

Искусственный интеллект: наука или технология?

А.И. Панов

Московский физико-технический институт
Лаборатория когнитивных динамических систем

Летняя школа «Комбинаторика и алгоритмы»

panov.ai@mipt.ru



Кратко о себе

Панов Александр Игоревич, к. ф.-м. н.

- Старший научный сотрудник лаборатории динамических интеллектуальных систем ИСА ФИЦ ИУ РАН (рук. Г.С. Осипов).
- Научный сотрудник и доцент ФКН ВШЭ.
- Доцент кафедры системных исследований и зам. зав. лаборатории когнитивных динамических систем МФТИ.
- Член Российской ассоциации искусственного интеллекта (РААИ).
- Член Сообщества биологически инспирированных когнитивных архитектур (BICA Society).
- Организатор Международной конференции по биологически инспирированным когнитивным архитектурам (BICA-2016 — Нью-Йорк, BICA-2017 — Москва), Международной школы по биологически инспирированным когнитивным архитектурам (Fierces on BICA, Москва) и школы молодых ученых по ИИ (ISyT 2017, Санкт-Петербург).
- Руководитель проектов РФФИ и РНФ.
- Ментор студенческой лаборатории по ИИ (SLabAI).



Искусственный интеллект как мечта

Писателей и ученых давно волновала идея создания машины, способной мыслить как человек.



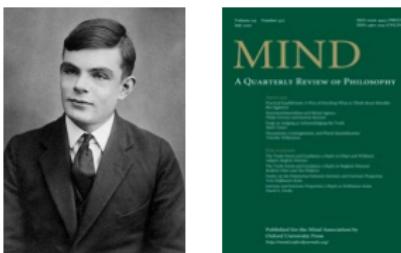
В 1832 г. российский изобретатель Семён Корсаков предложил серию «интеллектуальных машин», работающих на перфокартах. В 1942 г. американский писатель Айзек Азимов в рассказе «Хоровод» сформулировал знаменитые правила поведения для роботов — три закона робототехники.



Но возможна ли реализация этой идеи в принципе?

Искусственный интеллект как мечта

В 1950 г. в философском журнале «Mind» появилась статья английского ученого Алана Тьюринга под названием «Вычислительные машины и разум» (Computing Machinery and Intelligence), в которой предложил знаменитый тест.



Аргументы против ИИ:

- **Теологический:** мышление как функция бессмертной души.
- **Этический:** последствия создания ИИ непредсказуемы.
- **Математический:** теорема Гёделя и возражения Пенроуза.
- **Психологический:** что такое сознание.
- **Нейрофизиологический:** непрерывность нервной системы.
- **Экстрасенсорный:** отсутствие телепатии.

Искусственный интеллект как мечта

Возможно ли создание машины, способной мыслить как человек? Два ответа — две гипотезы:

- теория сильного искусственного интеллекта предполагает, что компьютеры могут приобрести способность мыслить и осознавать себя, хотя и не обязательно их мыслительный процесс будет подобен человеческому.
- теория слабого искусственного интеллекта отвергает такую возможность.

Мысленный эксперимент американского философа Джона Сёрла, предложенный им в 1980 г.



ИИ как научное направление

Определения искусственного интеллекта как научного направления:



ИИ - это научное направление, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными (Дмитрий Поступов).

ИИ как научное направление

Определения искусственного интеллекта как научного направления:



ИИ - это научное направление, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными (Дмитрий Поступов).



ИИ - это наука и технология создания интеллектуальных машин, особенно интеллектуальных компьютерных программ (Джон Маккарти).

ИИ как научное направление

Определения искусственного интеллекта как научного направления:



ИИ - это научное направление, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными (Дмитрий Поступов).

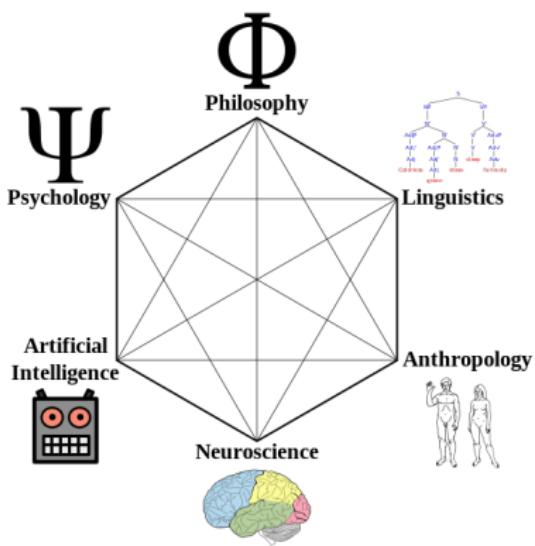


ИИ - это наука и технология создания интеллектуальных машин, особенно интеллектуальных компьютерных программ (Джон Маккарти).



ИИ - это наука об «интеллектуальных агентах», т.е. о некотором устройстве или программе, которая воспринимает свою среду и выполняет действия, которые максимизируют ее шансы на успех при достижении какой-то цели (Питер Рассел).

Когнитивные науки



Когнитивная наука (лат. *cognitio* «познание») –
междисциплинарное научное направление изучающее психику,
разум (mind) человека и реализующие его процессы.

Программная реализация ИИ

- **1954 г.** — аналитики *Рэнд Корпорейшн, Аллан Ньюэлл, Дж. Шоу и Герберт Саймон* решили написать программу игры в шахматы. В этой затее им вызвались помочь *А. Тьюринг* и *К. Шенон*, а также группа голландских психологов.
- **1956 г.** — Дартмутский семинар (конференция по вопросам искусственного интеллекта, проведённая летом 1956 года в Дартмутском колледже).
- **1957 г.** — программа для игры в шахматы (*NSS*) была написана. В основе работы *NSS* лежали эвристики — правила выбора в отсутствие теоретических оснований.

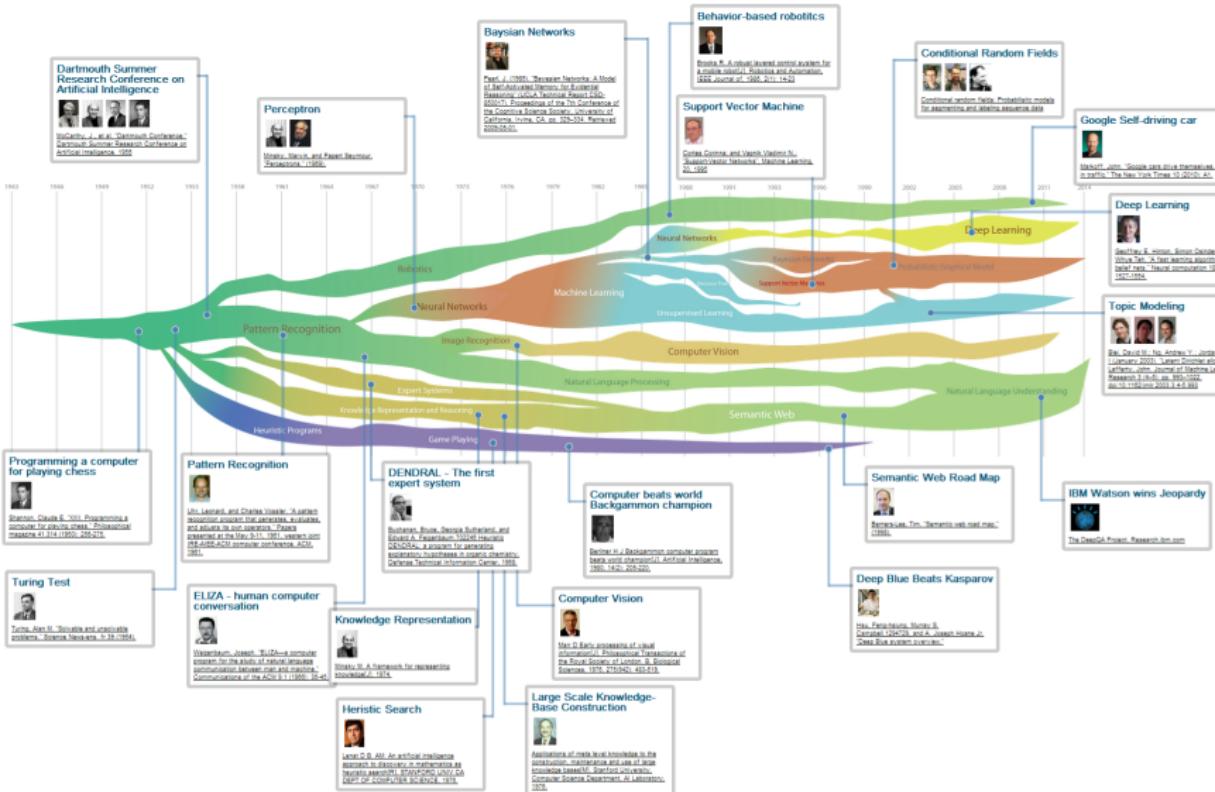
Дальнейшее развитие

- **1960 г.** — GPS («универсальный решатель задач», Аллан Ньюэлл и Герберт Саймон): вычисление неопределенных интегралов, головоломки и некоторые другие задачи. Программы автоматического доказательства теорем из планиметрии, решения алгебраических задач.
- **1960 г.** — возникновение эвристического программирования.
- **1963 г.** — Джон Маккарти — ЛИСП. Возникновение функционального программирования.
- **1968 г.** — книга «Семантическая обработка информации» Марвина Минского.

Поиск непереборных методов решения задач

- 1964 г. — В.Н. Пушкин и Д.А. Поспелов — модельная гипотеза мышления versus лабиринтной; методы решения переборных задач человеком.
- 1964 г. — С.Ю. Маслов — метод автоматического поиска доказательства теорем в исчислении предикатов (обратный метод).
- 1965 г. — Дж.А. Робинсон — метод автоматического поиска доказательства теорем в исчислении предикатов (метод резолюций).
- 1968 г. — возникновение логического программирования.
- 1971 г. — А. Колменрауэр — язык Пролог.
- 1976 г. — Гипотеза Саймона и Ньюэлла: «Физическая символьная система имеет необходимые и достаточные средства для произведения базовых интеллектуальных действий».

Современный ИИ



Современный ИИ

Сформировались два основных подхода к разработке систем искусственного интеллекта:

- **нисходящий** (Top-Down AI), символьный: создание экспертных систем, баз знаний и систем логического вывода, имитирующих высокоуровневые психические процессы: мышление, рассуждение, речь, эмоции, творчество и т. д.;
- **восходящий** (Bottom-Up AI), биологический: изучение нейронных сетей и эволюционных вычислений, моделирующих интеллектуальное поведение на основе биологических элементов, а также создание соответствующих вычислительных систем.

Перспективное направление - объединение: знаковые системы, нейросимвольные вычисления.

Искусственный интеллект — организационная структура

- Во многих странах есть ассоциации искусственного интеллекта (EurAI, AAAI, RAAI).
- Каждый или раз в два года проходят крупнейшие конференции по ИИ: ECAI, AAAI Conference, IJCAI, КИИ.
- РААИ - общероссийская общественная организация (261 индивидуальный член, 45 региональных отделений).
- Выходят тематические журналы (Artificial Intelligence, Cognitive Psychology, Autonomous Robots, Neural Networks, Искусственный интеллект и принятие решений).



ИСТОЧНИКИ

Онлайн: Coursera, Udacity, postnauka.ru, курсы ведущих университетов:

- Machine Learning and Artificial Intelligence в Принстоне (<https://www.cs.princeton.edu/courses/archive/fall116/cos402/>)
- Artificial Intelligence в CMU (<http://www.cs.cmu.edu/~./15381/>)
- Deep Learning for Self-Driving Cars в MIT (<http://selfdrivingcars.mit.edu>)
- Deep Reinforcement Learning в Беркли (<http://rll.berkeley.edu/deeprlcourse/>)

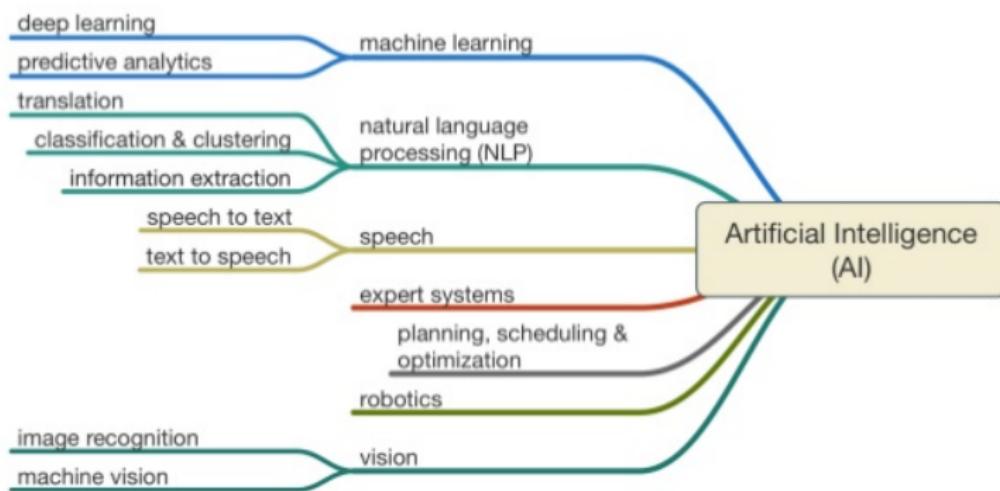
Книги:

- Nilsson N.J. Artificial Intelligence: A New Synthesis. San Francisco: Morgan Kaufmann, 1998. 513 р.
- Russell S., Norvig P. Artificial Intelligence: A Modern Approach, Prentice Hall, 2009 (3 edition), Р. 1152. (Искусственный интеллект. Современный подход)
- Осипов Г.С. Методы искусственного интеллекта. М.: ФИЗМАТЛИТ, 2011.
- Fireani D., Mattiussi C. Bio-Inspired Artificial Intelligence: Theories, Methods, and Technologies, The MIT Press, 2008, Р. 658.
- Поспелов Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов. М.: Радио и связь, 1989. 184 с.
- Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота. Модели поведения. М.: Наука, 1987. 288 с.
- Что-то «популярное» (Джефф Хокинс, Сандра Блейксли. Об интеллекте; Роджер Пенроуз. Новый ум короля)

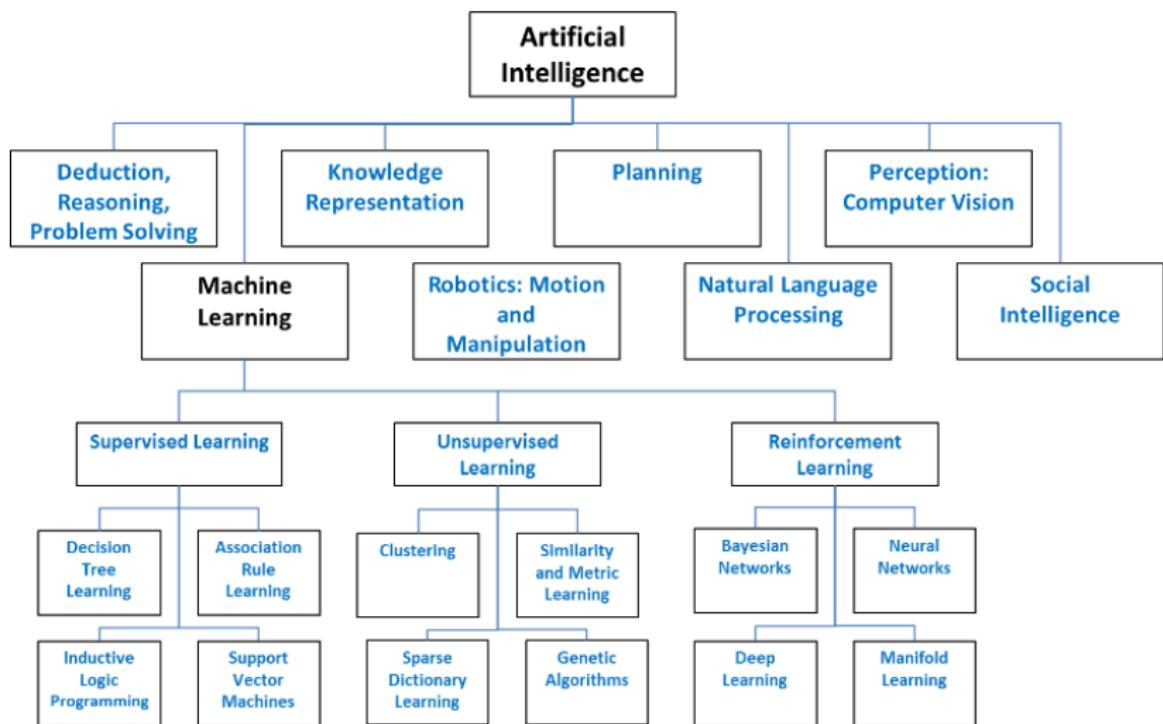
Профессиональные интернет-ресурсы (<http://open.ai>, <https://deepmind.com>)

Основные направления ИИ

Artificial Intelligence (AI)



Основные направления ИИ



Основные направления ИИ



Приобретение знаний, анализ данных и автоматическое порождение гипотез

Цель: создание методологий, технологий и программных средств обнаружения и переноса компетентности в базы знаний.

Методы приобретения знаний:

Машинное обучение и обучение по примерам

(методы построения деревьев решений, индуктивные методы построения правил; статистические методы, в частности, Байесовские сети; метод ближайших соседей, искусственные нейронные сети)

Приобретение знаний из текстов

Прямые методы приобретения знаний (автоматизированный диалог с экспертами)

Представление знаний

Предмет: разработка языков и программных средств для описания экспертных и эмпирических знаний.

Содержание:

- семантические сети, системы фреймов, системы правил (продукционные системы) и их гибриды;
- логики пространства и времени;
- онтологии – способ обмена знаниями;
- дескриптивные логики (теория баз знаний и онтологий).

Автоматизация рассуждений

Методы индукции, абдукции и аналогии, аргументации, рассуждения на основе прецедентов, на основе ограничений, рассуждения о действиях и изменениях, рассуждения с неопределенностью, немонотонные рассуждения.

Немонотонные рассуждения связаны с поиском эмпирических зависимостей в данных, обучением по примерам и рассуждениями в эмпирических теориях. Выделились в самостоятельный раздел логики.

Рассуждения о действиях исследуют связь действий и эффектов действий (результатов действий).

Рассуждения с неопределенностью — использование Байесовского формализма в моделях рассуждений.

Многоагентные системы

Изучаются интеллектуальные программные агенты, их коалиции и поведение.

Интеллектуальный программный агент — программная система, обладающая автономностью, социальными чертами, реактивностью и активностью.

Основные проблемы: коммуникация интеллектуальных агентов, разработка языков для этой цели, координация поведения агентов, распределение ролей в коалициях агентов, коллективное поведение агентов.

Роботы и автономные системы

- Диалоговое взаимодействие коалиций мобильных роботов.
 - Интерпретация команд, поступающих от человека.
 - Качественные логики пространства-времени.
 - Рассуждения, основанные на оценках.
 - Проблема символизации (symbol grounding problem).

Интеллектуальные динамические системы и автоматическое планирование поведения

Результат интеграции методов искусственного интеллекта с теорией динамических систем:

- планирование,
- моделирование,
- управление.

Обработка естественного языка, интерфейс и модели пользователя

- Семантический поиск в больших массивах текстов:
 - поиск документов (в полнотекстовой БД, в локальных и глобальных телекоммуникационных сетях);
 - извлечение данных из текстов; извлечение знаний из текстов.
 - Обработка текстов: сегментация, классификация, кластеризация, аннотирование или реферирование текстов. Перевод.
 - Диалоговые системы:
 - интеллектуальные вопросно-ответные системы;
 - системы общения конечных пользователей с БД, предоставляющие различные услуги (выполнение банковских операций по телефону, заказ товаров по каталогам);
 - голосовое управление техникой, кооперативное решение проблем (человек плюс интеллектуальная система).
 - Автоматическое обучение анализу текстов.

Нечеткие модели и мягкие вычисления

- Нечеткие схемы вывода по аналогии;
- теория нечетких мер;
- модели геометрических объектов;
- алгоритмы эволюционного моделирования с динамическими параметрами (например, время жизни и размер популяции);
- методы решения оптимизационных задач с использованием технологий генетического поиска, гомеостатических и синергетических принципов и элементов самоорганизации.

Вклад ИИ в другие науки

Развитие ИИ привело к **возникновению самостоятельных областей:**

- эвристическое программирование,
- функциональное программирование,
- логическое программирование,
- объектно-ориентированное программирование,
- теория немонотонных рассуждений и немонотонные логики,
- инженерия знаний,
- технология программирования, основанная на знаниях,
- прикладная семиотика.

В инженерном направлении:

- экспертические системы.

Перспективные направления ИИ

- Рассуждения, основанные на прецедентах.
- Рассуждения о пространстве — возрастающее значение для автономных мобильных устройств, анализа изображений (в частности, аэрофотоснимков), синтеза текстовых описаний по изображениям.
- Методы машинного обучения и автоматического формирования гипотез — решение практических задач: от обнаружения закономерностей в данных до повышения степени адаптивности различных технических устройств.
- Подходы, основанные на технологии интеллектуальных агентов перспективны при разработке больших программных систем.

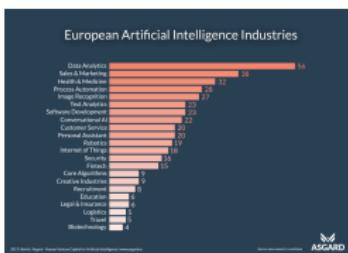
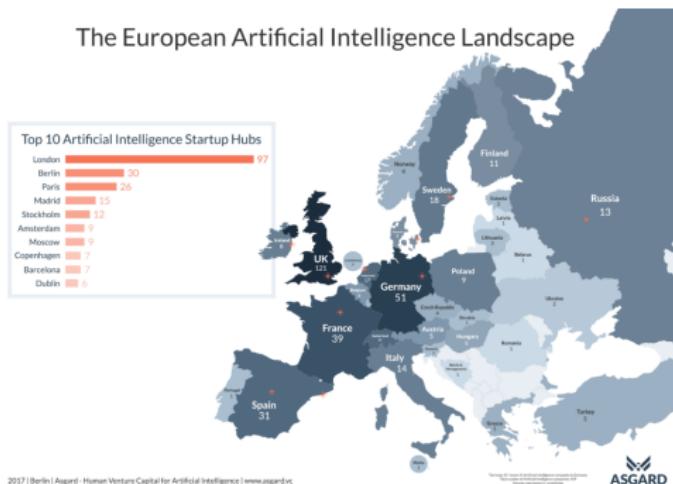
Перспективные направления ИИ

- **Влияние идей и методов ИИ на машинный анализ текстов на естественном языке** — коснется семантического анализа и методов синтаксического анализа — в этой области оно проявится в учете модели мира и использовании знаний о предметной области для уменьшения переборов на более ранних стадиях анализа.
- **Понимание текста.**
- **Автоматическое планирование и управление поведением.** Область применения - от бытовой техники до беспилотных аппаратов для исследования глубокого космоса.

Проблемы

- Переход от моделирования структурной организации к моделированию ментальных представлений, в частности, когнитивных функций, иначе говоря, от искусственного интеллекта — к искусенному сознанию.
 - Автоматическое (или полуавтоматическое) формирование интеллектуальными агентами модели мира, включая зрительные и слуховые образы предметов и их назначение.

ИИ в индустрии



ИИ стартапы

Machine Learning-Gen (123 Companies)



Machine Learning-App (260 Companies)



Computer Vision-Gen (106 Companies)



Computer Vision-App (83 Companies)



Smart Robots (65 Companies)



Virtual Personal Assistants (92 Companies)



Artificial Intelligence

Contact
info@venturescanner.com
 to see all 957 companies

Speech to Speech Trans. (15 Companies)



Context Aware Comp. (28 Companies)



Gesture Control (33 Companies)



Recommendation Eng. (60 Companies)

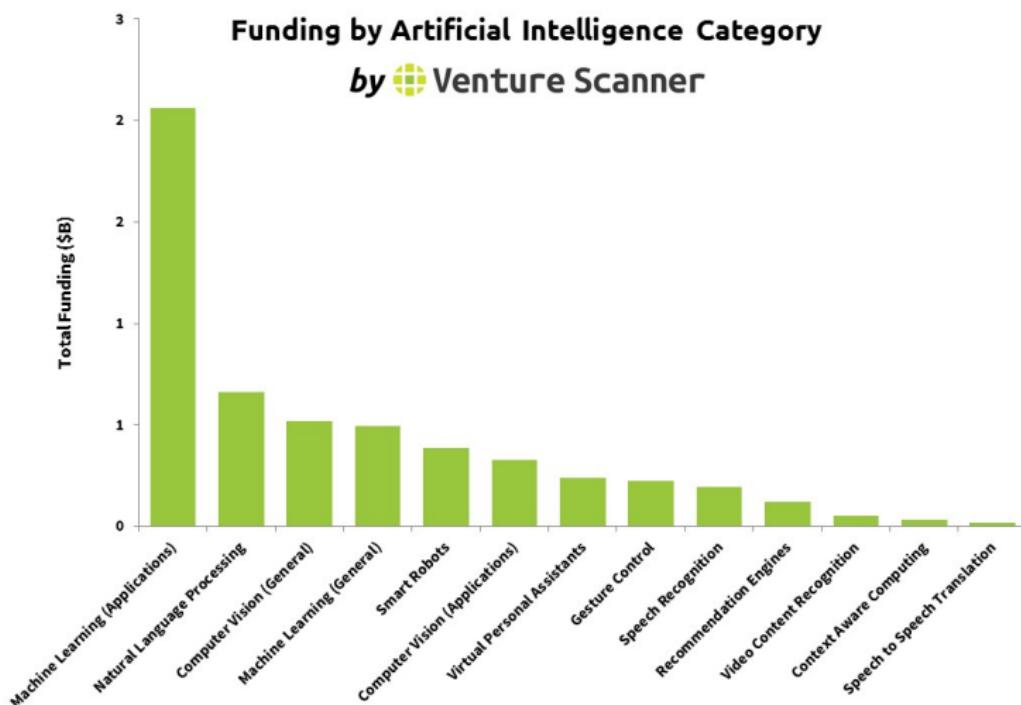


Video Content Recog. (14 Companies)



Venture Scanner

Финансирование ИИ



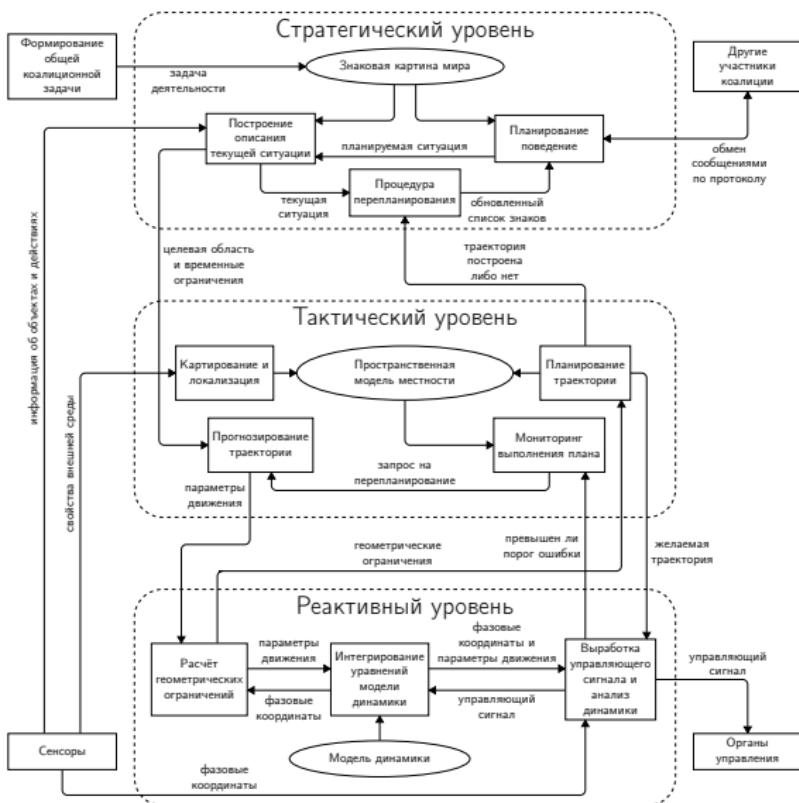
Некоторые результаты в России

- Разработан комплекс моделей поддержки принятия решений в конфликтных ситуациях с использованием когнитивных карт (ИПУ РАН).
- Для модели летательного аппарата «Этап» разработана теория бортовых интеллектуальных систем тактического уровня (БИС-Т/У), решающих задачи оперативного целеполагания и конструирования способа достижения оперативно назначеннной цели (ГосНИИАС).
- Созданы роботы, архитектура системы управления которых включает эмоциональную компоненту (НИЦ «Курчатовский институт»).
- Разработаны технологии поддержки интеллектуальных роботов-манипуляторов, способных автоматически принимать решения непосредственно во время работы (ИПМ РАН).

Некоторые результаты в ФИЦ ИУ РАН

- Создана семантическая поисковая машина нового поколения EXACTUS. Машина работает с запросами на естественном языке.
- Неоднократно занимала первые места по релевантности поиска на соревнованиях поисковых машин www.exactus.ru.
- Создана система прогнозирования социального стресса на основе анализа социальных медиа.
- Созданы системы
 - EXACTUS EXPERT — для семантического поиска и анализа качества научных публикаций,
 - EXACTUS PATENT — для семантического поиска и анализа патентной информации.
 - EXACTUS LIKE для обнаружения близких текстов и вычисления степени семантической близости;
 - TEXT Appliance — информационно-аналитическая система анализа неструктурированной информации.

Архитектура STRL

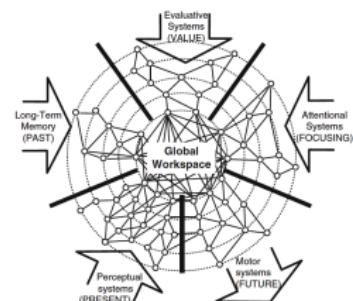
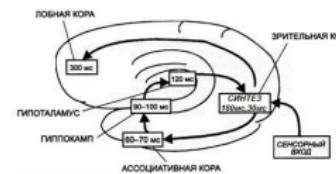
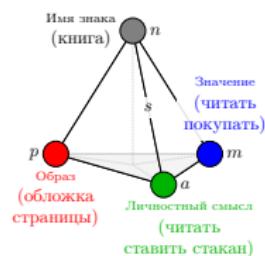


Картина мира субъекта деятельности

Картина мира субъекта деятельности - это представления субъекта о внешней среде, о своих собственных характеристиках, целях, мотивах, о других субъектах и операции (произвольные и непроизвольные), осуществляемые на основе этих представлений.

Элементом картины мира является знак:

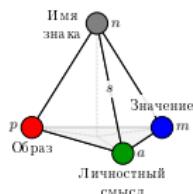
- в смысле культурно-исторического подхода Выготского-Лурии,
- выполняющий функции в соответствии с теорией деятельности Леонтьева.



В пользу существования такой структуры свидетельствуют:

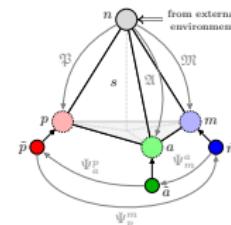
- нейрофизиологические данные (Эдельман, Иваницкий, Маунткастл и др.),
- другие психологические теории (например, трехкомпонентная модель Станович).

Три образующих элемента картины мира



Представляемая сущность описывается тремя причинно-следственными (каузальными) структурами:

- **структура образа** - представление взаимосвязи внешних сигналов и внутренних характеристик субъекта (агента) - сенсо-моторное представление,
 - **структура значения** - обобщенное знание о соотношениях во внешнем мире, согласованное в некоторой группе субъектов (агентов),
 - **структура личностного смысла** - ситуационная потребностно-мотивационная интерпретация знаний о соотношениях во внешней среде (значение для себя).



Процесс обучения —
образование новых знаков как
неподвижной точки операторов
замыкания

$$\Psi^m_{\mu} \Psi^a_m \Psi^P_a.$$

Реализация когнитивных функций — актуализация (активация) имеющихся знаков и формирование новых «сituационных» знаков — «протознаков» без конвенционального имени.

Спасибо за внимание!

МФТИ

panov.ai@mipt.ru