

Искусственный интеллект

Академична година 2023/2024

проф. Тодор Ганчев



Консултации



проф. Тодор Ганчев

tganchev@tu-varna.bg

104 ТВ



Целта на днешната лекция е ...

... да осмислим (не)интуитивното понятие AI...

... да се мотивираме -- за да продължим да мислим, да учим и да създаваме бъдещето на AI

(AI е предизвикателството на живота,
... т.к. AI е много повече от поредната нова технология)

The history of Artificial Intelligence is a history of fantasies, possibilities, demonstrations, and promise...

– Bruce G. Buchanan

“A (Very) Brief History of Artificial Intelligence”, AI Magazine, 26(4), 2005

“A (Very) Brief History of Artificial Intelligence”

Омир 750 BC → трикраки автомати, слуги на Hephaestus

...

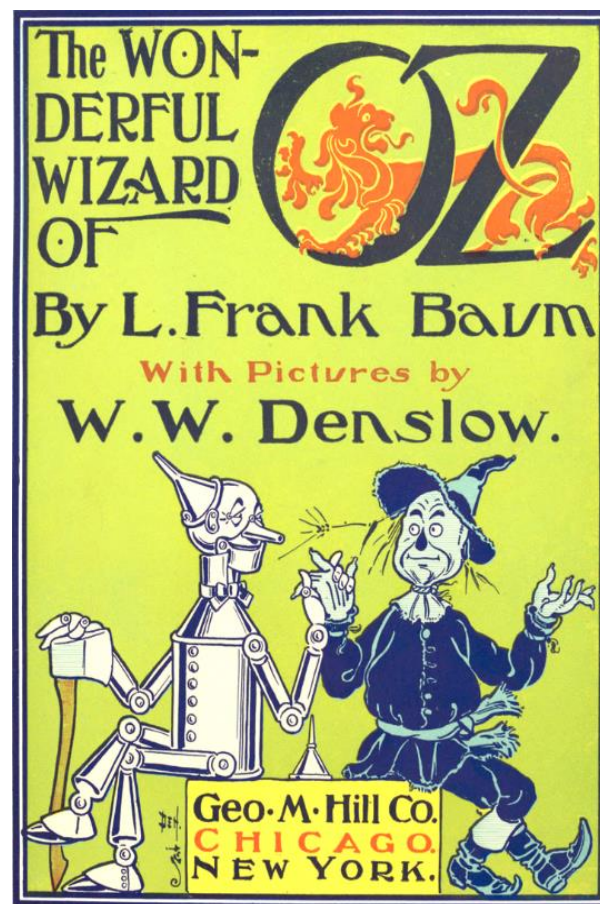
Lyman Frank Baum → механичният човек, герой в поредицата от 14 книги за магьосника от Оз, от която първата е “The Wonderful Wizard of Oz” (1900)

Тенекиеният човек е описан като:

“Изключително отзивчив, мислещ, перфектно говорещ механичен човек...

Мисли, говори, действа и прави всичко, освен да живее.”

Какво всъщност означава да е жив?



1 9 0 0

Размишления относно AI

Isaac Asimov – XX век

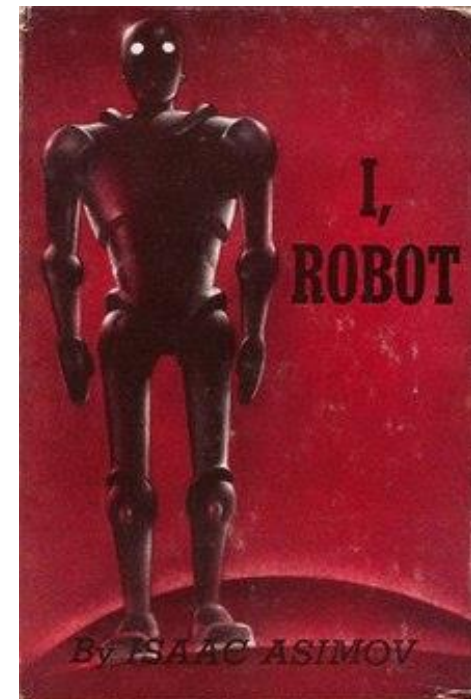
- „I, Robot“ (1950) – през 2035, роботи-хуманоиди служат в ежедневието на човечеството, което е защитено от трите закона на роботиката:

First Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.

Second Law: A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.

Third Law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

- „Foundation“, „Foundation and Empire“, and „Second Foundation“ (1951-1953) – Трилогия е изградена върху концепцията за разпада на Галактическата Империя, основополагащите принципи за запазване на цивилизацията, психо-историята (дисциплина, която използва статистически данни и модели, за да предскаже бъдещето на големи групи хора) и др. По-късно допълва трилогията с още 4 книги...
- „The Bicentennial Man“ (1976), (дом. помощник в 2020 г.)
- „The Positronic Man“ (1992)... (помощник за учени в 2022 г.)





AI и някои важни проблеми които трябва да се вземат в предвид

- AI все още не е достигнал „singularity point“ (към 2045 г. според Ray Kurzweil).
- Нерешените проблеми, са свързани главно със сложната същност на Етиката, Морала, Социалните стереотипи, Социалната приемливост на тази технология
- Някой скрити „усложнения“ поради които AI може да се счита за „неприемлив“...

=> Човечеството понастоящем не си дава сметка за моралните, социални и законови последствия от създаването на AI, ... *и какво всъщност следва след достигане момента на сингулярност ...*

AI & човешкото общество

Elon Musk за AI:

- ⇒ “The biggest mistake that I see AI Researchers making is assuming they are intelligent.
- ⇒ They are not compared to AI...
- ⇒ ...and so, a lot of them cannot imagine something smarter than themselves, but AI will be vastly smarter, vastly.”
- ⇒ “*So what do you do in situation like that?*” -- „Let’s hope AI will be nice with us (humans) ...“

Защо младите хора трябва да ги е грижа за AI ?

⇒ „...if you cannot beat AI, join AI...” 😊 -- Elon Musk

⇒ **Създаване на AI в помощ на човека** (Humanistic AI) ?– технологии които да помагат на човека да подобри когнитивните си способности, възможността да се справя с проблеми, да подсилва ума или тялото на човека, особено в случаите когато има установени дефицити

⇒ **Възможни са и друг вид „бъдеще“ за AI:**

- в едно от тези, AI се използва за водене на *войни*,
- в друго, като инструмент за *контрол* или за „въдворяване на ред“, *тирания* ...
- или за развитие на AI към човешките ценности (*тук възникват проблеми свързани с Етика, Морал, Социални норми*) ...

Основни теми в настоящия курс

Въведение. Основни понятия.
Исторически сведения. Философия
и етика на AI.

Интелигентни агенти. Рационален
агент използващ модел на околния
свят. Агенти изпълняващи зададен
цел. Агенти оптимизиращи целева
функция. Самообучаващи се
агенти.

Методи за машинно обучение.
Неврони без обратни връзки.
Класификация с невронни мрежи
без обратна връзка (FFNN, MLP NN,
PNN)

Класификация с невронни мрежи.
Неврони с обратни връзки.
Рекурентни и локално-рекурентни
невронни мрежи. (TDNN, DRNN,
LRPNN, GLRPNN)

Статистически методи за машинно
обучение. Смес от Гаусови
функции (GMM). Класификация с
опорни вектори (SVM).

Конволюционни невронни мрежи
(CNN). Архитектури на дълбоки
невронни мрежи (DNN).

Еволюционни методи за обучение
на модели. Диференциална
еволюция (DE). Оптимизация с
рояк от частици (PSO)

Ресурси необходими за
реализацията на системи с
изкуствен интелект. Апаратни
средства.

Съвременни средства за
реализация системи с изкуствен
интелект.

Оценяване за спец. КСТ – 3к

Редовно обучение:

Краен_брой_точки = Точки_{Сем.Контрол} + Точки_{Изпит}

$0 \leq \text{Точки}_{\text{Сем.Контрол}}, \text{Точки}_{\text{Изпит}} \leq 50,$

- Точки от Семестриален контрол (общо до 50 точки):
 - реферат/презентация/демо
 - активно участие в упражненията
- Точки Изпит-тест (общо до 50 точки):



Краен_брой_точки > [50...100] точки

Реферат

Съгласувайте с асистента по AI на лаб. упр. !!!

Примерни теми:

- Автоматичен превод от един език на друг.
- Автоматично генериране на програмен код по задание.
- Чат-ботове.
- Интелигентни асистенти – Siri, Alexa ...
- Анализ на видео – футбол, охрана, автомобили.
- Интелигентна навигация на робот, дрон...
- Финансови приложения на AI ...
- Приложения в роботиката и VR ...
- Компютърни игри с AI ...

Приложения на AI

Обработка на естествена реч

- Разпознаване, Интерпретация, Чат-ботове
- Синтез на реч, Промяна на идентичност (напр. voice morphing)

Класификация на обекти, събития, поведения...

- Емоционално състояние (от реч, ЕКГ, ЕЕГ, и др.)
- Опасни поведения (вандализъм, уличен бой...)
- Животински видове (насекоми, птици...)

Регресия и предсказване

- Ръст на човек по гласът му
- Обем на тялото

Търсене на информация, събитие, знания

Вземане на информирано/отговорно решение

Приложения на AI

Разпознаване на обекти, поведения и събития

- Букви/символи, Авторство,
- Реч/диктор/диалект,
- Животински видове от аудио/видео записи,
- Лица, пръстови отпечатащи, поведенчески биометрични показатели
- Типове дейности, Опасни поведения, Намерения,
- Събития, новост, отклонения от очакванията
- Емоционални състояния, диагностика на заболявания...

Машинен превод на писмени документи и реч

Управления и навигация в роботиката

Стратегическо планиране

Експертни системи -- подпомагане вземането на решения

Диагностични системи

Компютърни игри, Ботове, Изкуствен живот

Произведения на изкуството и творчество от AI с-ми

Военно дело, Икономика и финанси

Неврология, Психология, Лингвистика ...

Препоръчвана литература

- С. Стоянов, Т. Глушкова, Й. Тодоров, Изкуствен интелект: Решаване на проблеми чрез търсене. Издателство Изкуства. 2019 г. ISBN: 9786197243871
- С. Стоянов, Т. Глушкова, М. Папанчева, Изкуствен интелект: Представяне на знанията чрез логика. Логическо програмиране, Издателство Изкуства. 2021 г. ISBN: 9786197243970
- М. Тодорова, Изкуственият интелект: Кратка история на развитие и етични аспекти на темата, Издателство Изток-Запад. 2019 г. ISBN: 9786190105602



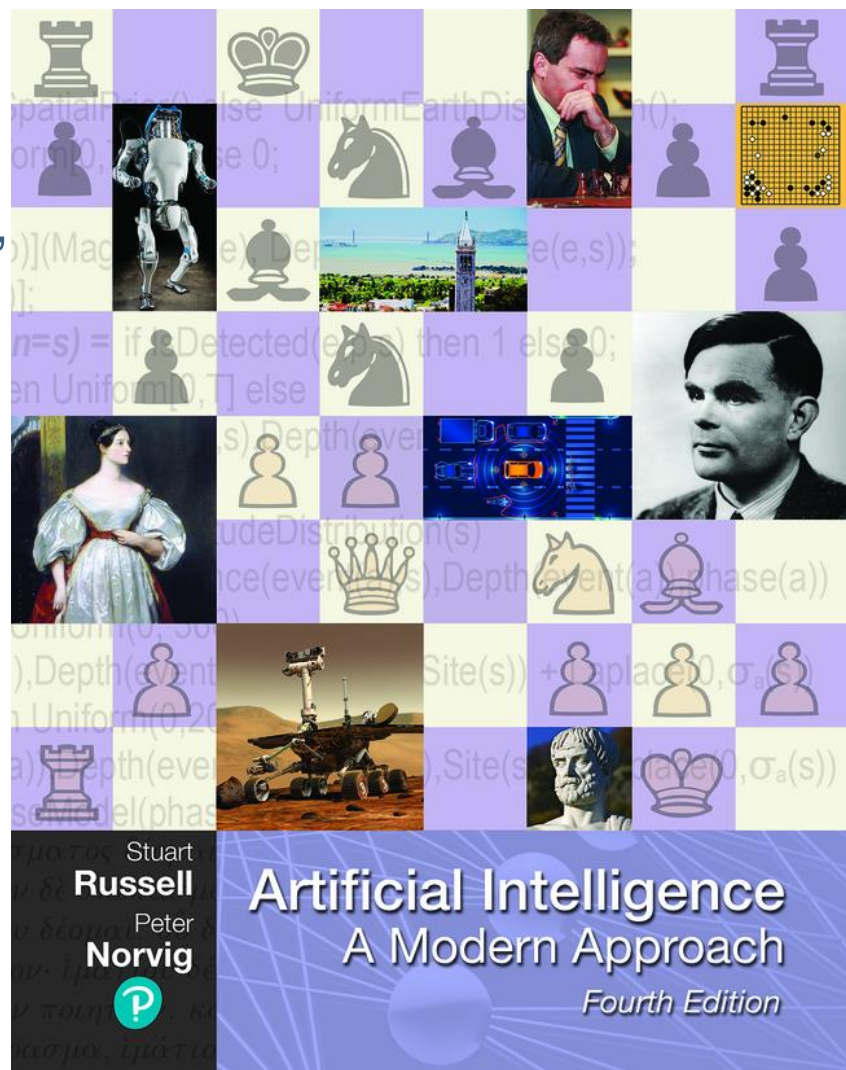


Препоръчвана литература на Английски език

- **S. Russell & P. Norvig, Artificial intelligence: A Modern Approach, 4th Ed. Prentice-Hall, 2021, ISBN: 9780134610993**
- D. Poole & A. Mackworth, Artificial Intelligence: Foundations of Computational Agents, 2nd edition, Cambridge University Press, 2017. ISBN: 9781107195394
- M.T. Jones, Artificial Intelligence: A Systems Approach, Infinity Science Press LLC, Hingham, Massachusetts, 2008.
- M. Negnevitsky, Artificial intelligence: A guide to intelligent systems, 3rd Edition, Addison-Wesley, 2011.
- T. Munakata, Fundamentals of the New Artificial Intelligence: Neural, Evolutionary, Fuzzy and More, 2nd Edition, Springer, 2008
- M. Minsky, The Emotion Machine. ISBN 0-7432-7663-9. Simon & Schuster, 2007.

Препоръчвана литература

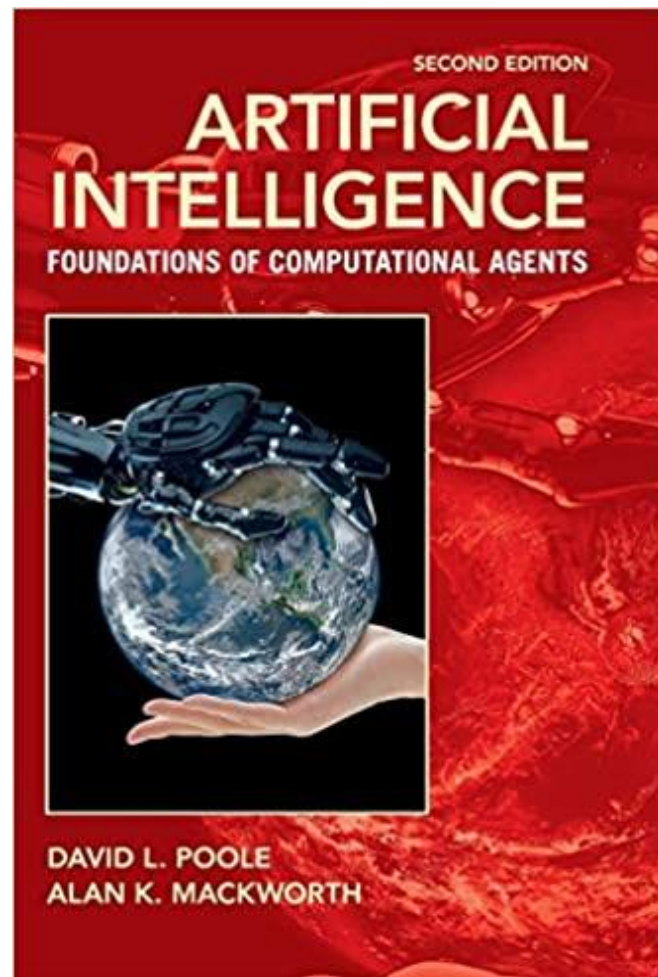
S. Russell & P. Norvig,
Artificial Intelligence: A Modern Approach,
4th Ed.,
Prentice-Hall, 2021.
ISBN: 9-780-1346-1099-3



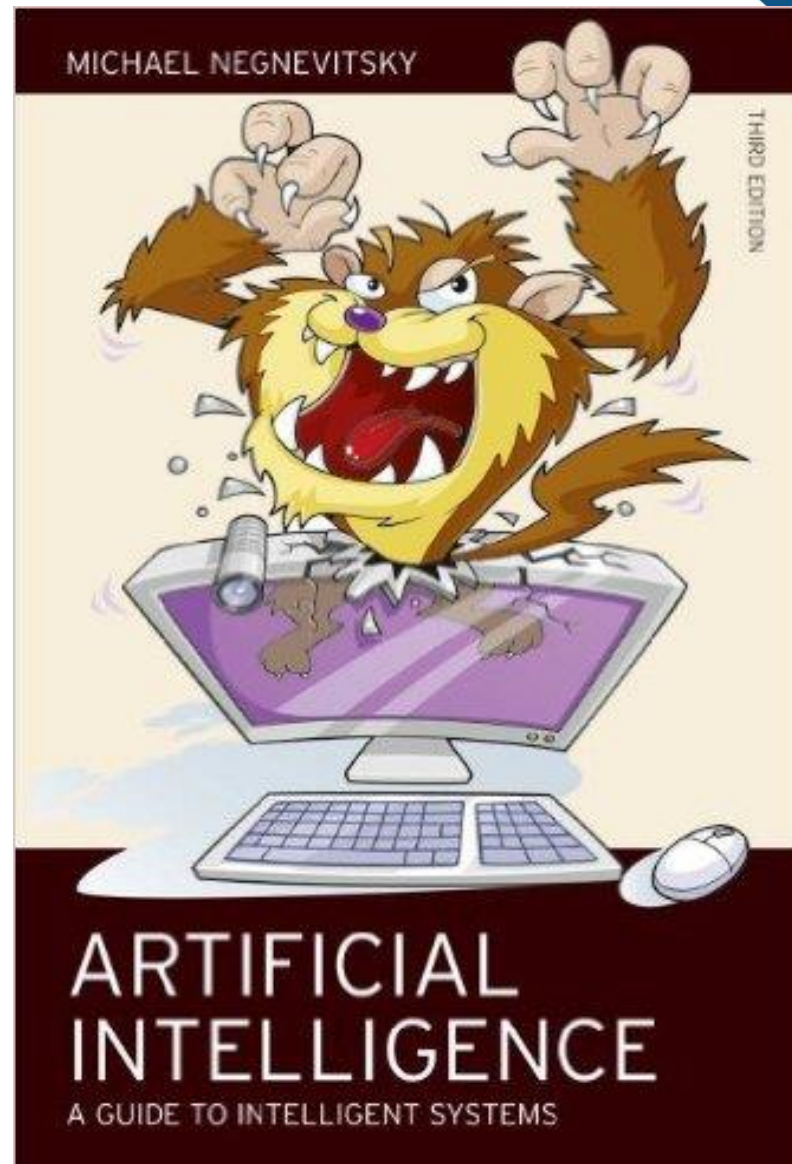
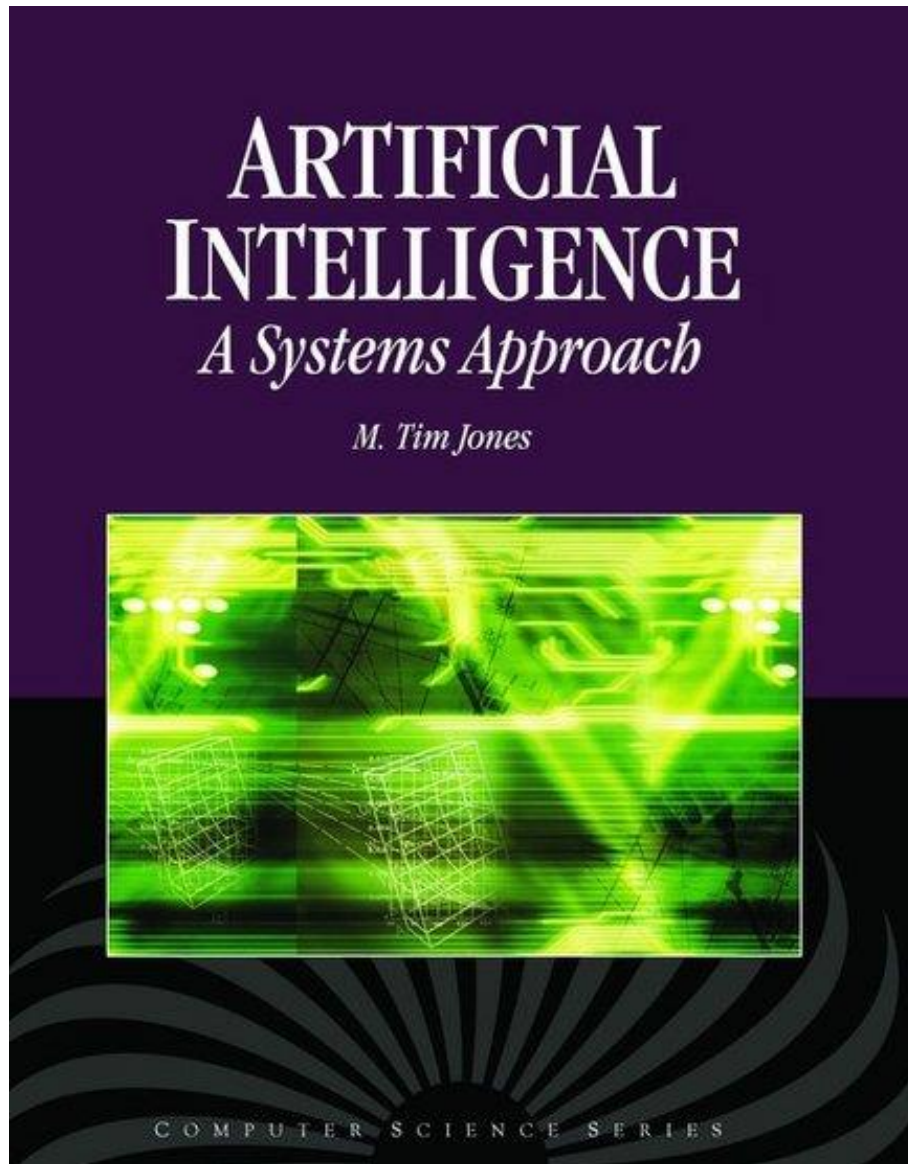
Препоръчвана литература

D. Poole & A. Mackworth, Artificial Intelligence: Foundations of Computational Agents, 2nd Ed., Cambridge University Press, 2017.

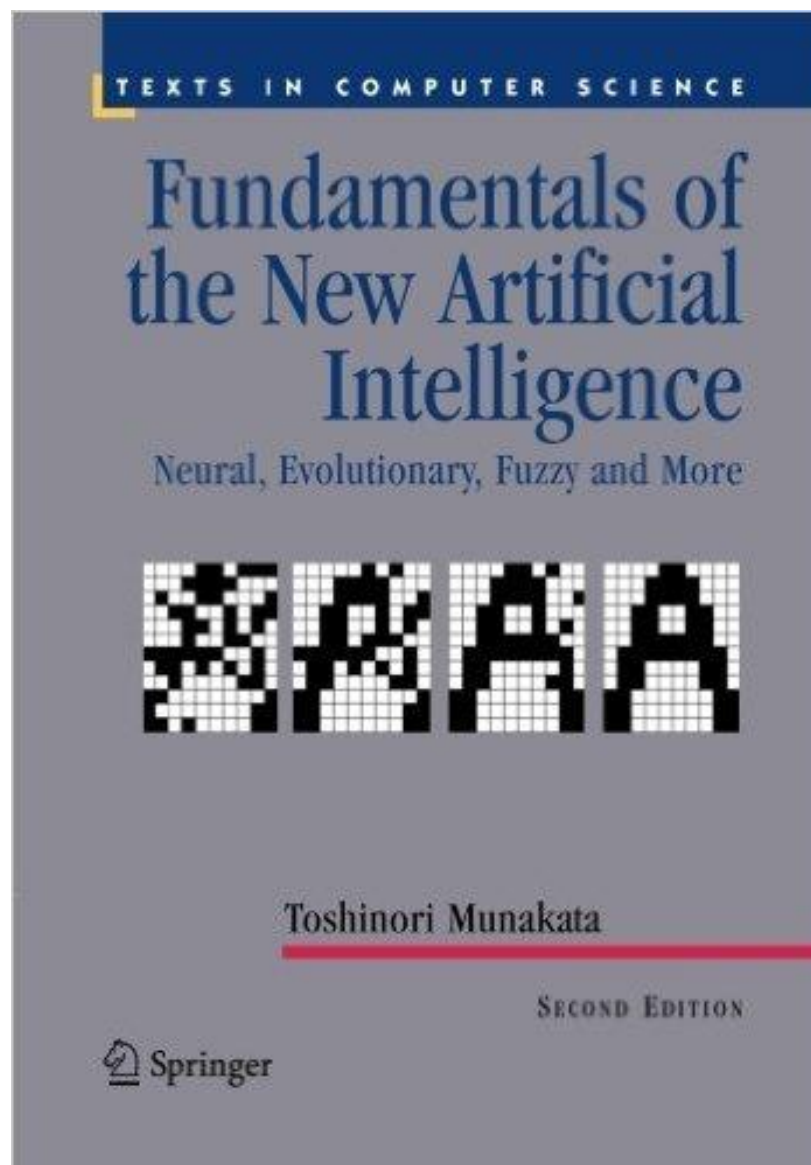
ISBN: 9-781-1071-9539-4



Препоръчвана литература

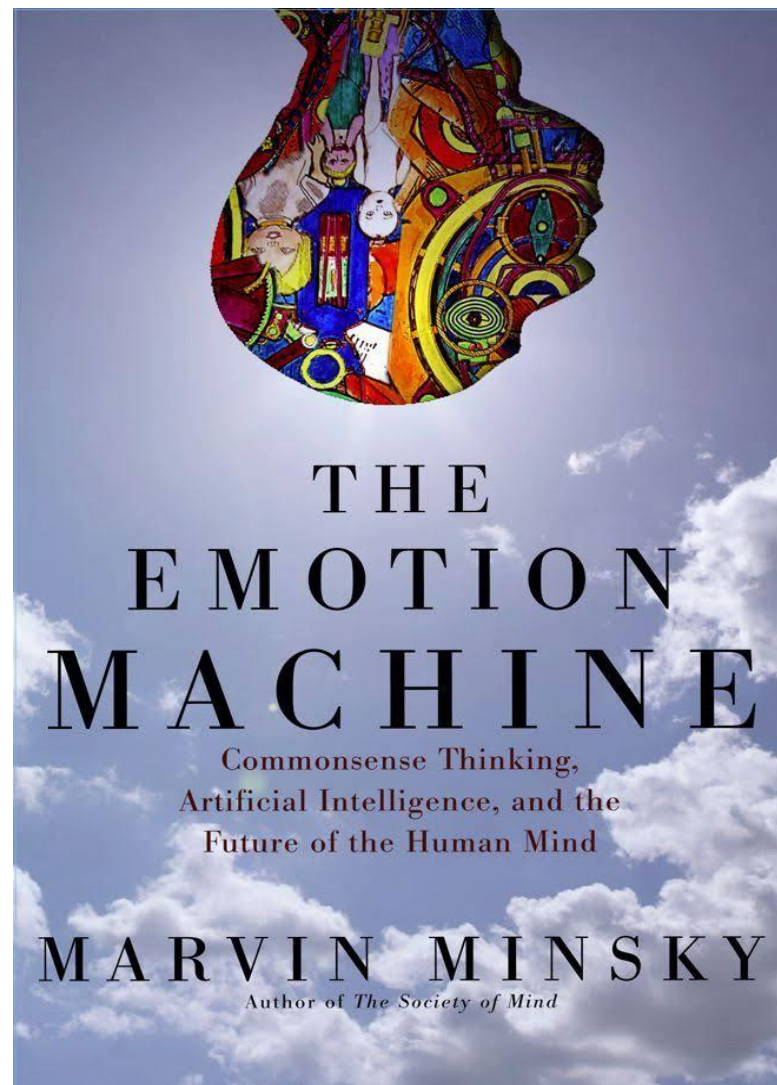


Препоръчвана литература



Препоръчвана литература

M. Minsky, The Emotion Machine.
Simon & Schuster, 2007.
ISBN 0-7432-7663-9



Изкуствен интелект

Тема 1:

Въведение в дисциплината





Тема 1: Въведение в дисциплината

Основни понятия

Гледни точки към науката „Изкуствен интелект“

История на AI

Достижения на AI

Текущи възможности на AI

Основни понятия

- Интелект ?
 - Интелигентност ???
 - Естествена интелигентност ???
 - Изкуствена интелигентност ???
 - Машинна интелигентност ???
 - Други видове интелигентност ?
 - Извънземна ????
 - Животинска ??
 - Универсална интелигентност?
 - Изкуствен интелект ???
- ...“**машини които могат да мислят, да се учат, и да създават**”...

Очаквания

Автономни системи които:

- се самообучават нови задачи/проблеми
- Играят шах, GO, решават кръстословици
- Доказват теореми, извличат знание, откриват действието на нови молекули и оценяват тяхната приложимост
- Пишат поезия, рисуват картини, създават музика
- Пилотират автомобили, кораби, самолети, космически апарати...
- Диагностицират заболявания
- Планират дейности, Оценяват рискове
- Управляват финансови ресурси...

... или която и да е друга интелектуална дейност ...

Гледни точки към AI

Човешко мислене

Рационално мислене

Човешки действия

Рационални действия

Системата е рационална ако взема „най-доброто решение“ и (извършва съответните действия) ...
...при наличната до момента информация.

Гледни точки към AI

Thinking Humanly

“The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense.” (Haugeland, 1985)

“[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . .” (Bellman, 1978)

Acting Humanly

“The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990)

“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)

Thinking Rationally

“The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985)

“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)

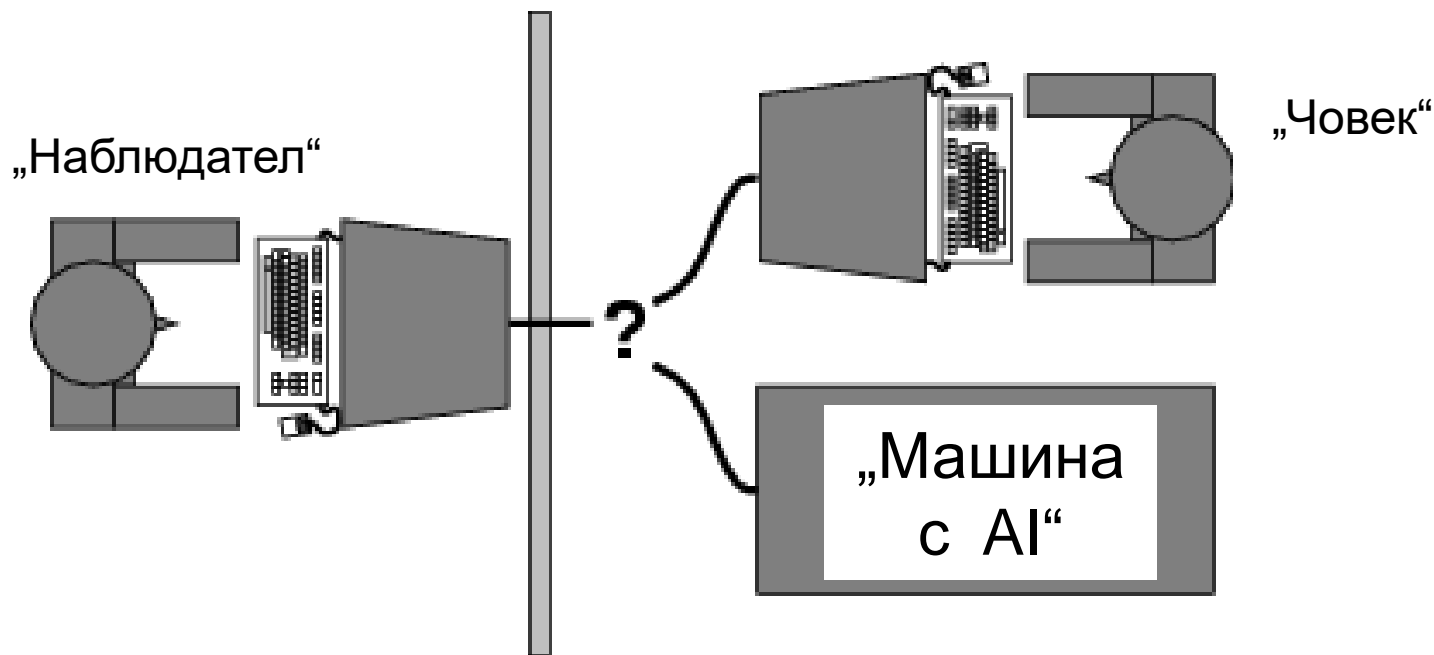
Acting Rationally

“Computational Intelligence is the study of the design of intelligent agents.” (Poole *et al.*, 1998)

“AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)

AI и човешкият начин на действие

- Тест на Turing (1950) "Computing machinery and intelligence,"
- "Могат ли машините да мислят?" → „Могат ли машините да имат интелигентно поведение ?,"
- Имитация на човешки поведения (The Imitation Game)



Машината притежава интелигентност ако наблюдателят не може да прецени дали комуникира с човек или от машина.

AI и човешкият начин на действие



Allan Turing:

- Прогнозира, че до 2000 г. машините може да имат 30 % шанс да заблудят необучен човек в рамките на 5 минути
- Предугажда всички основни аргументи против AI за следващите 50 години
- Описва всички основни компоненти на AI :
 - Система за представяне и съхранение на знание (памет)
 - Система за автоматично получаване на заключения (мислене),
 - Система за разбиране на естествена реч (комуникация),
 - Система за машинно обучение (учене и адаптация).

AI и човешкото мислене: „cognitive modeling“

- „Когнитивната революция“, през 60те год. на XX век: психология на обработката на информация
- Изисква научни теории за вътрешната работа на мозъка
- -- Как да се валидира това? Изисква се
 - 1) Предсказване и тестване на поведението на хора (top-down) или
 - 2) Директна идентификация от неврологични сигнали (bottom-up)

Тези два метода (Когнитивна наука -- **Cognitive Science** и Когнитивна неврология -- **Cognitive Neuroscience**) понастоящем не са част от науката за AI, според днешните формулировки

Рационално мислене: "laws of thought"

Аристотел разсъждава какво представляват коректните (правилни) аргументи и процеса на мислене?

Античните гръцки учени/философи дефинират различни форми на логиката:

система за означаване и правила за извличане на мисли, независимо от това дали се има в предвид идеята за механизация или не на процеса.

Съществува директна връзка между модерната наука AI и математиката и философията.

Проблеми:

1. Не всички интелигентни поведения се предават/проявяват посредством логически съждения.
2. Каква е целта и смисъла на мисленето?
3. Какви мисли трябва да имам?

Рационални действия: „rational agent“

- Рационално поведение: doing the right thing!!! Правим това което е правилно...

- Кое е правилно?

Правилното е това което се очаква да максимизира постигането на заложена цел, при наличната към момента информация.

- За целта не е задължително да присъства „мислене“,

но ако има такова мисленето трябва да е в услуга на рационалното действие ...

- Рационалните агенти се очаква да имат рационално поведение.

Рационален агент

- Агент → дефиниция?
- Рационален агент ?
- Абстрактно записано, агентът е функция която превръща последователност от наблюдения и възприятия в действия:

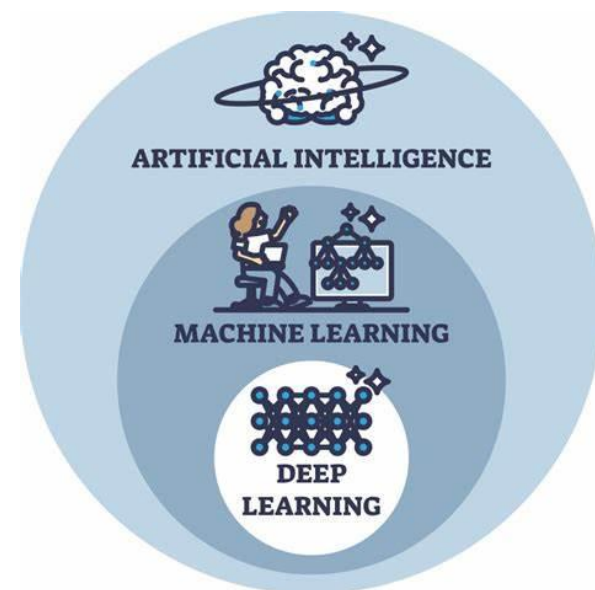
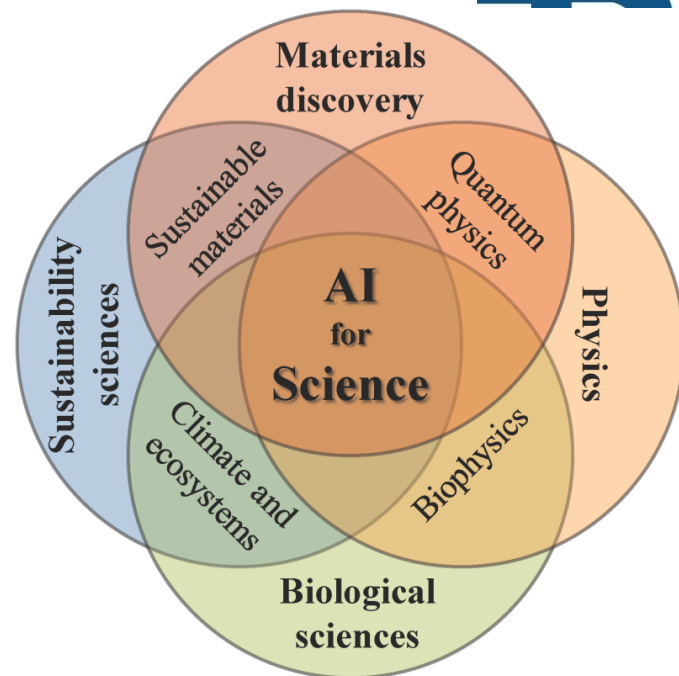
$$[f: \mathcal{P}^* \rightarrow \mathcal{A}]$$

За всяка среда и задача, ние търсим агент (или категория от агенти) с най-добри характеристики

- **Ограничение:** Изчислителната тежест прави недостижимо създаването на идеален агент
 - Създай най-добрата програма в рамките на ресурсите на конкретния хардуер

Предистория на AI

- **Философия** - Логика, методи за обосновка и вземане на решения, умът като физична система в основата на обучението, език, рационалност.
- **Математика** - Математическо представяне на алгоритми за доказване, изчисления, вероятност, възможност или невъзможност за обосноваване на решение, податливост или неподатливост за формализиране на проблема.
- **Икономика** - полезност, теория за вземане на решения.
- **Неврология** - физическа основа за умствена дейност.
- **Психология** - феномени на възприятие и двигателен контрол, експериментални техники.
- **Компютърно инженерство** - създаване на свръх-бързи компютри.
- **Теория на управлението** - проектиране на системи, които максимизират обективна функция във времето.
- **Лингвистика** - начини за представяне на знание, граматика.



Събития от историята на AI

1940-1950s – Зараждане на AI

1943 McCulloch & Pitts: Модел на мозък от логически елементи

1950 А. Тюринг: Компютри и интелигентност -- "Computing Machinery and Intelligence"

1950-1970s – Първоначален ентусиазъм

1950s Ранни AI програми: Checkers program, Logic Theorist, Geometry Engine

1956 Семинар в Дортмунд: терминът "Artificial Intelligence"

1965 Цялостен алгоритъм за логически разсъждения (Робинсън)

1970-1990s – подходи основаващи се на знание (Knowledge-based)

1969—79 Ранни разработки по системи базирани на знание (knowledge-based systems)

1980-- Експертните системи стават индустрия, AI е предимно статистическо моделиране

Събития от историята на AI

1990s- наши дни – научни методи, AI става наука

1987-- Вероятностен подход, справяне с несигурност, AI става наука

1995-- Появяват се интелигентните агенти

XXI век - AI технологии в ежедневието

2002-- прахосмукачката iRobot Roomba

2011-- Watson, Siri

2014-- Софтуер (почти) удовлетворява тест на Тюринг –> 1/3 съдиите са убедени, че разговарят с 13 годишно момче Eugene Goostman. През същата година се появява Alexa.



Някой достижения на AI

- 1997 – IBM “Deep Blue” побеждава Гари Каспаров, тогавашния световен шампион по шах
- 1999 – Proverb решава кръстословици по-добре от повечето хора
- 2011 – IBM Watson побеждава на Jeopardy!, побеждавайки легендарните шампиони Brad Rutter и Jen Jennings.
- 2015 – DeepMind’s AlphaGo побеждава Европейския шампион на китайската игра Go.
- Google-Waymo демонстрира кола без водач (driverless cars*)
Chryslers, Toyota, Honda, ..., Apple.

* [Society of Automotive Engineers \(SAE\)](#)

Level 0 - no automation

Level 1 - hands on/shared control

Level 2 - hands off

Level 3 - eyes off

Level 4 - mind off

Level 5 - steering wheel optional.

Някои достигновения

Европейски парламент

Изкуственият интелект е ключова технология за бъдещето, която намира все повече приложения днес. Научете какво представлява той и как променя живота ни.

Изкуственият интелект

Използване днес и в бъдеще

Няколко примера как вече използваме изкуствения интелект и какви възможности предлага той

Дигитални лични асистиенти на смартфони или компютри

Интелигентни климатични инсталации

Интернет на нещата: свързани с Мрежата прахосмукачки, хладилници, часовници...

Автономни коли

Пазаруване и рекламиране онлайн

Умно земеделие: напояване, хранене на животни, роботи, премахващи плевели

Роботи във фабриките

Търсене в интернет

Машинни преводи

Електронна сигурност

Борба с дезинформацията

Оптимизиране на продукти и на вериги за продажба

Някой достижения





Agent-AI

Analyzing information issued by people and understanding intentions and emotions in that information



Person
to
Machine

AI

Machine
to
Machine



Analyzing unconscious and unnoticeable aspects of a person's mind and body as well as understanding deep psychological, intellectual, and instinctual states in that person

Heart-Touching-AI

Ambient-AI

Analyzing and understanding just about anything in the world (objects, people, the environment) and instantaneously predicting and controlling those things



Organically connecting and cultivating multiple types of AI and optimizing the entire social system

Network-AI

AI-Някой достижения



Доказва математическо предположение (Robbins conjecture) нерешено от десетилетия

AI за военно планиране – нападение, отбрана, тилова подкрепа от товари и жива сила...

Алгоритми за планиране и насочване на космически апарати и боеприпаси.

Машинен превод на текст (Google, Microsoft, OpenAI...)

Skype – превод в реално време (voice translation)

Създава програмен код на различни езици за програмиране

Създава произведения на изкуството (поезия, проза, картини, видеа... по поръчка)

Elon Musk – AI технологии за финансовите пазари ... neuro-chip за BCI, нови HMI за управление на обекти и др.

Ресурси: сравнение човек-машина

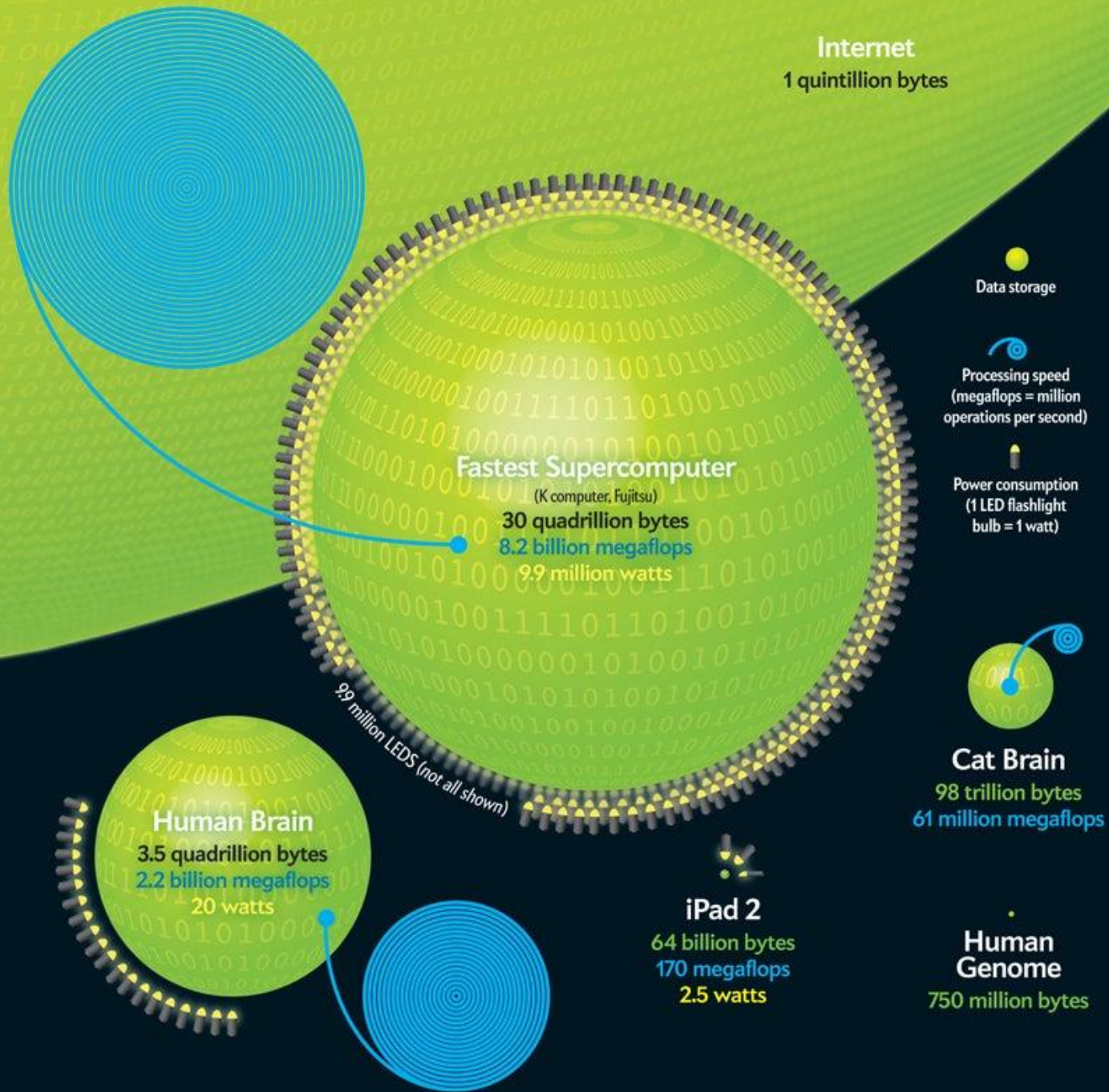


Съгласно Ръсел & Норуиг, 2021:

	Supercomputer	Personal Computer	Human Brain
Computational units	10^6 GPUs + CPUs	8 CPU cores	10^6 columns
	10^{15} transistors	10^{10} transistors	10^{11} neurons
Storage units	10^{16} bytes RAM	10^{10} bytes RAM	10^{11} neurons
	10^{17} bytes disk	10^{12} bytes disk	10^{14} synapses
Cycle time	10^{-9} sec	10^{-9} sec	10^{-3} sec
Operations/sec	10^{18}	10^{10}	10^{17}

Figure 1.2 A crude comparison of a leading supercomputer, Summit (?); a typical personal computer of 2019; and the human brain. Human brain power has not changed much in thousands of years, whereas supercomputers have improved from megaFLOPs in the 1960s to gigaFLOPs in the 1980s, teraFLOPs in the 1990s, petaFLOPs in 2008, and exaFLOPs in 2018 (1 exaFLOP = 10^{18} floating point operations per second).

Ресурси: котка човек компютър Интернет



Какво може AI днес?

...Понастоящем AI позволява да се доближим на $\frac{3}{4}$ от функционалността на човешката нервната система.



Компютрите са
по-успешни от
човека

- Входни и изходни интерфейси
- Обработка на информация
- Памет
- ...

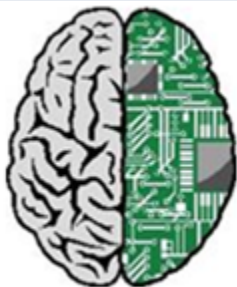
Какво може AI днес?

...Понастоящем AI позволява да се доближим на $\frac{3}{4}$ от функционалността на човешката нервната система.



Компютрите са
по-успешни от
човека

- Входни и изходни интерфейси
- Обработка на информация
- Памет
- ...



Близки
ВЪЗМОЖНОСТИ

- Сложни движения
- Зрение
- Езикови способности
- Решаване на структурирани проблеми

Какво може AI днес?

...Понастоящем AI позволява да се доближим на $\frac{3}{4}$ от функционалността на човешката нервната система.



Компютрите са по-успешни от човека

- Входни и изходни интерфейси
- Обработка на информация
- Памет
- ...



Близки възможности

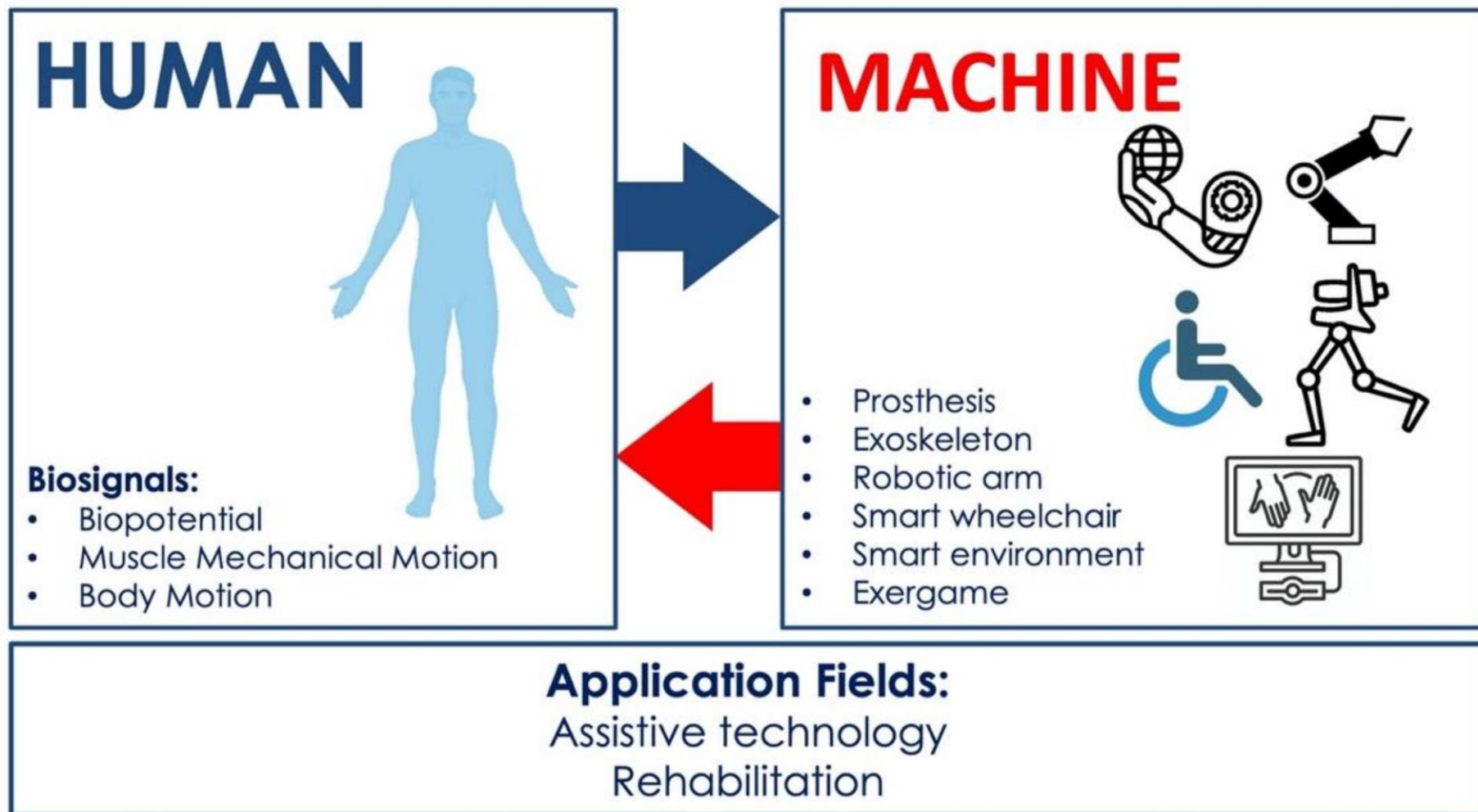
- Сложни движения
- Зрение
- Езикови способности
- Решаване на структурирани проблеми



Човешкият мозък е по-успешен

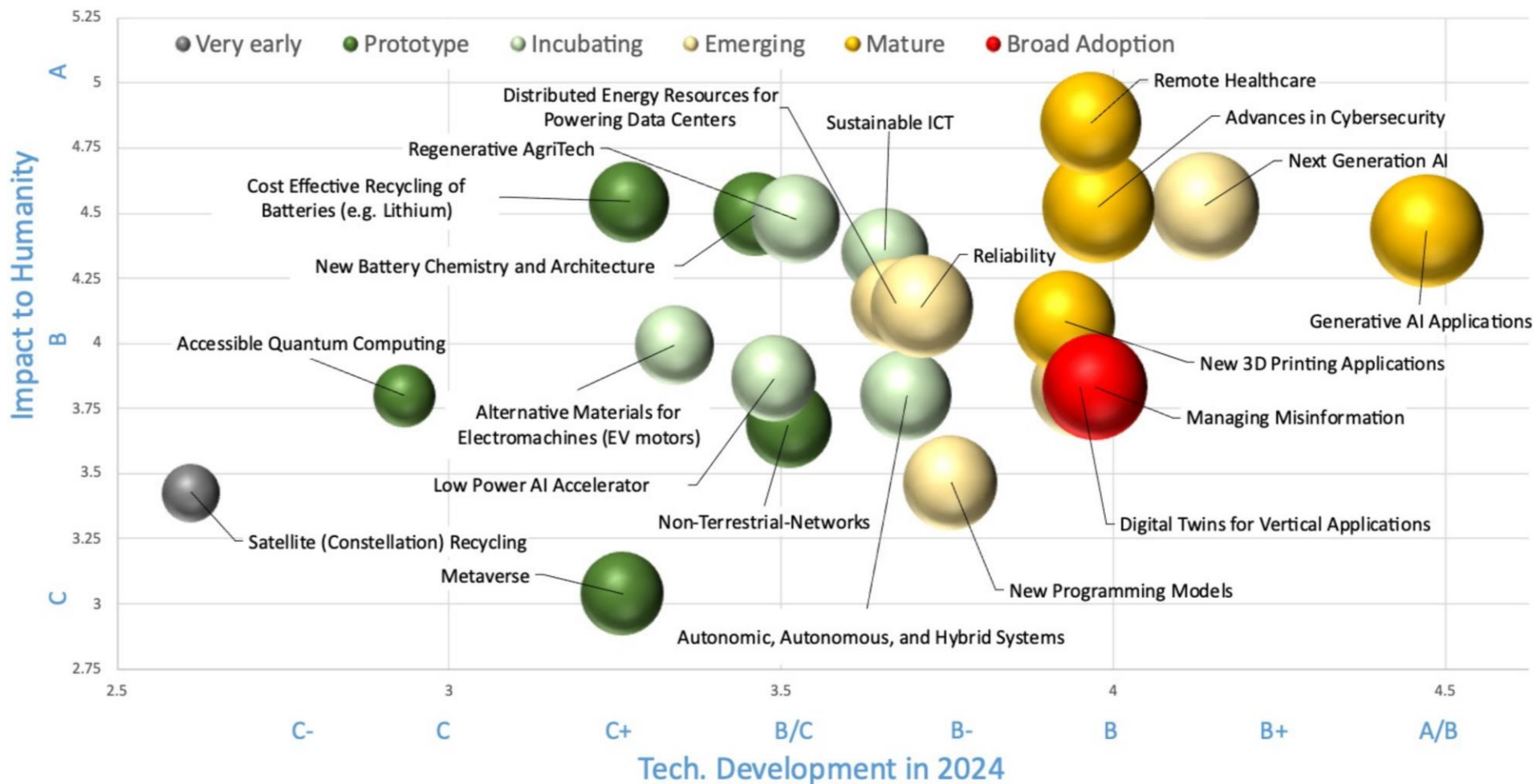
- Творчество, съзидание
- Емоции и емпатия
- Планиране и вземане на решения
- Съзнание и подсъзнание

Сливане в едно Човек-Машина???



Настоящето -> 2024 в прогнози

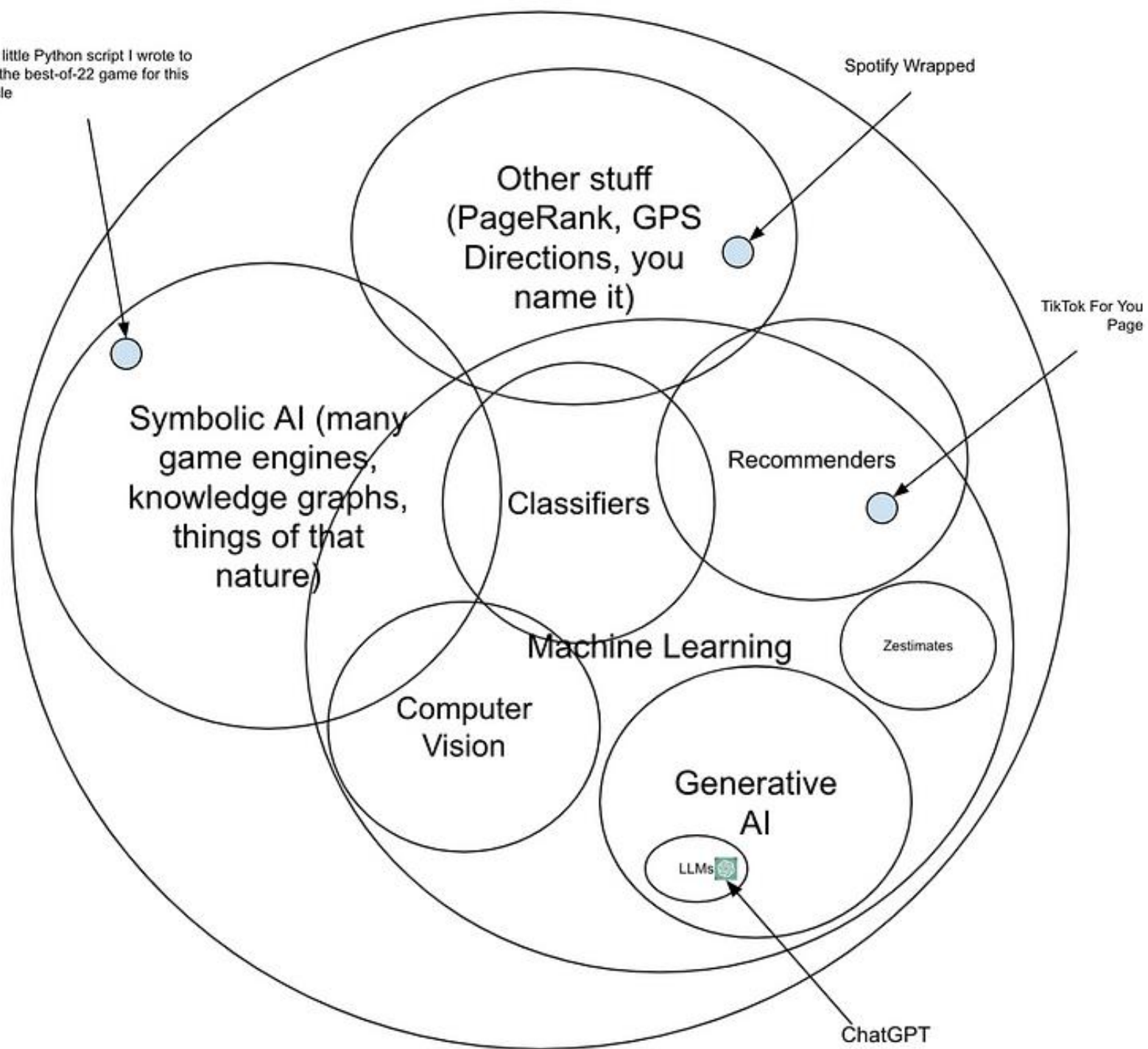
PREDICTION: Tech. Development in 2024 (x-axis) vs Impact to Humanity (y-axis)
(size of bubble proportional to relative market adoption)





AI

The little Python script I wrote to win the best-of-22 game for this article



Въпроси?

