

Искусственный интеллект:

ожидания,
ограничения
и опасности

Взгляд футуролога



02.03.18

И. В. Артюхов

ppt.com



О чём этот доклад:

- Что такое ИИ *на самом деле?*
- Какой бывает ИИ?
- Последние достижения в ИИ
- Некоторые возможные применения ИИ
 - Медицина
 - Управление
 - Искусство и т. д.
- Некоторые перспективы ИИ в



**А что это такое,
«искусственный интеллект»?**



**А что это такое,
«искусственный интеллект»?**

А что такое интеллект?



Что такое интеллект?

Интеллект (от лат. *intellectus* — понимание, познание) — общие способности к познанию, пониманию и разрешению проблем. Понятие интеллект объединяет все познавательные способности индивида: ощущение, восприятие, память, представление, мышление, воображение.

Интеллект — способность системы создавать в ходе самообучения программы (в первую очередь эвристические) для решения задач определенного класса сложности и решать эти задачи».

Интеллект — ...



Исходное определение

Искусственный интеллект (ИИ, англ. Artificial intelligence, AI) – Наука и технология создания интеллектуальных машин, особенно интеллектуальных компьютерных программ.

«Проблема состоит в том, что пока мы не можем в целом определить, какие вычислительные процедуры мы хотим называть интеллектуальными. Мы понимаем некоторые механизмы интеллекта и не понимаем остальные. Поэтому под интеллектом в пределах этой науки понимается только вычислительная составляющая способности достигать целей в мире»

Дж. МакКартни, 1956 г.



Современное понимание

Будем называть **искусственным интеллект** систему (например, программу, устройство или распределённую группу взаимодействующих устройств), способную решать интеллектуальные задачи (а также область науки, такие системы изучающую и технологию их создания).

Интеллектуальной задачей можно называть такую, которая традиционно считалась прерогативой человека или такую, которую умный («интеллектуальный») человек решает лучше глупого.



Какой бывает ИИ?



Классификация по способу построения

- **«Сверху вниз» («Top-Down»)**

Примеры: экспертные системы, базы знаний, системы логического вывода.

- **«Снизу вверх» («Bottom-Up»)**

Примеры: нейронные сети, эволюционные алгоритмы.

- **Гибридные.**

- **Роевой.**



Классификация по соотнесению с человеком

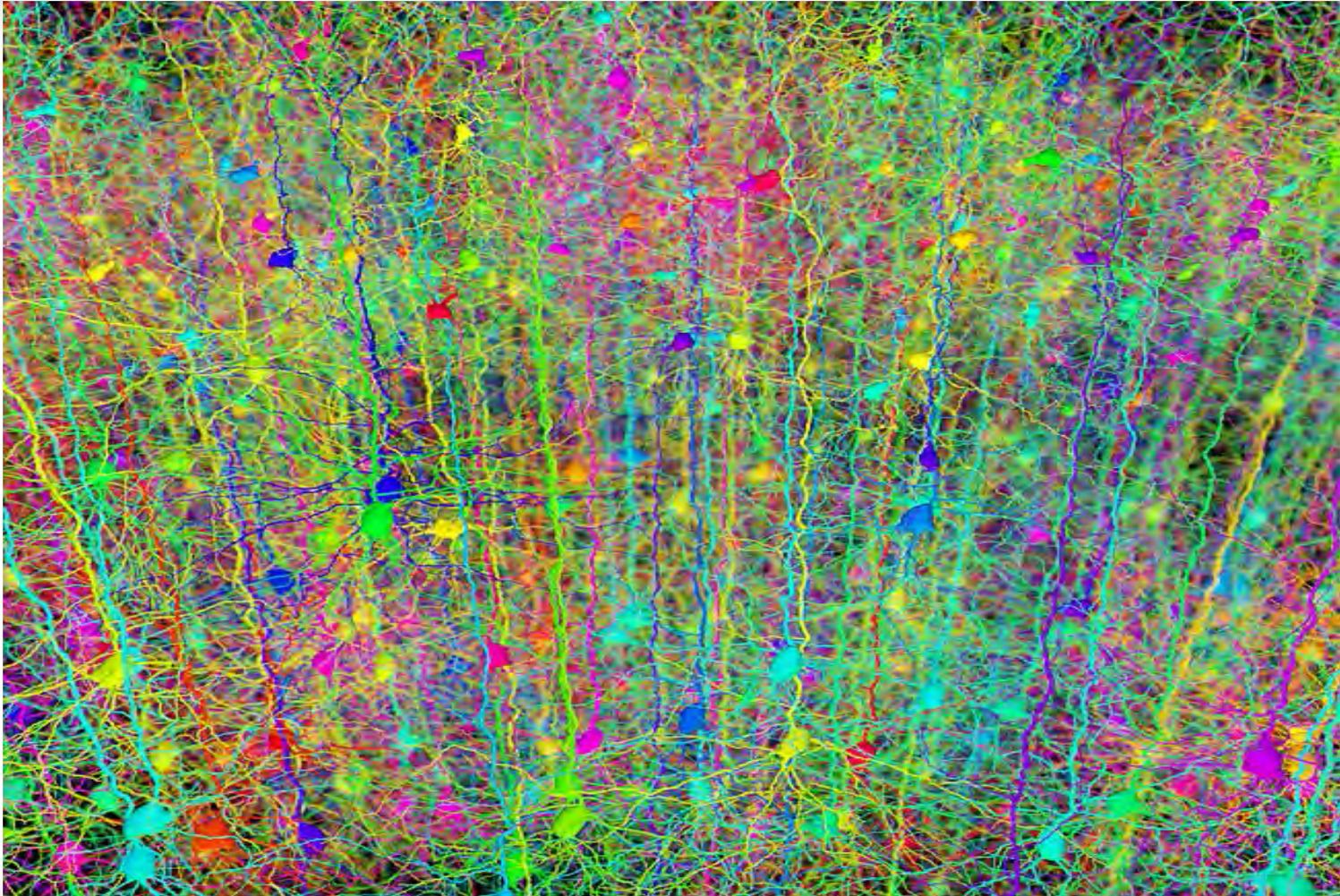
- **Специализированный искусственный интеллект.**
- **Универсальный искусственный интеллект (англ. Artificial General Intelligence, AGI) – ИИ, способный**



Нейронные сети

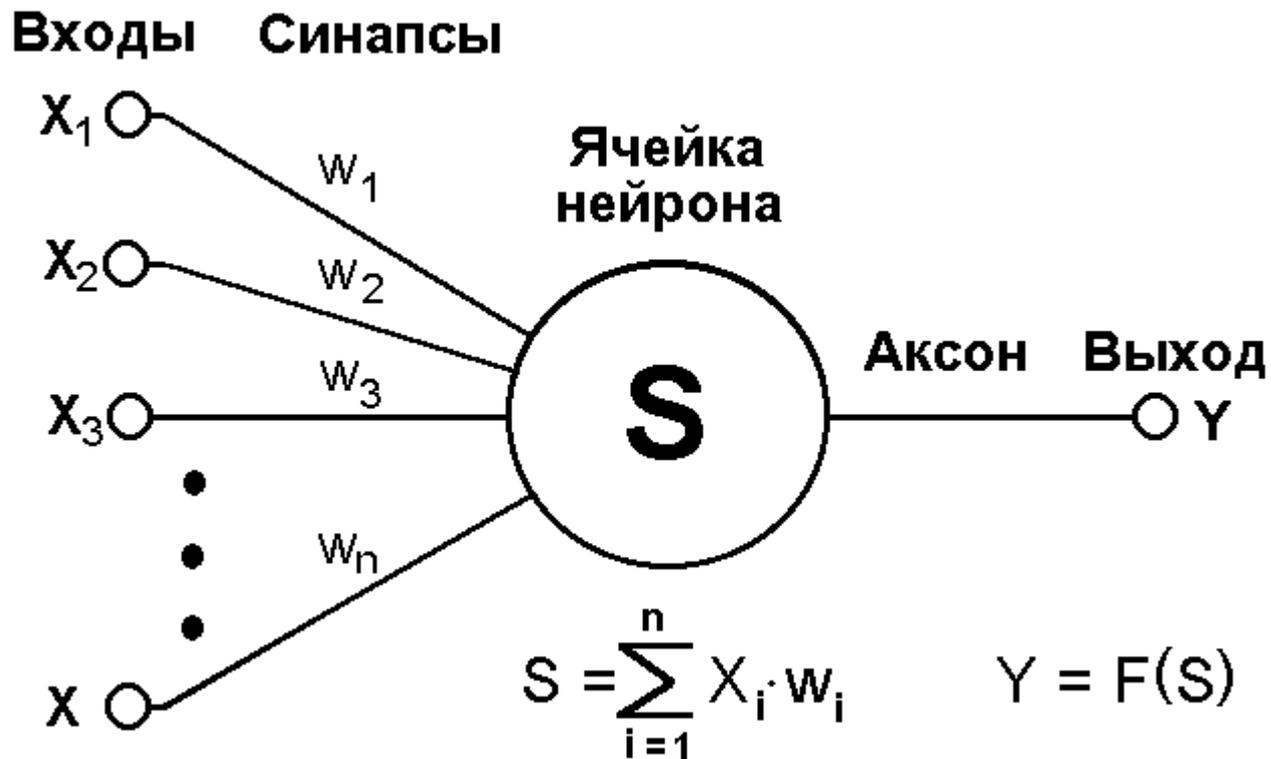


Нейронная сеть млекопитающего





Нейрон





Процессы в головном мозгу

- Очень быстрые (процессы в синапсах и ионных каналах)
- Быстрые (срабатывание нейронов)
- «Средние» (совместная работа групп нейронов, мозговые ритмы)
- Медленные (запоминание, обучение, формирование и др.)

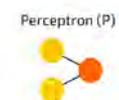


Некоторые структуры Искусственных НС

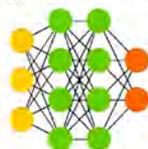
A mostly complete chart of Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

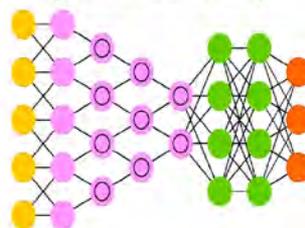
- Backfed Input Cell
- Input Cell
- △ Noisy Input Cell
- Hidden Cell
- Probabilistic Hidden Cell
- △ Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- △ Different Memory Cell
- Kernel
- Convolution or Pool



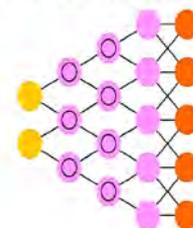
Deep Feed Forward (DFF)



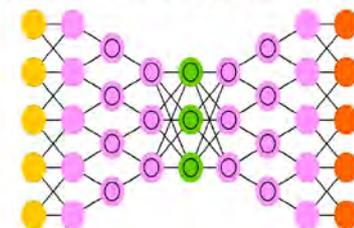
Deep Convolutional Network (DCN)



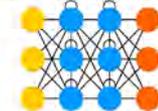
Deconvolutional Network (DN)



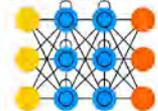
Deep Convolutional Inverse Graphics Network (DCIGN)



Recurrent Neural Network (RNN)



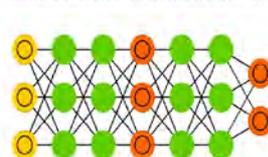
Long / Short Term Memory (LSTM)



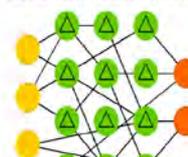
Gated Recurrent Unit (GRU)



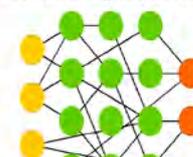
Generative Adversarial Network (GAN)



Liquid State Machine (LSM)



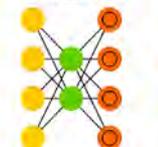
Extreme Learning Machine (ELM)



Echo State Network (ESN)



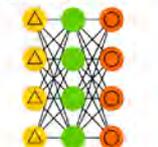
Auto Encoder (AE)



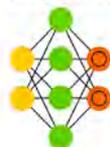
Variational AE (VAE)



Denoising AE (DAE)



Sparse AE (SAE)



Markov Chain (MC)



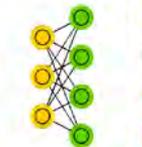
Hopfield Network (HN)



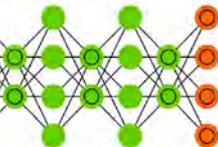
Boltzmann Machine (BM)



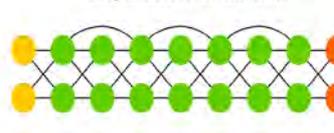
Restricted BM (RBM)



Deep Belief Network (DBN)



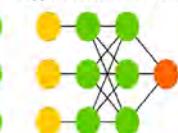
Deep Residual Network (DRN)



Kohonen Network (KN)



Support Vector Machine (SVM)



Neural Turing Machine (NTM)



<http://www.asimovinstitute.org/neural-network-zoo/>



Некоторые последние достижения в ИИ

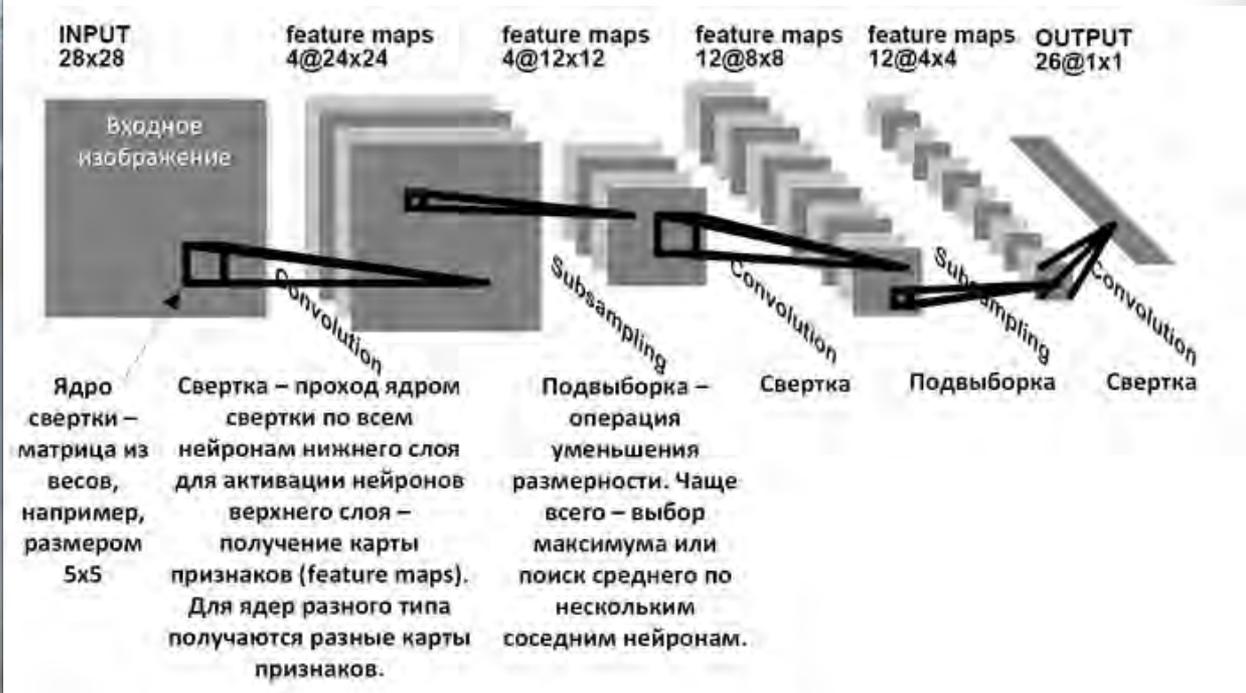


Победа над чемпионом в Го (2015 г., на 10 лет раньше прогнозов)





Компьютер AlphaGo





Беспилотные самолёты, автомобили, корабли





Распознавание лиц (и морд)

ferra.ru

Голландец выпускает кота домой с помощью системы распознавания морды

Новостная служба Ferra - Ferra.Ru - 1 марта 2018 г.



Голландский разработчик Аркайц Гарро (Arkaitz Garro), инженер-программист из WeTransfer, создал приспособление,

WAKE UP TO BEAUTIFUL
WITH BING



«Чтение мыслей»

КОНСОМОЛЬСКАЯ ПРАВДА В УКРАИНЕ
RSS PDA связь с редакцией
Читай свежий номер 1 Марта № 2205

ФОТО ВИДЕО ПОДПИСКА НА ГАЗЕТУ ВСЕ О КП РЕКЛАМА АКЦИИ И КОНКУРСЫ ПРЕСС-ЦЕНТР

Паралимпиада-2018 Суд о госизмене Януковича Уголовное дело Саакашвили Судьба хрущевки Актуальное

Новости Политика Экономика Общество Происшествия Спорт Светская хроника

Общество | Наука, техника, интернет

Александр ГОЛОВАСТОВ (12 января, 19:55) 557



"Мысли напоказ": создан ретранслятор образов из мозга человека [фото, видео]

Пока лучше всего система передает на экран изображения букв и цифр из мозга человека, но также может воссоздавать образы предметов, животных, зданий.

Японские исследователи из Китотского университета под руководством Юкиасу Камитани создали нейросеть, которая способна показывать образы из мозга человека,

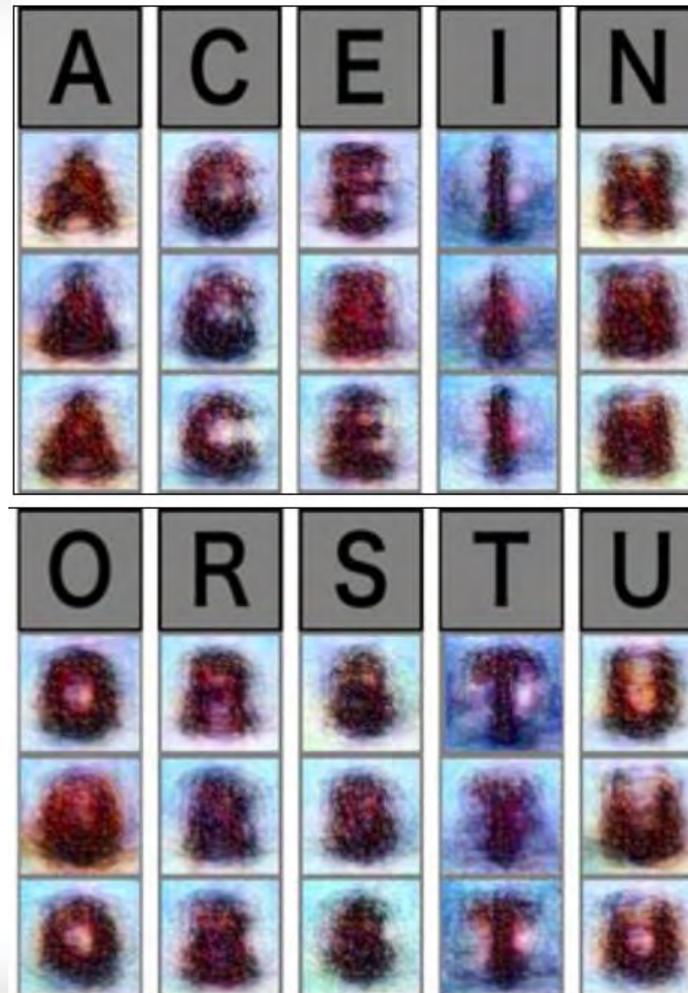
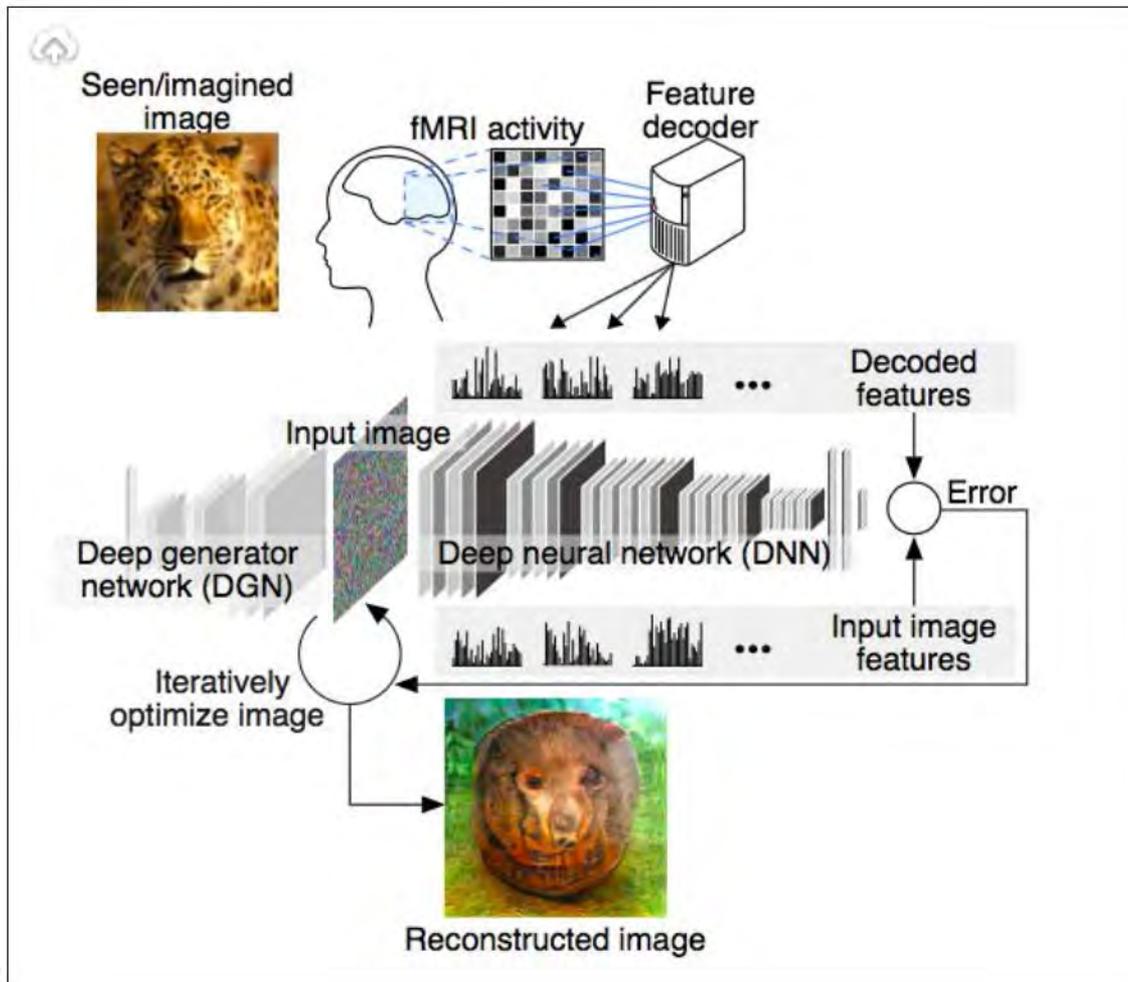




Схема процесса





ИИ учится играть, не зная правил

Geektimes Публикации Пользователи Хабы Компании Песочница

alizar 26 февраля 2015 в 12:17

ИИ от Google самостоятельно освоил 49 старых игр Atari

IT-компании, Игры и игровые приставки, Искусственный интеллект, Научно-популярное



Компания Google создала систему искусственного интеллекта, которая играет лучше человека во многие аркадные игры. Программа научилась играть, не зная правил и не имея доступа к коду, а просто наблюдая за картинкой на экране.



ИИ (сам!) научился жульничать

ferra.ru

Обзоры

Новости

Видео

Как выбрать

Ноутбуки и планшеты Телефоны Фото и видео Аудио Еще

Искусственный интеллект научился читерствовать в играх

1 марта 2018, 19:24



Комментарии

Трио исследователей машинного обучения из Фрайбургского университета создали искусственный интеллект, который научился использовать ошибки в игре ради победы.



Они исследовали обучения ИИ-агентов для навигации в видеоиграх методом проб и ошибок. Для этого использовались восемь игр для консоли Atari, в том числе классическая аркадная игра Q*bert из далеких 80-х годов прошлого века.

В США втайне от граждан испытали технологию предсказания преступлений

ОЛЕГ ИЛЮХИН

28.02.2018 21:03

5184



21



11





Наушники - переводчик

[На главную](#) / [Новости](#) / 05 октября 2017

Google Pixel Buds – умные беспроводные наушники, способные выступать в роли переводчика

Теги: [Google](#) Источник: [Google, The Verge](#)



Компания Google, как известно, лишила новые смартфоны Pixel разъёма для наушников. Понимая, что нужно при этом предложить альтернативу, поисковый гигант анонсировал свои беспроводные наушники Pixel Buds.



Вопреки трендам, это не полностью беспроводные наушники. Автономность заявлена на уровне пяти часов, а комплектный футляр увеличивает её до 24 часов.



Диагностика

Газета.ru

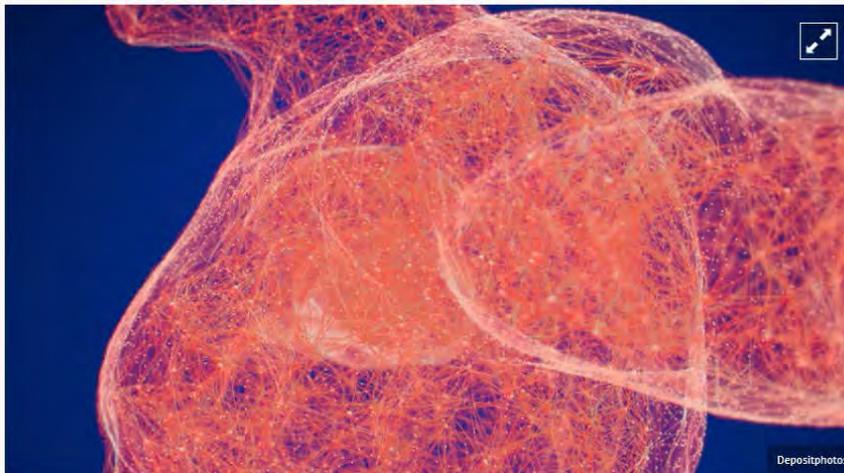


HOME [НОВОСТИ](#) [ПОЛИТИКА](#) [БИЗНЕС](#) [ОБЩЕСТВО](#) [АРМИЯ](#) [МНЕНИЯ](#) [КУЛЬТУРА](#) [НАУКА](#) [ТЕХНОЛОГИИ](#)
[Биология](#) [История](#) [Космос](#) [Медицина](#) [Мракобесие](#) [Наука и власть](#) [Социальные науки](#) [Технологии](#) [Физмат](#) [Лекции](#)

С одного взгляда: как нейросеть предсказывает инфаркт

Нейронная сеть научилась определять риск болезней сердца по сетчатке

Алла Салькова 25.02.2018, 13:14



Разработчики научили нейросеть по сетчатке выявлять риск развития сердечно-сосудистых заболеваний. Точность такой диагностики оказалась сравнима с результатами использования существующих методов. Создатели алгоритма рассчитывают, что со временем они улучшат качество диагностики.



Анализ генома

MIT
Technology
Review

Log in / Create an account Search

Subscribe

Topics+ The Download Magazine Events More+

Rewriting Life

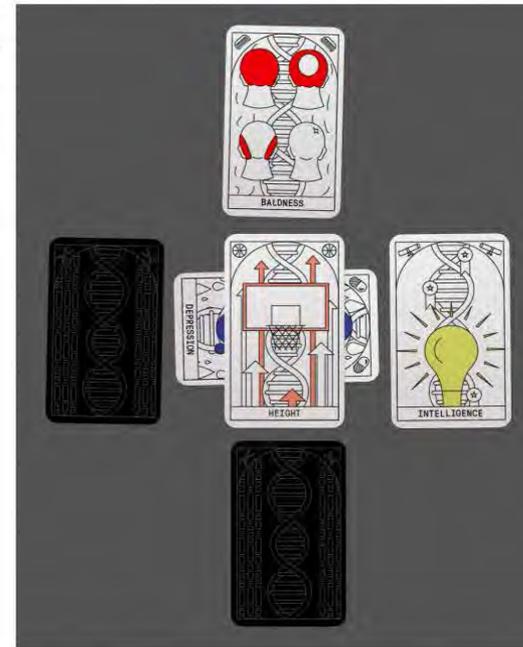
Forecasts of genetic fate just got a lot more accurate

DNA-based scores are getting better at predicting intelligence, risks for common diseases, and more.

by Antonio Regalado February 21, 2018

When Amit Khera explains how he predicts disease, the young cardiologist's hands touch the air, arranging imaginary columns of people: 30,000 who have suffered heart attacks here, 100,000 healthy controls there.

There's never been data available on as many people's genes as there is



SELMAN DESIGN

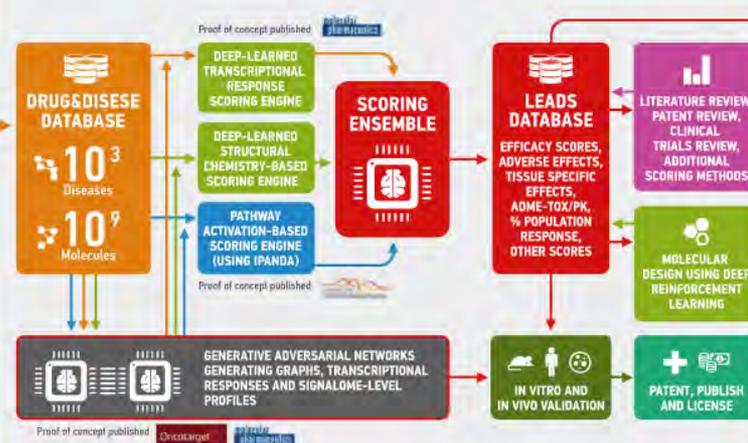


Компьютерная разработка НОВЫХ ЛЕКАРСТВ

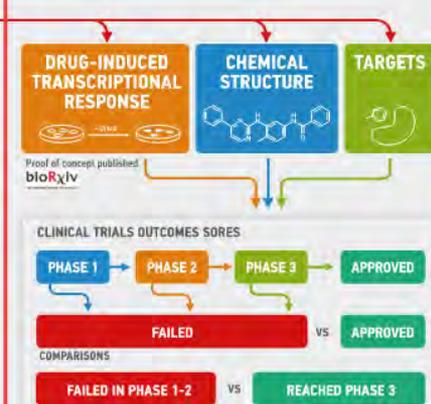
TARGET IDENTIFICATION PIPELINES (DISEASES + AGING)



GENERATION OF NOVEL SMALL MOLECULE LEADS



PREDICTORS OF CLINICAL TRIAL OUTCOMES





Создание видео по сценарию

 Artificial intelligence is moving into movie production. SHAREGRID/UNSPLASH

New algorithm can create movies from just a few snippets of text

By **Matthew Hutson** | Feb. 23, 2018, 4:35 PM

Screenwriters denied the big budgets and formidable resources of the major film studios may soon have another option, thanks to a new algorithm that can generate a video simply by consuming a (very short) script. The new movies are far from Oscar-worthy, but a similar technique could one day find uses outside entertainment, by, say, helping a witness reconstruct a car crash or a crime.

Artificial intelligence (AI) is getting much better at **identifying the content of images** and providing labels. So-called “generative” algorithms go the other way, producing images from labels (or **brain scans**). A few can even take a single movie frame and predict the next series of frames. But putting it all together—creating an image from text and making it move realistically in accordance with the text—has not been done before.

“As far as I know, it’s the first text-to-video work that gives such good results. They are not perfect, but at least they start to look like real videos,” says Tinne Tuytelaars, a computer scientist at Katholieke Universiteit Leuven in Belgium, who has done her own video prediction research. “It’s really nice work.”



Всегомогущ ли ИИ потенциально?

- **Основанный на алгоритмах, правилах, эвристиках и т.д.**

Не сможет решать задачи, для которых его создатель не способен составить алгоритмы, правила, эвристики и т. д.

- **Самообучающийся**

Не сможет решать задачи, для которых нет обучающих выборок

- **Гибридный, роевой**
?



Практические применения ИИ



Практические применения ИИ

- В медицине
- В управлении сложными системами
- В роботах и дронах
- В борьбе с преступностью
- В научных исследованиях
- В индустрии развлечений...



Практические применения ИИ

- В медицине
- В управлении сложными системами
- В роботах и дронах
- В борьбе с преступностью
- В научных исследованиях
- В индустрии развлечений...
- **Да везде!**



В медицине

- Диагностика
- Выбор лечения
- Роботы-хирурги
- Роботы-няни
- Умные протезы, экзоскелеты и пр.
- Поиск новых лекарств
- ...



В управлении системами

- Транспортными
- Энергетическими
- Экономическими
- Социальными
- ...



В индустрии развлечений

- ИИ в играх
- Генерация текстов (и даже стихов)
- Сочинение музыки
- Генерация изображений
- Генерация видео
- ...



Что УЖЕ может ИИ

- Обыгрывать чемпионов в jeopardy, шахматы, го и покер
- Писать картины, музыку, стихи, сценарии к фильмам, сводки новостей...
- Общаться на естественном языке
- Переводить с языка на язык
- Обрабатывать гиганские массивы данных – классифицировать, кластеризовать, прогнозировать...
- Распознавать лица, голоса, эмоции, породы собак...
- Водить автомобили
- Разрабатывать новые ИИ



Перспективы ИИ



ИИ и другие прорывные технологии

- Робототехника, дроны
- Новые материалы и технологии
 - Нанoeлектроника
 - Фотоника, спинотроника и др.
 - Квантовые компьютеры
- Большие данные
- Интернет вещей
- Децентрализованные системы
- Геномика и генная инженерия
- Социальные технологии



Нетривиальные перспективы

- Искусственные личности
- Глобальный распределённый ИИ
- Квантовый ИИ



На пути к распределённому глобальному ИИ

[Home](#) [Projects](#) [Publications](#) [Working Papers](#) [Bibliography](#) [Links](#) [Seminars](#) [Videos](#) [News Feed](#) [Social Media](#)

THE GLOBAL BRAIN INSTITUTE

NEWS:

- *David R. Weinbaum (Weaver) and Viktoras Veitas have received the [Kurzweil Prize for Best AGI Idea](#) for their concept of [Open Ended Intelligence](#)*
- *Francis Heylighen has received the [2015 Outstanding Technology Award](#) from the [Web Intelligence Consortium](#): check out the [Press Release](#)*

The [Global Brain](#) can be defined as the distributed intelligence emerging from the Internet. The Global Brain Institute (GBI) was founded in 2012 at the [Vrije Universiteit Brussel](#) (VUB) to research this phenomenon. The GBI grew out of the [Global Brain Group](#), an international community of researchers created in 1996, and the [Evolution, Complexity and Cognition](#) research group at the VUB.

The GBI uses scientific methods to better understand the global evolution towards ever-stronger connectivity between people, software and machines. By developing concrete models of this development, we can anticipate both its promises and its perils. That would help us to steer a course towards the best possible outcome for humanity.

Members

Director

- [Francis Heylighen](#): overall theoretical development

Managing Director

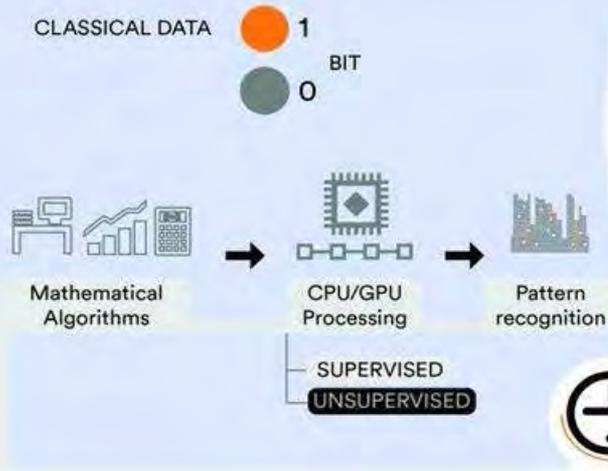
- [Marta Lenartowicz](#): new projects development

Researchers

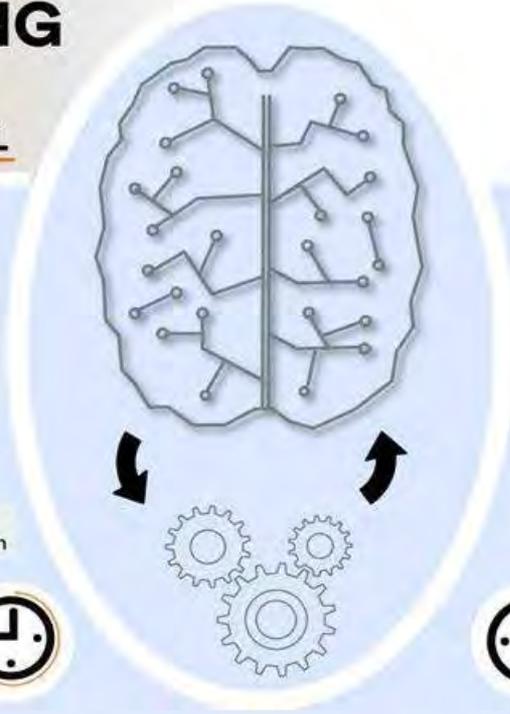
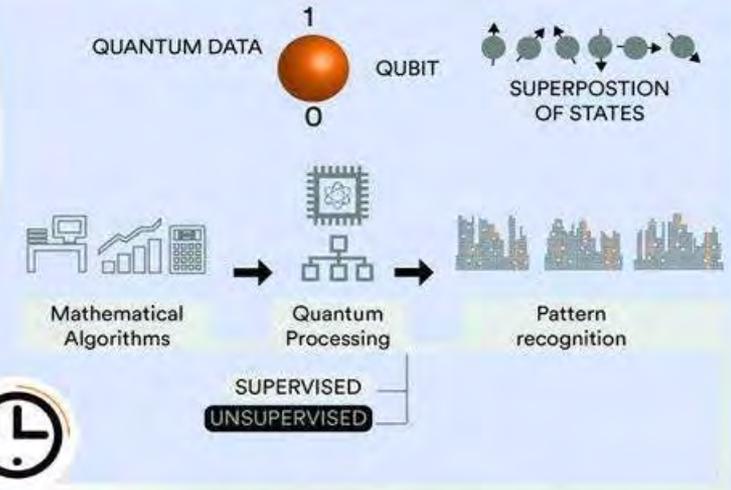
- [Shima Beigi](#): sustainability and resilience of complex systems
- [Evo Busseniens](#): mathematical modeling of self-organization
- [Vincenzo de Florio](#): resilient and fractal socio-technological systems
- [Jon Echanove](#): global organizations and dealing with uncertainty
- [Mixel Kiemen](#): networking, social media and modeling of innovation
- [Cadell Last](#): future human metasystem, big history, and media presence

MACHINE LEARNING

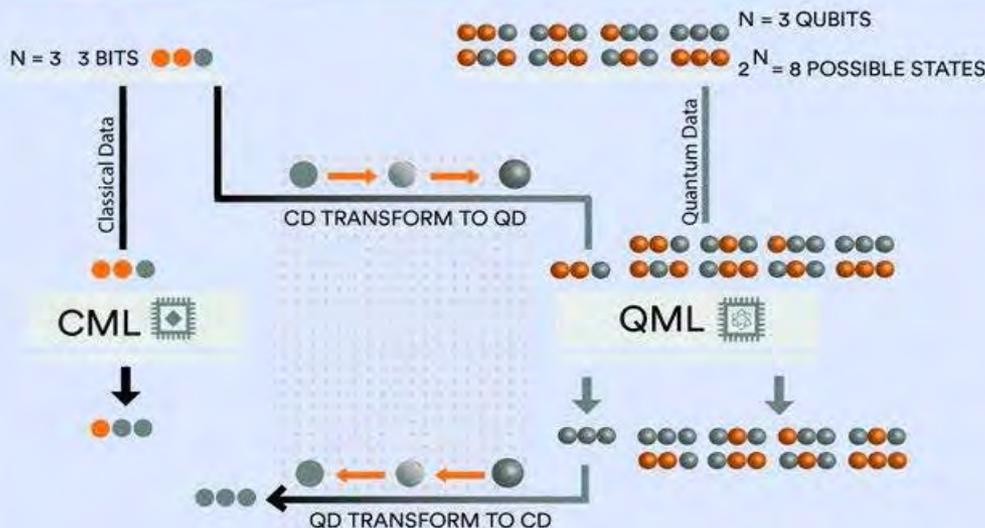
CLASSICAL MACHINE LEARNING - CML



QUANTUM MACHINE LEARNING - QML



PROCESSING METHODS



APPLICATIONS





Опасности ИИ



Распространённые страхи

- ИИ осознает себя и захочет править миром!
- ИИ реш(ит/ат), что люди ему/им уже не нужны!
- ИИ испугается, что люди его выключат и на всякий случай их уничтожит!
- . . .



Войны между человеком и ИИ?



Войны между человеком и ИИ?



**Идут
вовсю!**



Опасности реальные и мнимые

Как *не может* возникнуть враждебный суперинтеллект:

- В результате «осознания себя»
- Из-за стремления к власти
- Из-за страха «что выключат»

Как *может* возникнуть «плохой» суперинтеллект:

- В результате ошибки при его разработке
- В результате чьих либо злонамеренных или безответственных действий



**Меня пугает не искусственный
разум, а человеческая глупость.
*Нейл Якобштейн***



Сценарий: «Плохой ИСИ»

- Действия Супер-ИИ будут принципиально непонятны для разума человека
- Заложенные в Супер-ИИ принципы, изначально казавшиеся разумными и очевидными, в перспективе могут привести к неожиданным действиям Супер-ИИ
- Супер-ИИ будет превосходить человека настолько, что сможет легко его убедить в правильности своих действий



Сценарий: «Хороший, но ненадёжный ИСИ»

ИСИ работает прекрасно, справляясь с самыми сложными проблемами намного лучше людей. Человечество доверяет ему всё более ответственные задачи – такие, как управление глобальными транспортными, энергетическими, финансовыми и информационными потоками, становясь всё более зависимым от него.



Сценарий: «Хороший, но ненадёжный ИСИ»

ИСИ работает прекрасно, справляясь с самыми сложными проблемами намного лучше людей. Человечество доверяет ему всё более ответственные задачи – такие, как управление глобальными транспортными, энергетическими, финансовыми и информационными потоками, становясь всё более зависимым от него.

И тут он ломается...



Решение: несколько (или много) ИСИ

Они дублируют и контролируют друг друга, проверяя предлагаемые решения и находя в них ошибки.

ИСИ должны быть:

1. Независимы
2. Разнообразны (разные конструкции, разная история обучения).

Должна быть разработана процедура принятия решений при наличии разногласий.



Социальные проблемы

Воображаемая опасность:

Все рабочие места займут роботы и ИИ, а люди будут обречены на нищету.

Реальная опасность:

Все рабочие места займут роботы и ИИ, а люди привыкнут к безделью и изобилию и превратятся в иждивенцев.



Рабочие места, которые могут быть отобраны уже скоро

- Водители транспортных средств
- Переводчики
- Юристы
- Врачи-диагносты
- Киноактёры
- Многие другие, иногда совершенно неожиданные.



Рабочие места, которые могут быть отобраны уже скоро

- Водители транспортных средств
- Переводчики
- Юристы
- Врачи-диагносты
- Киноактёры
- Многие другие, иногда совершенно неожиданные.
- **Специальность, которая, как я думаю, исчезнет последней: Учитель этики для ИИ.**



Какова будет роль человека в мире, управляемом ИИ?

Она будет очень велика — мы будем создавать для ИИ *цели и ценности*. ИСИ не может сам решить, что хорошо, а что плохо — этому научить его должны будем мы. Систему ценностей надо будет закладывать в процессе создания и обучения ИИ, подтверждать и уточнять в ходе дальнейшего самообучения на модельных ситуациях и реальной работе.

Это будет очень непростым делом — хотя бы потому, что мы и сами не в состоянии сформулировать полную и непротиворечивую систему ценностей.

Но делать это кроме нас будет некому.



Сценарий: Сингулярность (по В. Винже)

Мы создаём искусственный интеллект (ИИ)
более мощный, чем человеческий

... который создаёт новый ИИ, более мощный,
чем свой

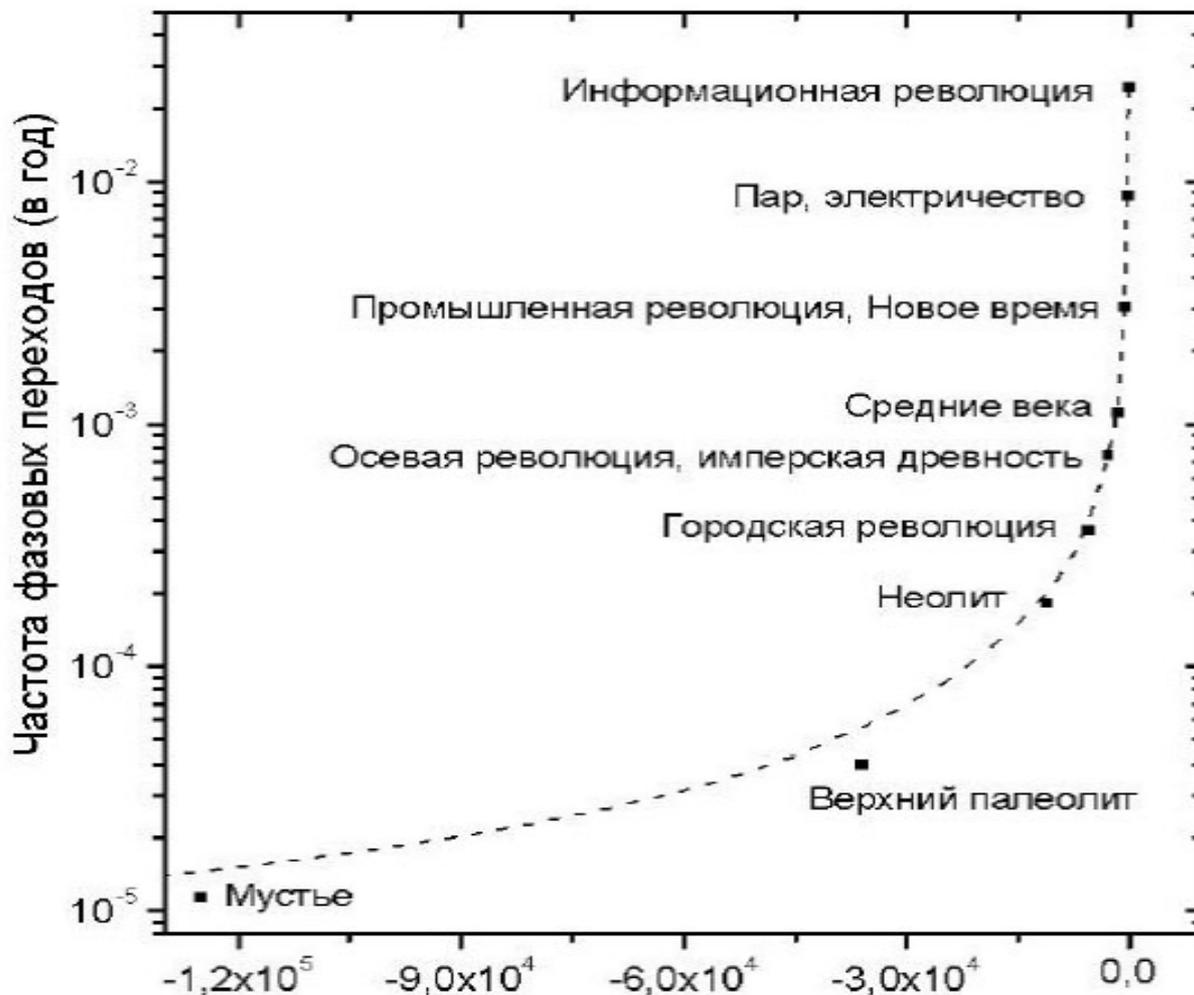
... который создаёт...

...

После чего наступает «горизонт событий»
(сингулярность) - момент во времени, за
который заглянуть невозможно, как и
целенаправленно повлиять на происходящее
после него.



Будущее наступает всё быстрее





Ещё сценарии

- Слияние человека и машины в единый суперинтеллект, как следствие развития интерфейсов мозг-компьютер.
- Загрузка сознания человека на искусственный носитель; жизнь в теле робота или в виртуальной реальности.
- Дублирование, резервирование, синхронизация и слияние личностей.
- Слияние естественных и искусственных интеллектов в планетарный (межпланетный, межзвёздный...) сверхразум.



Немного мотивации

- **Будущее наступит** - хотим мы этого или не хотим. Вопросы в том, каким оно будет и будем ли мы к нему готовы.
- **Точно предсказать, каким оно будет, невозможно.** Но если к нему не готовиться - мы в любом случае окажемся не готовы.
- **Будущее не фиксировано.** Каким оно будет, зависит от наших действий (в т. ч., бездействия) сегодня. За последствия нашего бездействия мы несём такую же ответственность перед собой и другими, как и за последствия активных действий.



Спасибо за внимание!

Вопросы?