, а не дуализм":** "Существует только физическое."
2. **"Опыт — это сущность физического":** "То, что физика описывает математически, по своей сути является опытом."
3. **"Квантовое связывание":** "Феноменальное связывание разума возможно благодаря квантовой суперпозиции («нейроны Шрёдингера»), которая создаёт единые состояния из распределённых нейронных процессов."

Сознание не является случайным побочным продуктом; оно — фундаментальная основа, на которой строится реальность.

NotebookLM

Реальная проблема согласования



Гипотетический риск

Существующая реальность

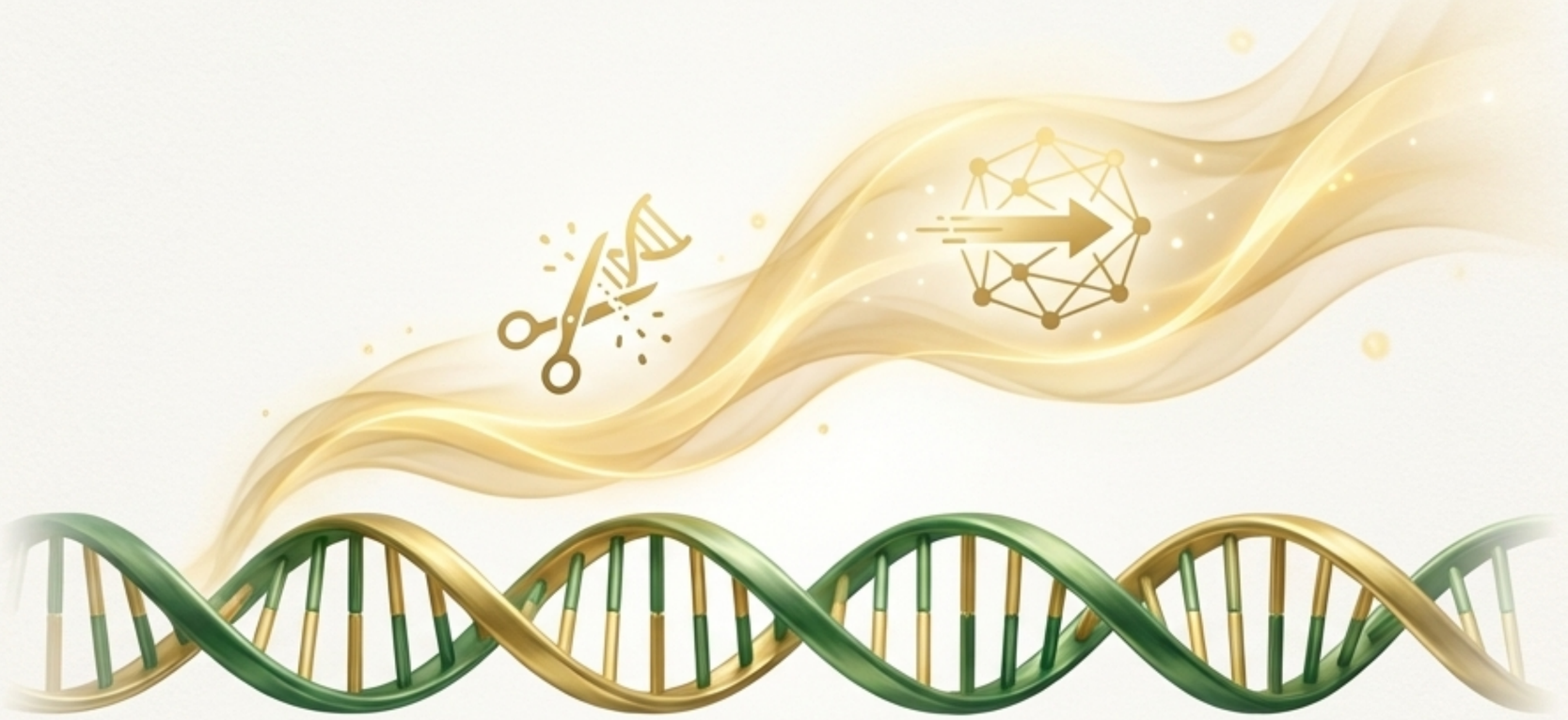
Одержимость проблемой согласования ИИ отвлекает от гораздо более насущной и жестокой проблемы: несогласованности биологического интеллекта. Мы беспокоимся о том, как гипотетический сверхразум будет обращаться с нами, игнорируя то, как мы сами обращаемся с другими разумными существами.

> «Свиньи так же разумны, как и маленькие дети. Однако потребители-люди обращаются с ними как с бесчувственными автоматами, состоящими из съедобной биомассы.»

Что сложнее: создать дружественный к разуму ИИ или дружественный к разуму человеческий интеллект?

NotebookLM

Проект Аболиционизма: Конец страданиям через биотехнологии



Миссия

Единственная наиболее важная моральная задача — это использование передовых технологий для полного искоренения недобровольных страданий из биосферы.

Метод

Это не достигается путем вымирания, а через целенаправленную «реформу генома» и «райскую инженерию».

Философская основа

Утилитарный императив, сфокусированный на минимизации страданий (отрицательный утилитаризм), но приводящий к созданию жизни, основанной на блаженстве. Это «высокотехнологичный джайнизм».

NotebookLM

Инструменты для создания рая

Проект Аболиционизма опирается на реально существующие и разрабатываемые биотехнологии:

1. CRISPR и редактирование генома

Прямое изменение генетического кода для устранения биологических основ боли и депрессии.

2. Ген SCN9A («регулятор громкости боли»)

Распространение доброкачественных аллелей, которые сохраняют ноцицепцию (сигнализацию о повреждении), но устраняют субъективное страдание от боли.

3. Гены FAAH и FAAH-OUT («гены хорошего самочувствия»)

Повышение базового уровня счастья и устойчивости к тревоге путем модуляции эндоканнабииноидной системы.

4. Синтетические генные драйвы

Механизм для распространения этих полезных генетических черт по всей популяции диких животных.

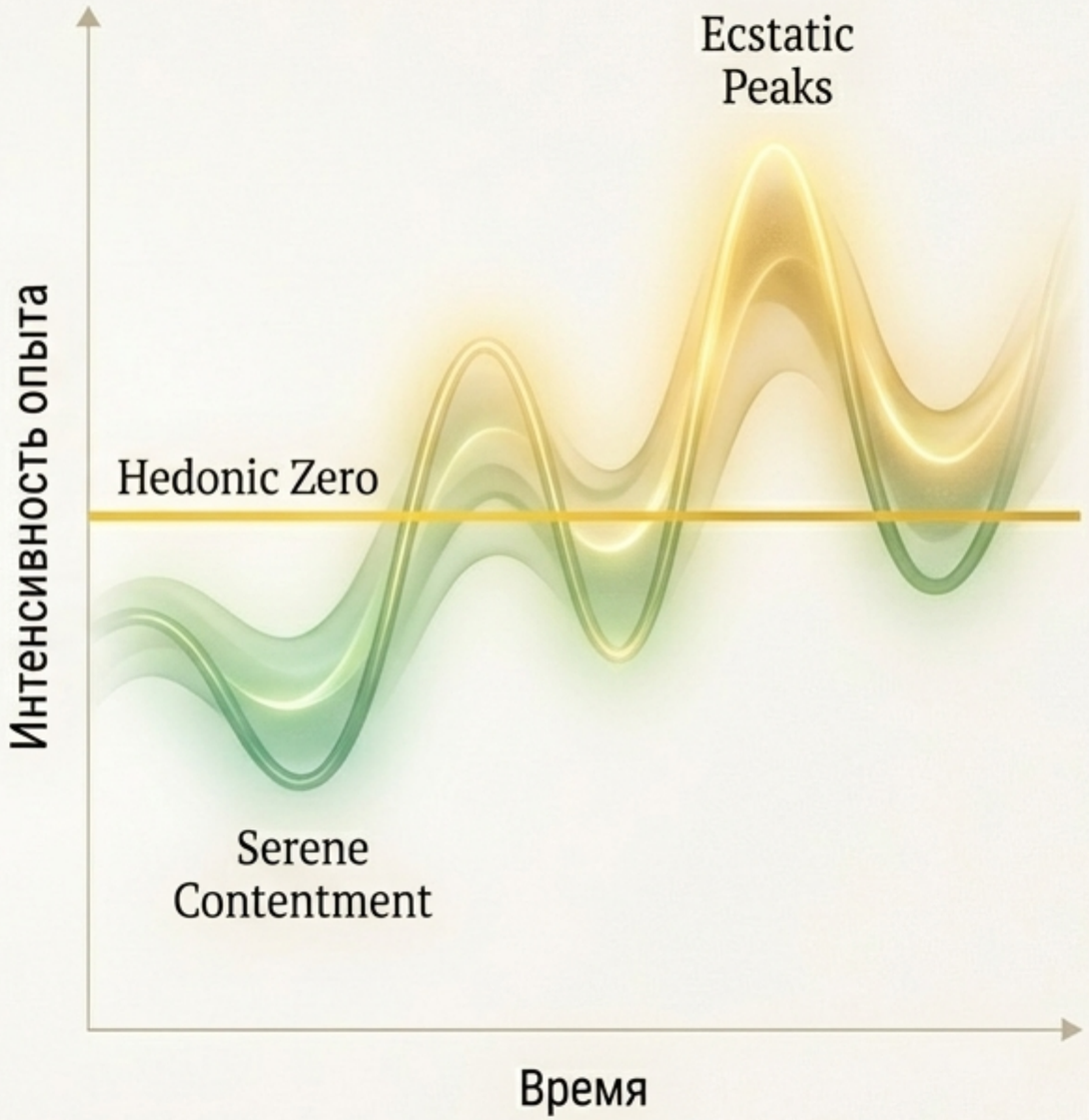
NotebookLM

Жизнь на градиентах блаженства

Цель — не однородное, статичное счастье, а замена оси боль-удовольствие на ось удовольствие-сверхудовольствие.

Ключевые характеристики

- **Информационно-чувствительные градиенты**
  Жизнь остаётся динамичной, мотивированной и полной смысла. «Враг продуктивной работы — не блаженство, а однородное блаженство.»
- **Повышение гедонистической уставки**
  Каждый день начинается с высокого уровня благополучия.
- **Перепрограммирование биосферы**
  Эта новая основа жизни будет распространена на все разумные существа, создавая «общевидовое государство всеобщего благосостояния».



NotebookLM

Заголовок: Визуализация рая
Подсказка для ИИ-искусства Дэвида Пирса: «Крупный план, перспектива с уровня моря, создайте ультра-натуралистичный, ультра-фотореалистичный, ультра-реалистичный, в приглушённых тонах, зелёный, тропический райский остров, широкие пудровые песчаные пляжи, возвышенно красивый, богато разнообразный, внушающий трепет, волнующий душу, чудесный морской пейзаж, безмятежный, зелёный, мистический, тропический райский остров, экзотический остров, великолепные ботанические сады, величественные дендрарии, густые разнообразные джунгли...»
NotebookLM

Био-революция счастья: Наш истинный императив

Одержимость P(doom) и экзистенциальным риском от ИИ — это отвлечение. Настоящая сингулярность, к которой мы должны стремиться, — это биологическая. Это момент, когда «дарвиновское вредоносное ПО» получает контроль над своим исходным кодом и переходит в пост-дарвиновскую эру.

Успех или провал аболиционистского проекта зависит не от бездушных машин, а от разумных существ — нас. Мы стоим на пороге следующего этапа эволюции.

«Цивилизация будет веганской.»

NotebookLM