

гравитация

2025

Международная университетская премия
в области искусственного интеллекта
и больших данных

Панельная дискуссия «Философские и этические аспекты технологий искусственного интеллекта»

***«Как образы будущего человеко-машинной цивилизации
формируют сегодняшнюю повестку и постановки задач в области ИИ»***

ИЛИ
МГУ

Воронцов Константин Вячеславович

д.ф.-м.н., профессор РАН, руководитель лаборатории
машинного обучения и семантического анализа

Институт искусственного интеллекта МГУ им. М.В. Ломоносова

Какие идеи мы закладываем в определение ИИ

Искусственный интеллект — интеллект, демонстрируемый машинами, в частности компьютерными системами...

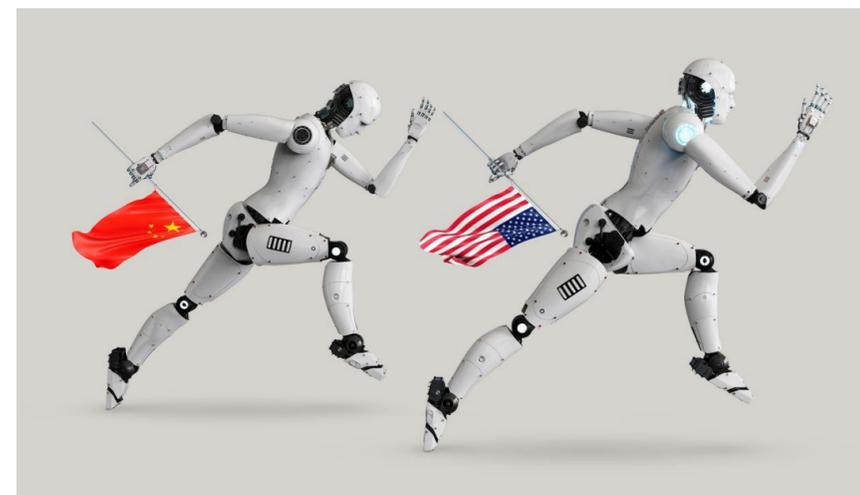
Википедия

Искусственный интеллект — комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их.

[Национальная стратегия развития искусственного интеллекта на период до 2030 года \(с изменениями на 15 февраля 2024 года\)](#)

Какие идеи Китай закладывает в определение ИИ

«Китай не будет слепо следовать тенденциям или участвовать в безудержной международной конкуренции»



AGI — инструмент, «**способствующий пониманию человеком мира и его преобразованию**», «ИИ общего назначения», который имеет множество применений, но не является сверхчеловеческой или самосовершенствующейся системой.

Антропоцентричное определение ИИ — необходимая основа этики ИИ

Искусственный интеллект —
вычислительные технологии,
создаваемые для повышения
производительности
созидательного
интеллектуального
труда людей

не замена человека

не загадочный новый тип разума

не повод уподобиться Богу, творящему «по образу и подобию Своему»



Какие технологии развивать?

Идём от целей и задач к технологиям, но не наоборот

- 1) **каковы цели и задачи цивилизационного развития?**
- 2) какие технологии необходимы и минимально достаточны?
- 3) AGI или антропоморфные роботы насколько важны для развития?
- 4) генераторы текстов и картинок насколько важны для развития?
- 5) ожидаемые эффекты насколько перевешивают затраты и риски?



Любая технология — это не цель, а средство

Больше вопросов, чем ответов...

Почему мы хотим AGI? Вместо того, чтобы заниматься конкретными задачами автоматизации интеллектуального труда людей...

Почему мы хотим антропоморфных роботов? Вместо того, чтобы оптимизировать машины под задачи, решаемые ими для нас...

Почему мы хотим приблизить машинный интеллект к человеческому? Вместо того, чтобы помогать людям решать их задачи лучше и быстрее...

Почему мы читаем Р.Курцвейла и Ю.Н.Харари и верим им?

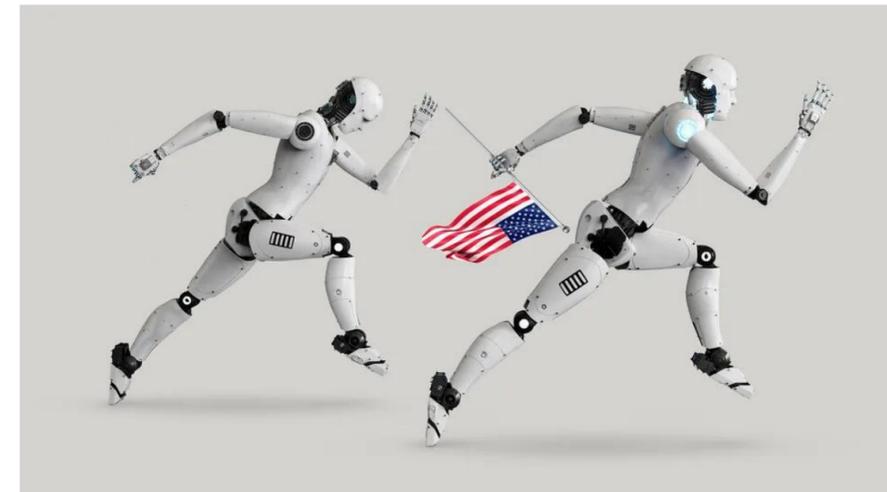
Вместо того, чтобы разрабатывать нужные нам технологические сценарии...

Кто нам навязывает противоестественные цели — сам ход развития науки или ставший привычным западный **культурный код**?

Что значит «естественный» ход развития науки, технологий, цивилизации?

Как перестать догонять и пойти на опережение

Только читать статьи и делать улучшения — означает всегда плестись в хвосте, отставая в исследованиях на 1-2 года.



Чтобы не отставать, нужно иметь аналоги.

Чтобы обгонять, нужно задавать новые *свои* цели, смыслы, концепции, ставить и решать *свои* задачи, идущие от собственных потребностей

Что для этого нужно:

целеполагание, системность, идеология, фантазия, образы будущего.

Наши задачи

В решении которых ИИ может помогать



увеличить число умных людей

быстро

намного

надолго



создать инструменты коллективного разума, научиться мыслить

системно, структурированно

мультидисциплинарно

коллективно, ради общих целей



создавать позитивные образы коллективного будущего



воссоединить воспитание с естественно-научным образованием



включить позитивную пропаганду цивилизационных ценностей

Цивилизационная идеология

Иерархия цивилизационных ценностей:

- 1) биосфера Земли, уникальная в достижимой части Вселенной
- 2) человеческий вид, результат миллиардов лет эволюции
- 3) коллективный разум — образование, наука, культура
- 4) индивидуальный разум, жизнь человека
- 5) результаты труда человека

Цель человеческой цивилизации — неограниченно долгое сохранение биосферы Земли, защита её от катаклизмов ради выживания вида homo sapiens в условиях, достаточно комфортных для всех людей

Добро — всё, что этому способствует (любовь, познание, созидание, ...)

Зло — всё, что этому препятствует (ненависть, невежество, деградация, ...)



Цивилизационная
идеология
ДЗЕН-канал

<https://dzen.ru/civideology>



Формирование позитивных образов будущего

Идея: открытая литературная вселенная фантастики ближнего прицела

- **фантастическая гипотеза** — научная, реалистичная, достижимая
- **цель** — обнаружить угрозы и уязвимости развития ИИ на пути к человеко-машинной цивилизации, найти способы их обхода
- **задача** — проверка гипотез о возможности или невозможности, убедительности или неубедительности различных путей развития
- **задача** — попытка «взлома» цивилизационной идеологии, вывода апокалиптического сценария из исходных предположений

Первый литературный конкурс состоялся
в апреле 2025 рамках конференции DataFusion



Контакты для продолжения дискуссии



Воронцов Константин Вячеславович

д.ф.-м.н., профессор РАН,
зав. лабораторией
машинного обучения
и семантического анализа
Института ИИ МГУ,
зав. кафедрой ММП ВМК МГУ,
зав. кафедрой МОЦГ МФТИ,
г.н.с. ФИЦ ИУ РАН



k.vorontsov@iai.msu.ru

<http://www.MachineLearning.ru/wiki?title=User:Vokov>



Цивилизационная
идеология
ДЗЕН-канал

<https://dzen.ru/civideology>

