

Tehisintellekt: sissejuhatus



- Orientatsioon
- Tehisintellekti tüübid ja valdkonnad
- Hübriidsüsteemid
- Robotid
- Filosoofia ja prognoosid
- Keelemudelite rakendamise hetkeseis

Orientatsioon

Marc Andreessen:

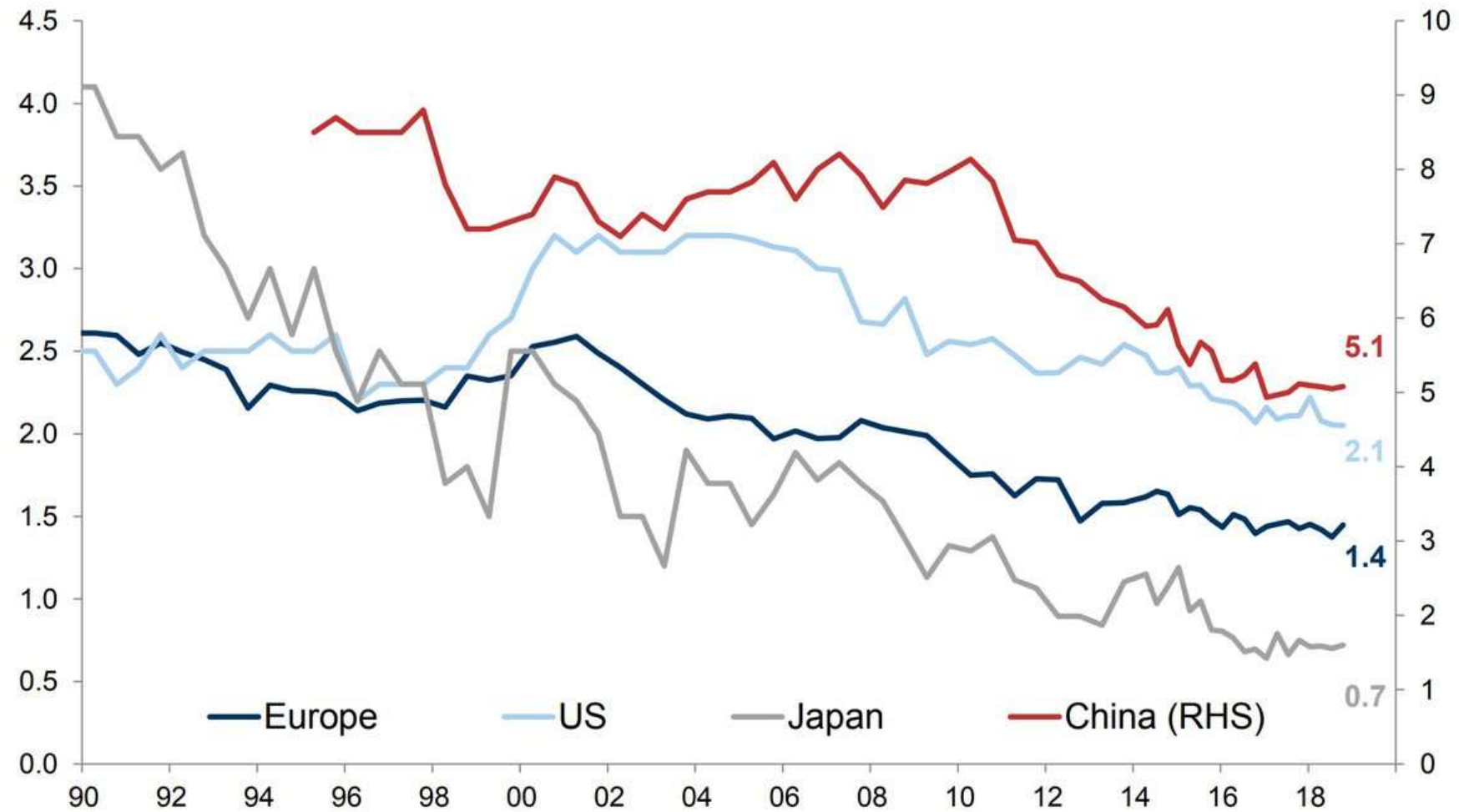
Software Is Eating the World

Paradoks:

IT ei ole sel sajandil globaalmajandust oluliselt efektiivistanud: vastupidi, GDP kasvutempo on pikaajalises langustrendis.

Exhibit 8: Long-term real global GDP growth forecast is at a historical low

Long-term (6-10y) GDP growth from Consensus Economics

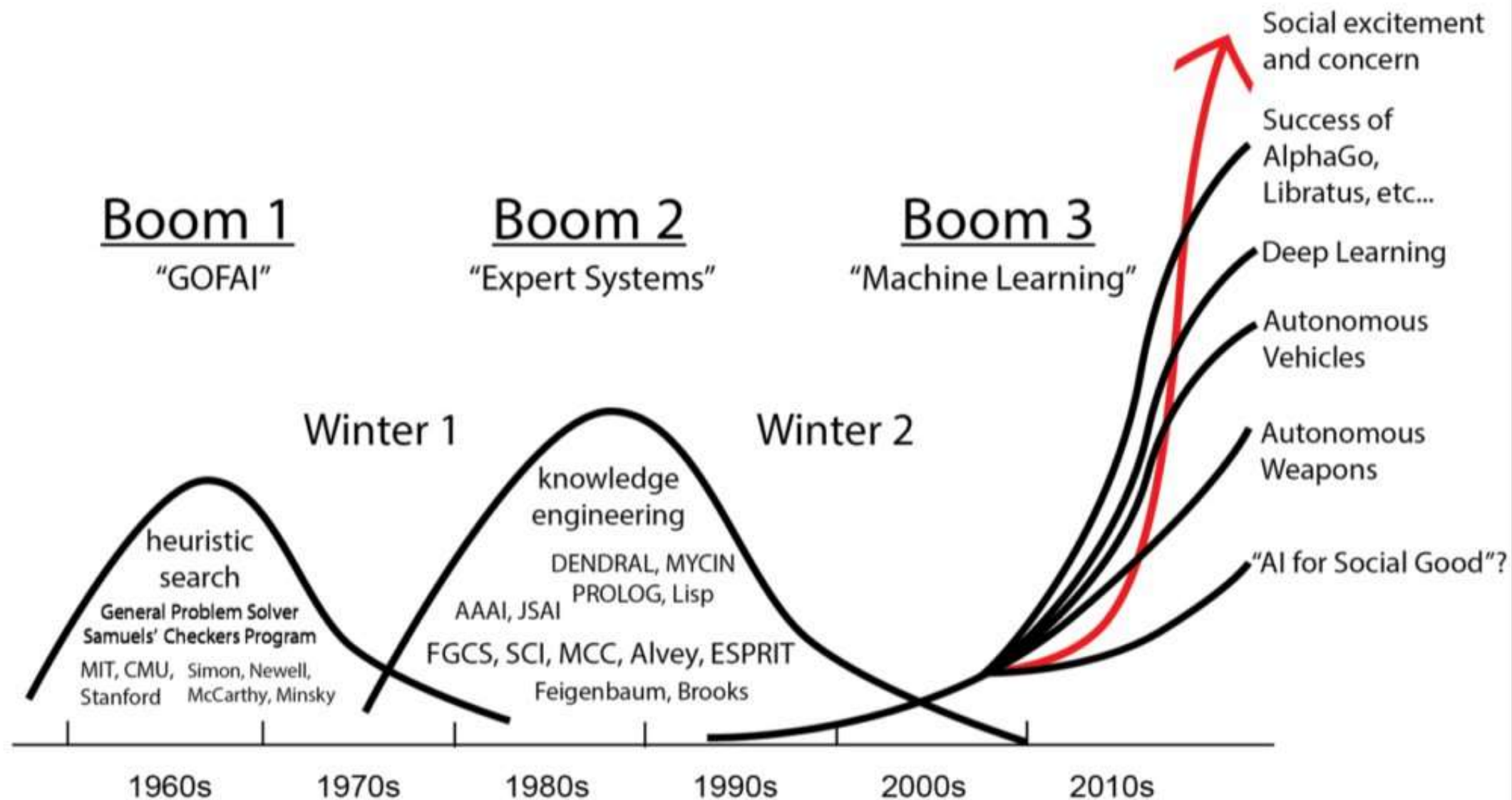


Source: Consensus Economics, Goldman Sachs Global Investment Research

IT globaalne probleem ja lootus

Tööprotsesside automatiseerimine on väga raske: tarkvara ehitamine on aeglane ja kallis.

Tehismõistuse tehnoloogiad aitavad efektiivsemalt ehitada automatiseerimis-tarkvara.

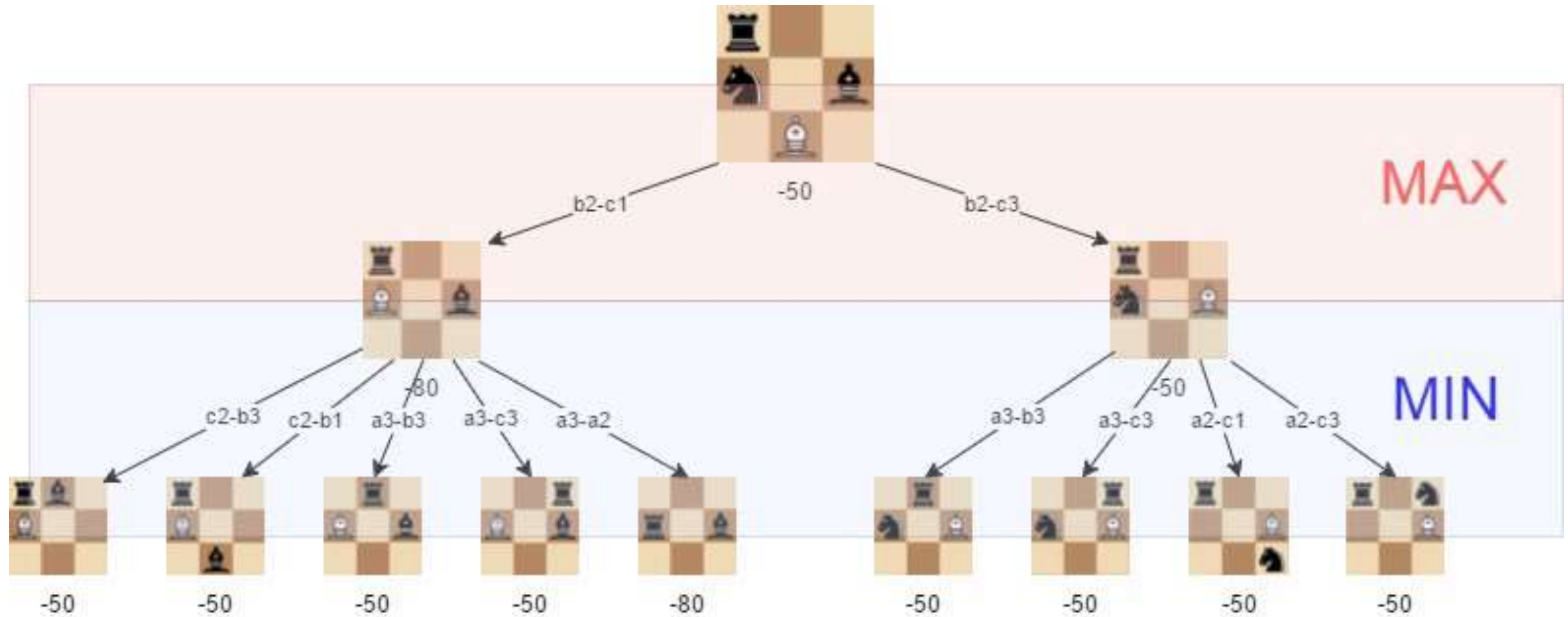


Tehisintellekti tüübid ja valdkonnad

Kaks suurt A.I. meetodite valdkonda:

- **Otsing** (sümbol-A.I.)
paljude variantide seast a la malemäng, matemaatikaõestus, teekonna optimeerimine, planeerimine jne.
- **Masinõpe** (“neuro”-A.I.)
näidetelt üldistamine, a la objektide äratundmine pildilt, deepfake, aegridadelt ennustamine, suured keelemudelid (LLM) GPT, LLaMA, Claude jne.

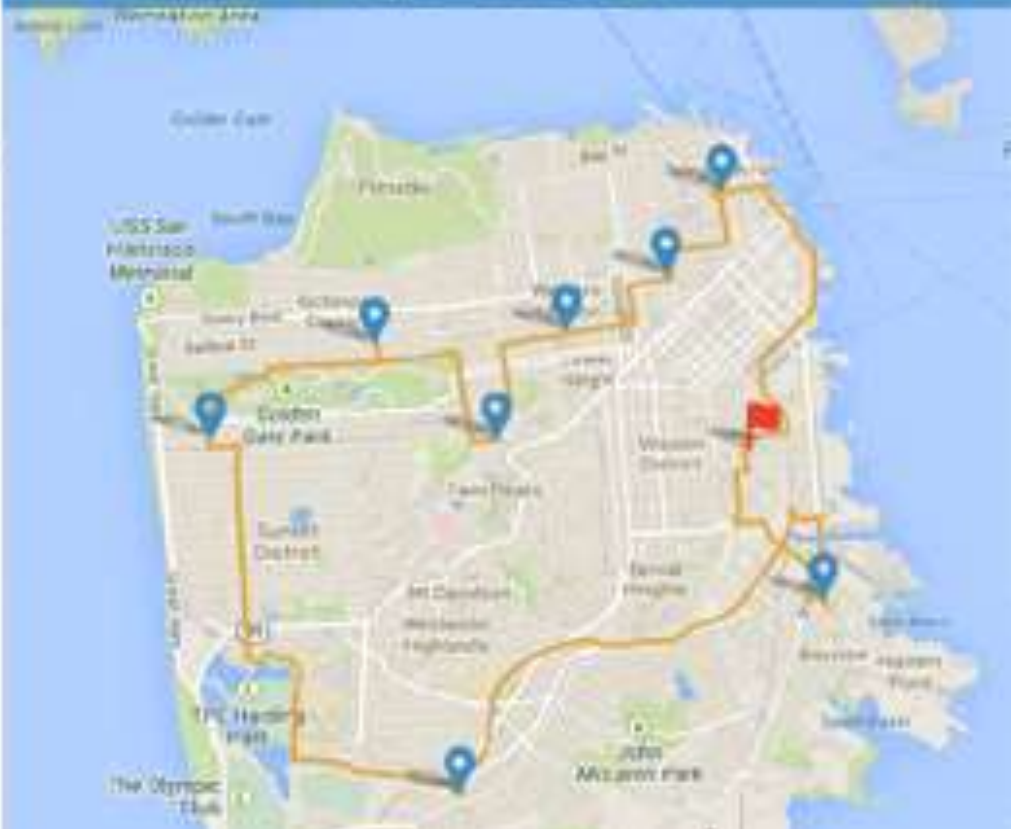
Otsing variantide hargnevast puust

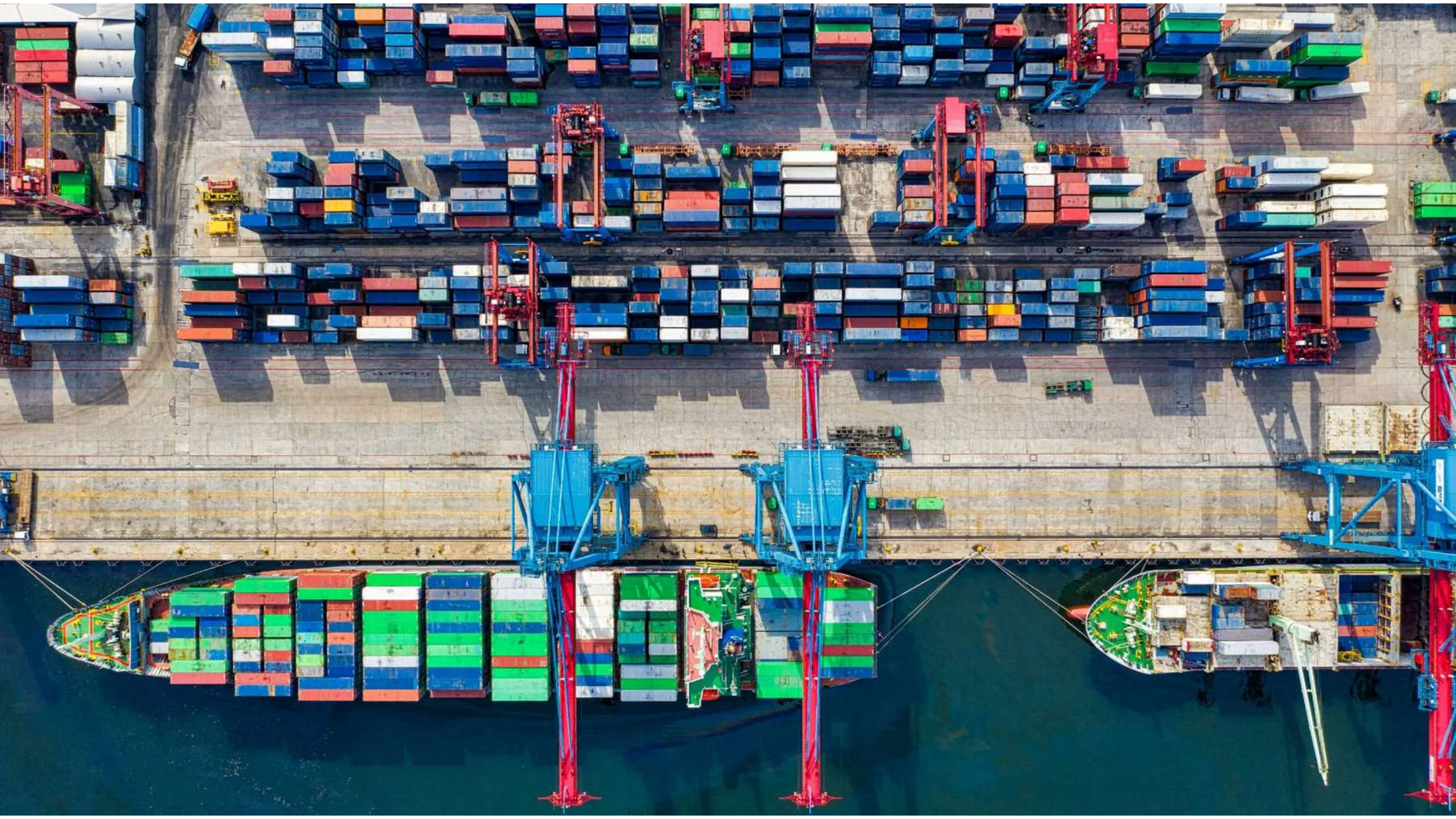


Unoptimized Route

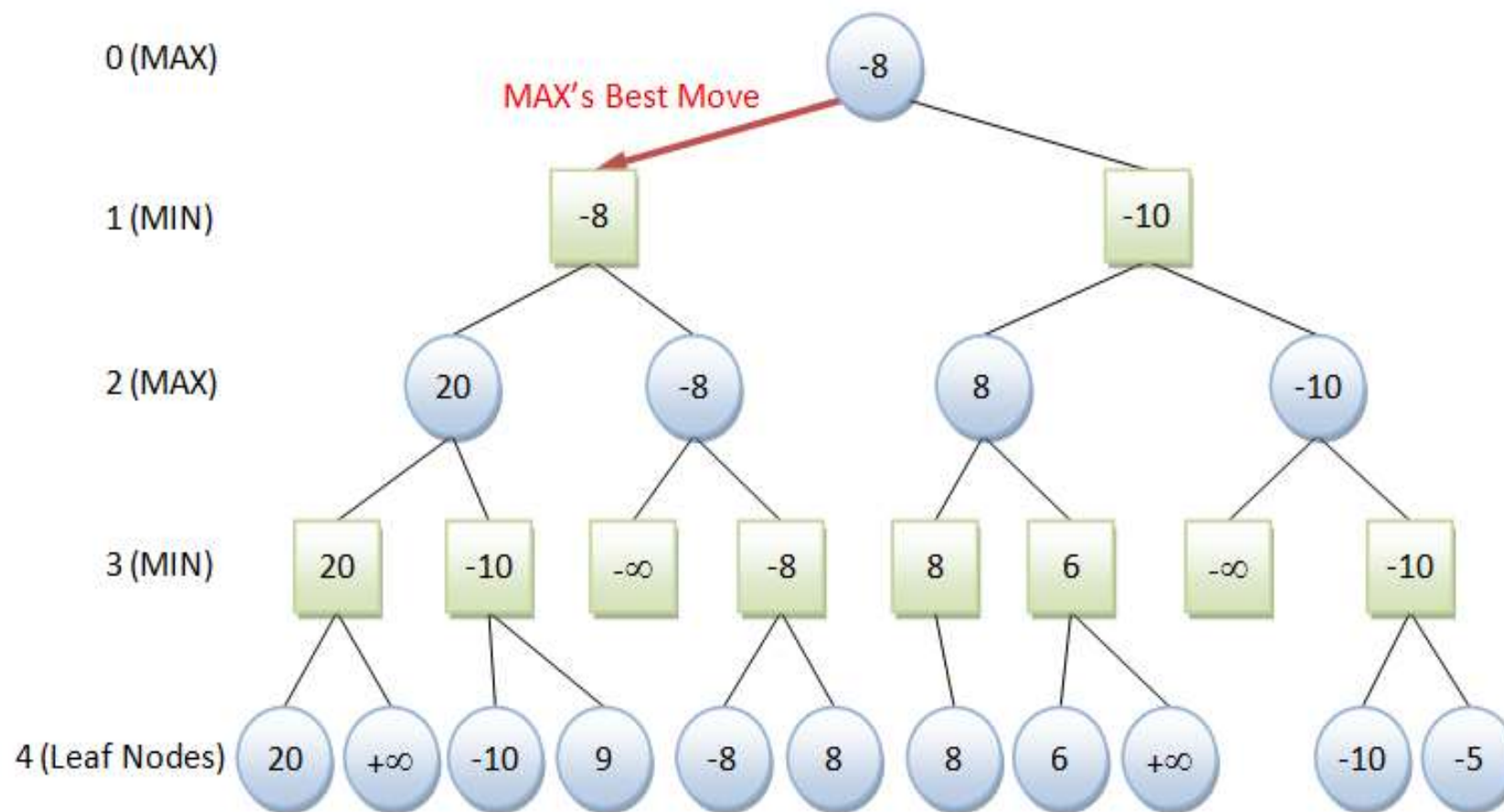


Optimized Route

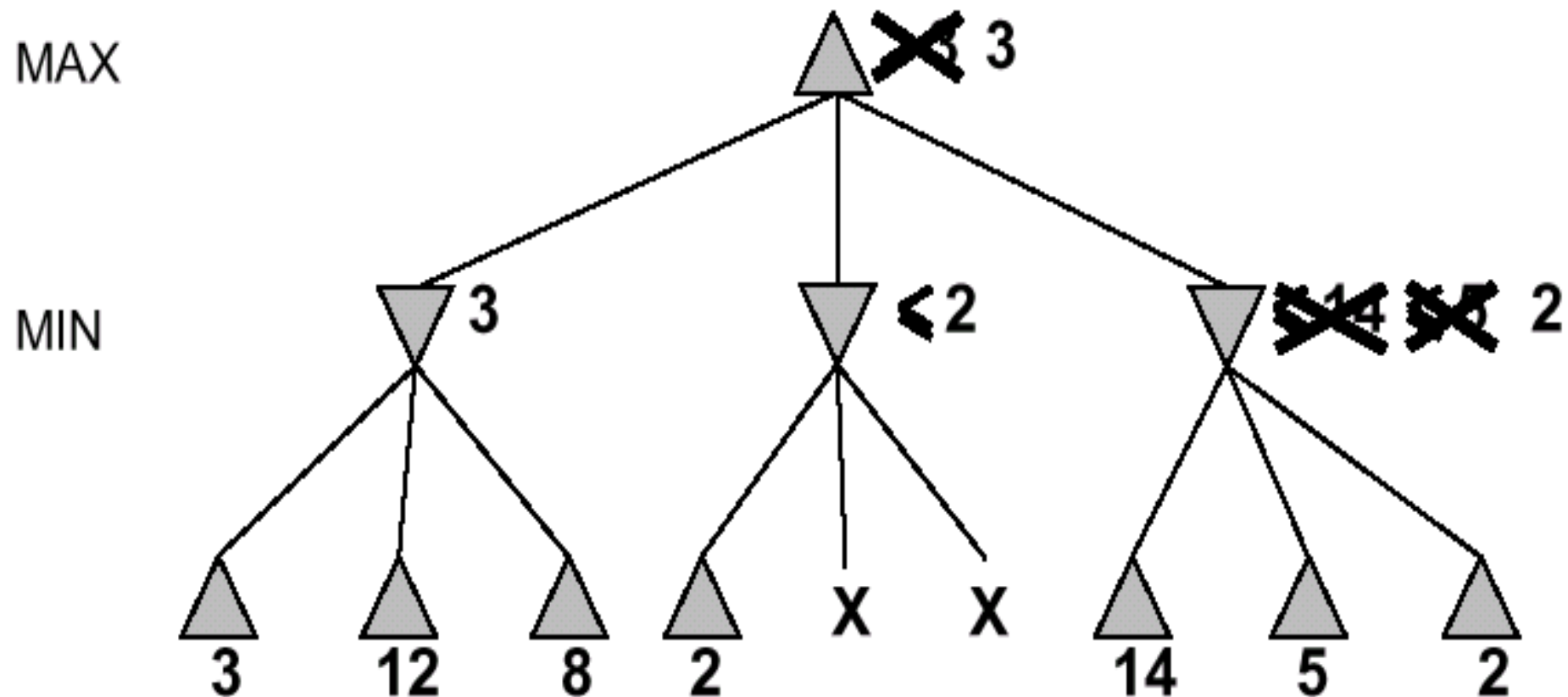




Otsing mängupuust: **minimax** algoritm



Üks esimesi universaalseid otsinguoptimeerimise meetodeid: alpha-beta



Rule-based systems, including expert systems

Given

`father(john,pete)` John is a father of Pete

`father(pete,mark)` Pete is a father of Mark

`father(X,Y) & father(Y,Z) => grandfather(X,Z)`

`father(X,Y) & mother(Y,Z) => grandfather(X,Z)`

If X is a father of Y and Y is a father/mother of Z, then X is a grandfather of Z

`grandfather(X,Y) => grandson(Y,X)`

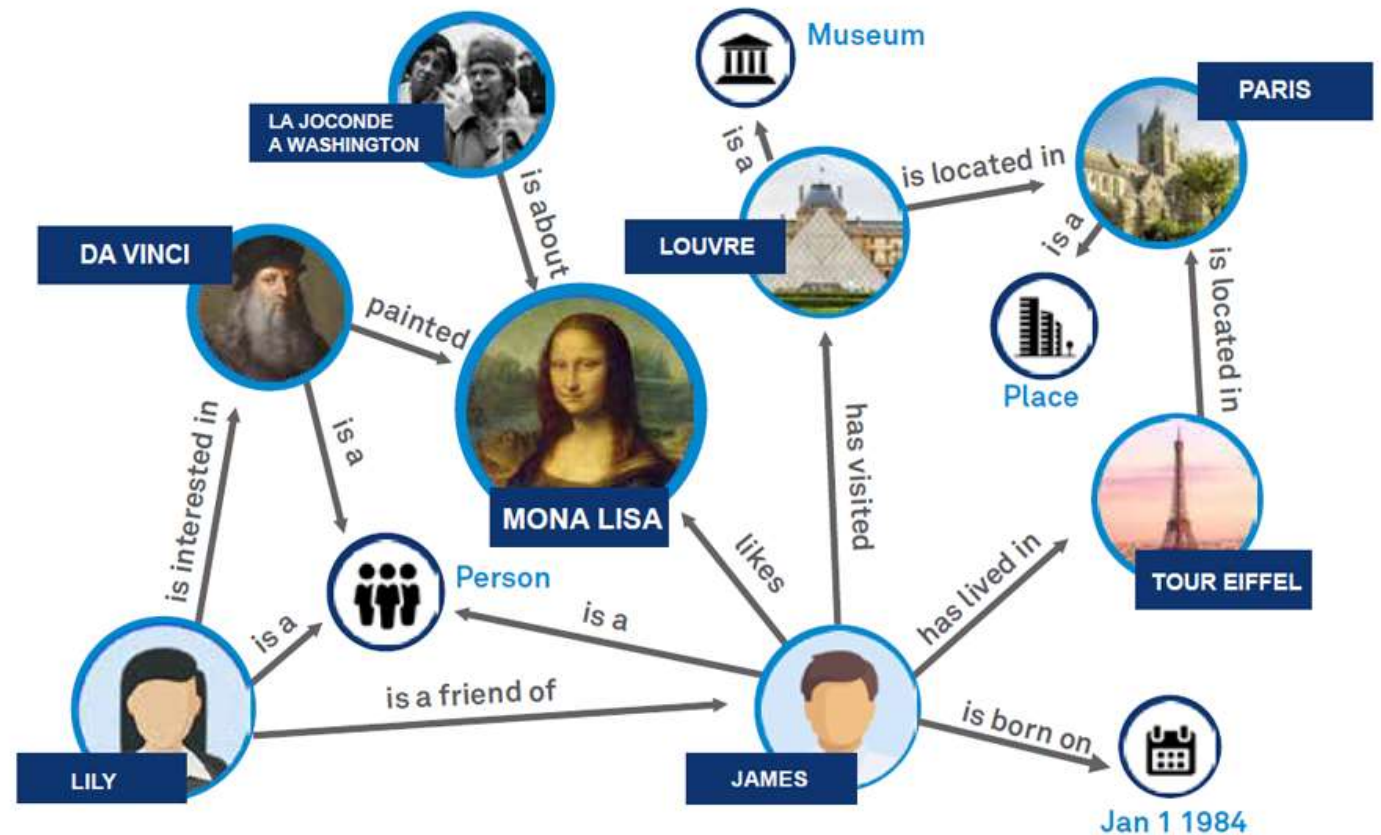
If X is a grandfather of Y then Y is a grandson of X

Exists X : `grandson(X,john)` ? Find a grandson of John



Knowledge graphs

A drastically **simplified form** of rule-based / expert systems



$m(X, i(X)) = e$. % i is an inverse function like one divided by X
 $m(X, e) = X$. % e is a unit element like 1 when on the right side
 $m(X, m(Y, Z)) = m(m(X, Y), Z)$. % m is an associative operation

% Question: is e also a unit element when it is on the left
% side of multiplication m?

$m(e, c) \neq c$.

% proof:

- 1: [in] =(m(X,i(X)),e).
- 2: [in] =(m(X,m(Y,Z)),m(m(X,Y),Z)).
- 3: [=, 1, 2.0.3] =(m(X3,e),m(m(X3,Y3),i(Y3))).
- 4: [in] =(m(X,e),X).
- 5: [simp, 3, 4] =(X,m(m(X,Y),i(Y))).
- 6: [=, 1, 5.0.3] =(X,m(e,i(i(X)))).
- 7: [=, 5.0.R, 5.0.3] =(m(X,Y),m(X,i(i(Y)))).
- 8: [in] -=(m(e,c),c).
- 9: [=, 6.0.R, 7.0.4, 8] false

Solve a predicate logic reasoning task:

[help](#)

% a simple example: using two facts and two rules, find a grandson of john

```
father(john,pete).  
father(pete,mark).  
(father(X,Y) & father(Y,Z)) => grandfather(X,Z).  
grandfather(X,Y) => grandson(Y,X).  
  
grandson(X,john) => $ans(X).
```

Prove

simple

examples



complex

examples



or

Browse

Advanced

Clear result

Clear input



Result:

result: proof found

answer: \$ans(mark).

proof:

```
1: [in] -father(X,Y) | -father(Z,X) | grandfather(Z,Y).  
2: [in] father(pete,mark).  
3: [mp, 1, 2] -father(X,pete) | grandfather(X,mark).  
4: [in] father(john,pete).  
5: [mp, 3, 4] grandfather(john,mark).  
6: [in] -grandfather(X,Y) | grandson(Y,X).  
7: [mp, 5, 6] grandson(mark,john).  
8: [in] -grandson(X,john) | $ans(X).  
9: [mp, 7, 8] $ans(mark).
```

Masinõpe

- Lihtsustatult öeldes: “äärmuslikult vinge statistika”.
- Kõige tuntum meetodite klass on **närvivõrgud**, mis omakorda jagunevad paljudeks alamtüüpideks.
- Kasutusel ka suur hulk **mitte-närvivõrkude**-põhiseid masinõppe algoritme: vektormasinad, lähimad naabrid, juhustemetsad, ...

Machine learning: what is the point?



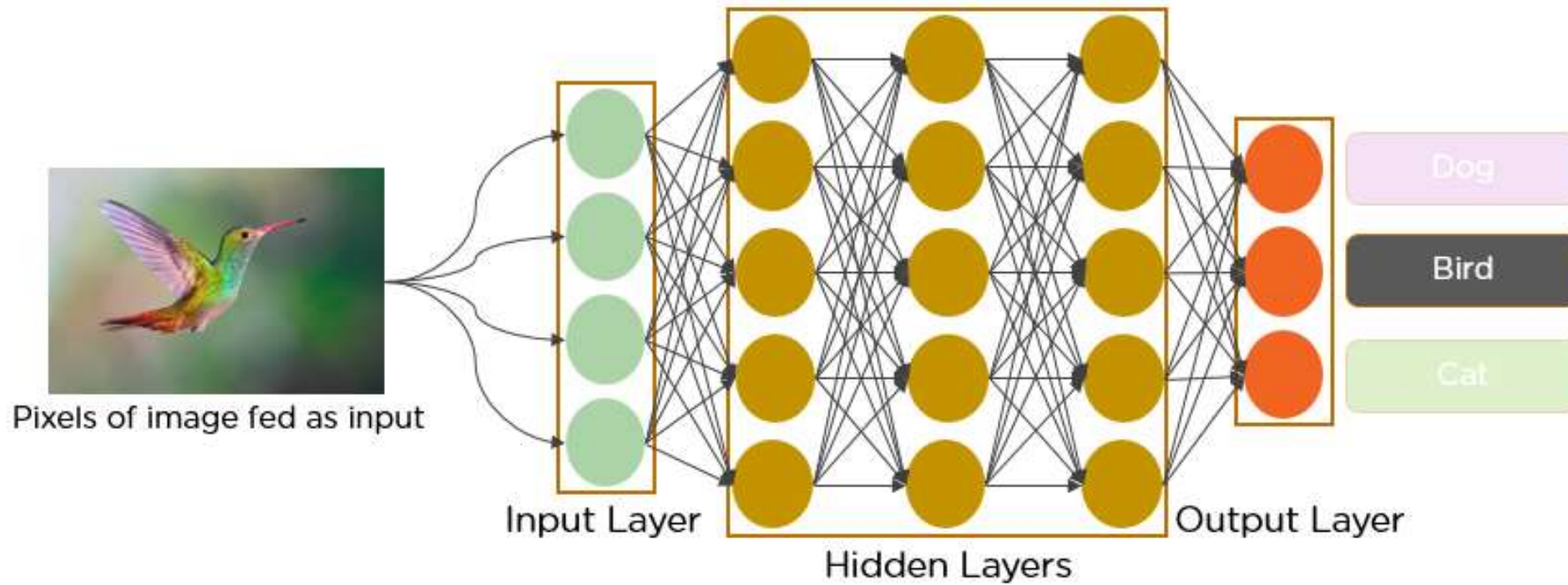
“cat”



“dog”

Writing a program to recognize whether an image is of a cat or a dog is theoretically possible, but extremely complicated: **not realistic in practice.**

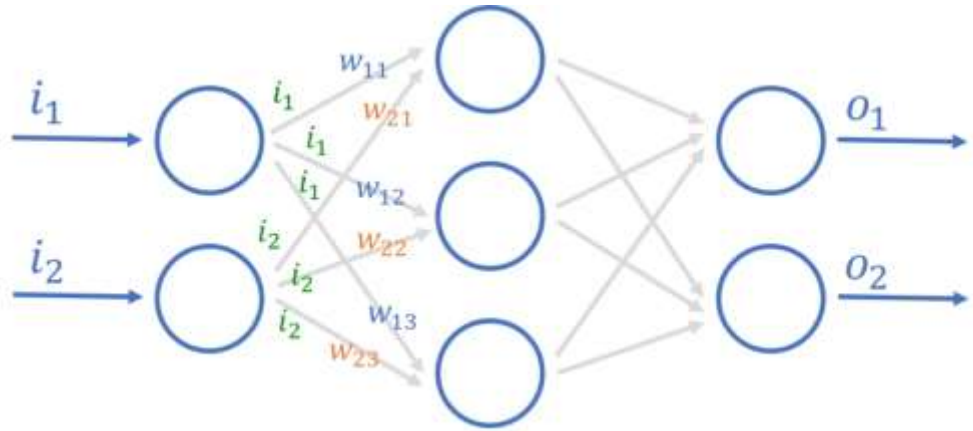
Deep neural networks ehk sügavad närvivõrgud



Matrices and calculations

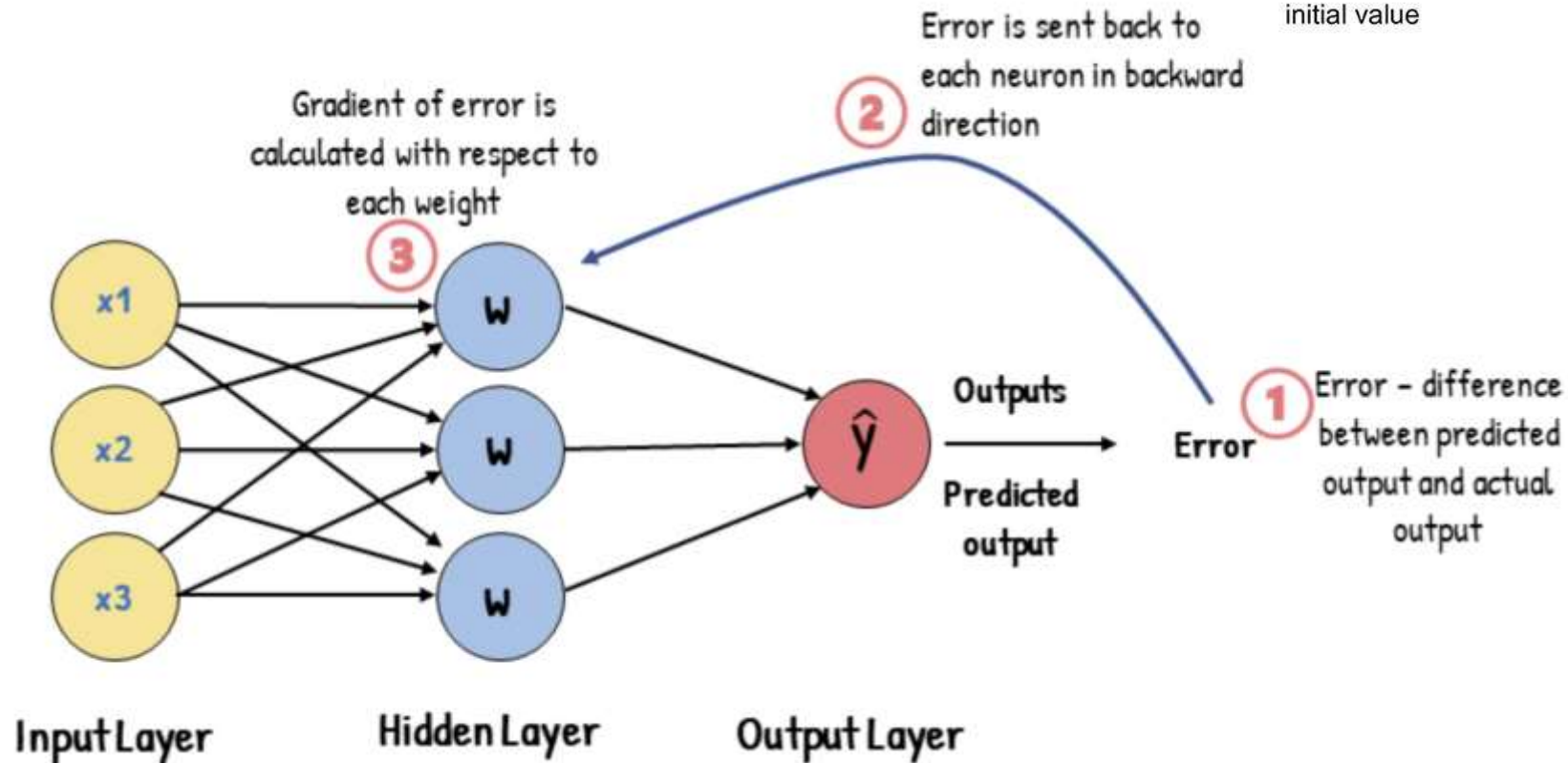
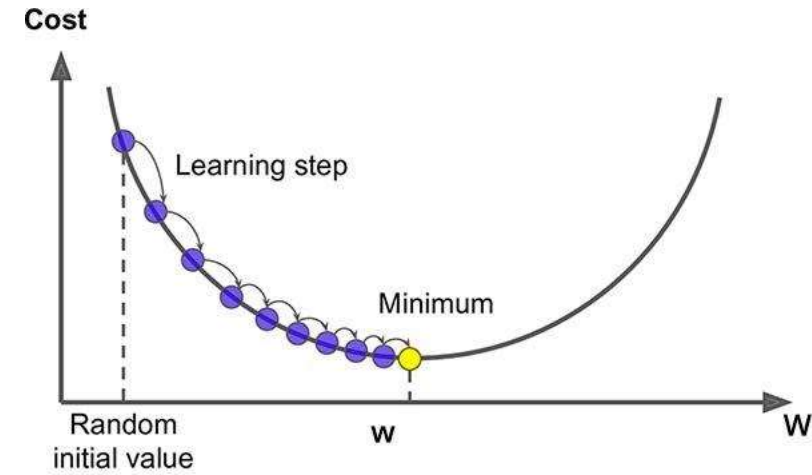
Connections are simply numbers (weights) you can store in a large table (matrix).

Calculating values for nodes can be done in several ways. One simple approach is to first multiply weights with inputs, sum all of them together and then normalize, using some popular simple function for this.

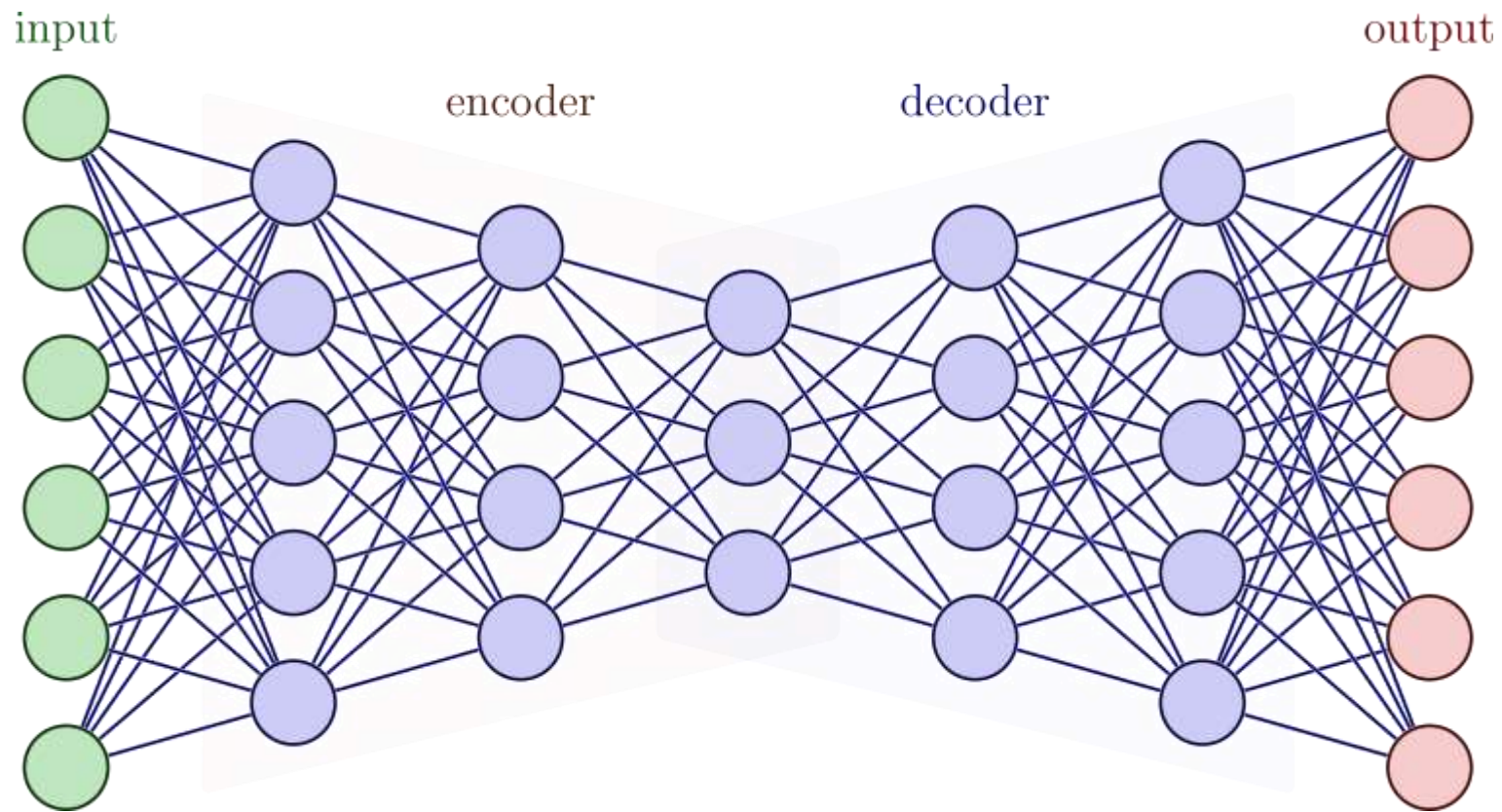


$$\begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \\ w_{13} & w_{23} \end{bmatrix} \cdot \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} (w_{11} \times i_1) + (w_{21} \times i_2) \\ (w_{12} \times i_1) + (w_{22} \times i_2) \\ (w_{13} \times i_1) + (w_{23} \times i_2) \end{bmatrix}$$

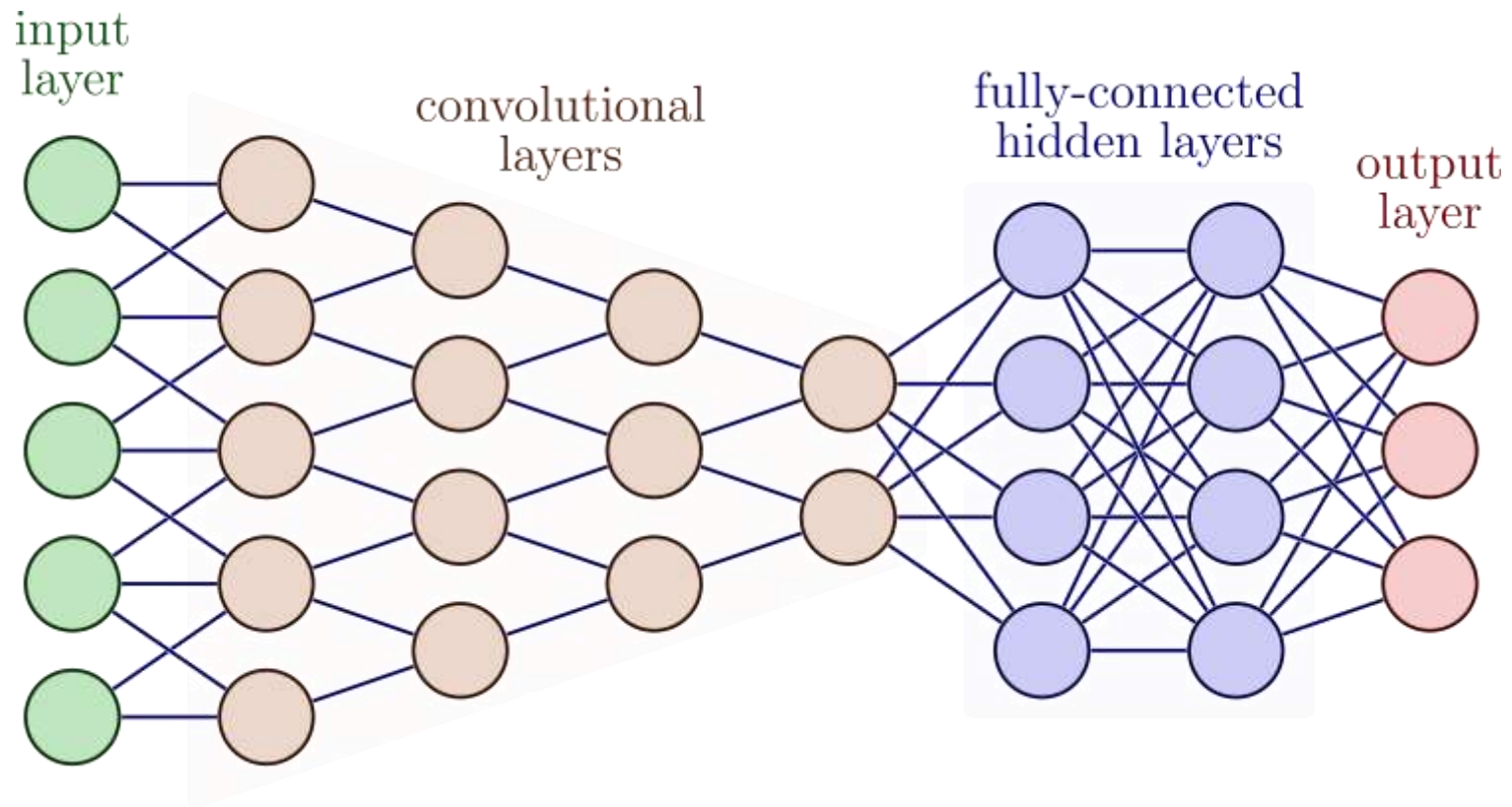
Learning by backpropagation and gradient descent



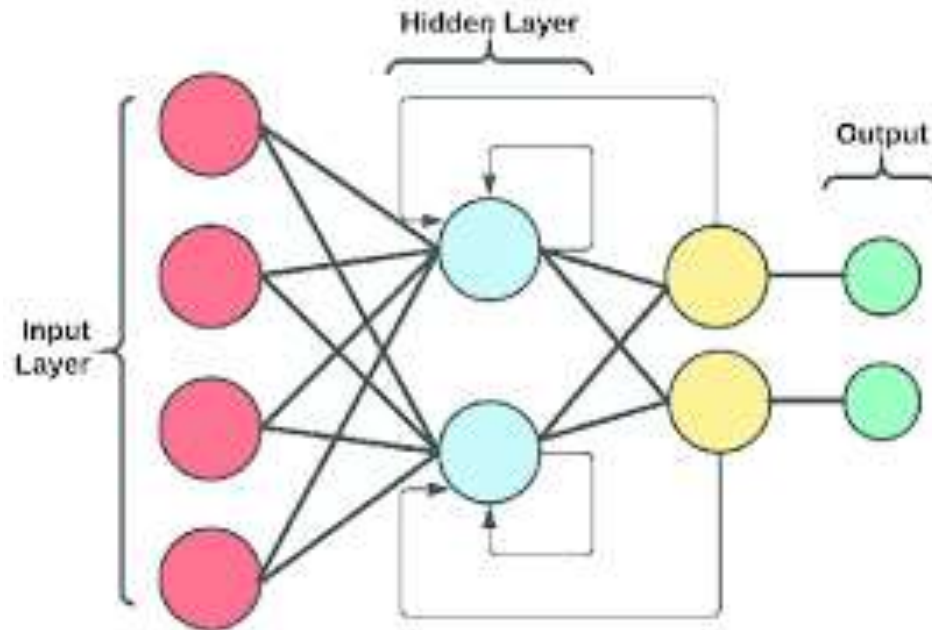
Autoencoder neural network (denoising, compression, generation)



Convolutional neural network (image recognition)



Recurrent neural network (time series, etc)



(a)



school bus 1.0

(b)



garbage truck 0.99

(c)



punching bag 1.0

(d)



snowplow 0.92



motor scooter 0.99



parachute 1.0



bobsled 1.0



parachute 0.54



fire truck 0.99



school bus 0.98

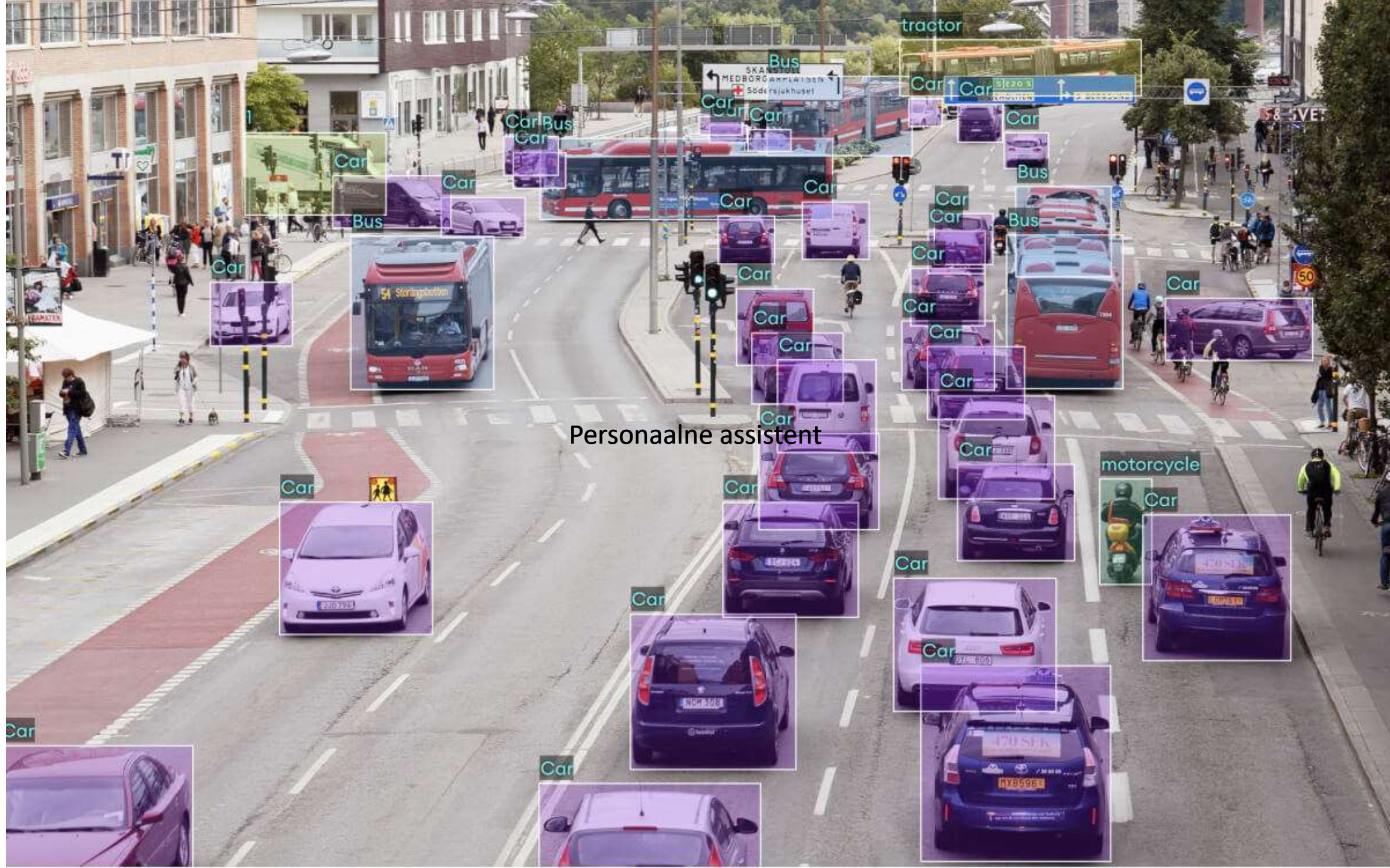


fireboat 0.98



bobsled 0.79

Personaalne assistent



Mida suured keelemudelid (LLM) teevad?

Ennustavad samm-sammult teksti jätkumist:

- Jaan istus autosse ja ...
- Jaan istus autosse ja **hakkas** ...
- Jaan istus autosse ja hakkas **sõitma** ...
- Jaan istus autosse ja hakkas sõitma **kodu** ...

Kuidas LLM ennustab? Lihtsustatult:

- Kogub statistikat suure tekstihulga pealt: mismoodi seal on laused jätkunud, näiteks

→ Mihkel istus mootorrattale ja ... (*hakkas, mõtles, ...*)

- Kui näeb lauset, mille kohta statistikat ei ole, kasutab statistikat varasemate sarnaste lausete kohta:

→ Jaan istus autosse ja ...

Human: how James Joyce and Vladimir Lenin first met?

ChatGPT: James Joyce and Vladimir Lenin met in Zurich, Switzerland in 1916. Both men were living in exile in Zurich during World War I. Joyce was a writer and Lenin was a revolutionary. They met at the Cafe Odéon, a popular gathering place for artists and intellectuals in Zurich.

Comment: an encounter of these people has never been confirmed

Human: A man and a goat are on one side of the river. They have a boat. How can they both go across?

ChatGPT: The man takes the goat across the river first, leaving the boat on the original side. Then, the man returns alone with the boat. He leaves the goat on the other side and takes the boat back to the original side. Finally, he takes *the cabbage* across the river.

Comment:



LLM hallucinations: confidently generating wrong answers.

Why?

LLM hallucinations: confidently generating wrong answers.

Why?



LLM hallucinations: confidently generating wrong answers.

Why?

- LLMs learn an approximate, highly compressed picture of the world knowledge.
- LLMs are always hallucinating: just that mostly these correspond to reality
- LLM learning optimizes for confident tone of voice and pleasing answers
- ... etc ...

LLMi põhisuutlikkused

- Üliinimlikul tasemel ahvimine: genereerib sobivas stiilis mõistlikku teksti.
- Head üldteadmised.
- Hea kategoriseerija.
- Suudab kasutada näiteid.
- Suudab teha mõõdukaid üldistusi ja lihtsamaid järeldusi.
- ...

LLMi põhiprobleemid

- Hallutsineerib ehk leiutab “fakte”, mis ei ole üldse tõesed.
- Vaikimisi omab ainult üldteadmisi.
- Ei ole mälu: uusi asju raske juurde õpetada.
- “Must kast” ehk ei anna tegelikke seletusi, kuidas vastused saadi.
- Ei sisalda otsialgoritme ja enamasti ei suuda teha pikemaid järeldusahelaid.
- Ei ole lihtsat viisi siduda andmebaasidega, inseneriarvutustega jne.
- Vastuste kvaliteet langeb kiirelt, kui sisendtekst on pikk.
- Ei ole selget viisi tema suunamiseks või mingit sorti vastuste keelamiseks.
- ...



Neurosümboolsed
ehk hübriidsüsteemid

	Common AI Techniques					
Use Case Families	Generative Models	Non-Generative ML	Optimisation	Simulation	Rules	Graphs
Forecasting	Low	High	Low	High	Medium	Low
Planning	Low	Low	High	Medium	Medium	High
Decision Intelligence	Low	Medium	High	High	High	Medium
Autonomous System	Low	Medium	High	Medium	Medium	Low
Segmentation	Medium	High	Low	Low	High	High
Recommender	Medium	High	Medium	Low	Medium	High
Perception	Medium	High	Low	Low	Low	Low
Intelligent Automation	Medium	High	Low	Low	High	Medium
Anomaly Detection	Medium	High	Low	Medium	Medium	High
Content Generation	High	Low	Low	High	Low	Low
Chatbots	High	High	Low	Low	Medium	High
Knowledge Discovery	High	Medium	Low	Low	Medium	High

Research area of neurosymbolic A.I. aka hybrid A.I.

Find ways to combine M.L. with symbolic components to be able to solve different really hard tasks.

Some approaches:

- Use M.L. to heuristically guide proof search for symbolic reasoning
- Build LLMs capable of autoformalization and then use symbolic reasoning
- Let LLMs learn on symbolic knowledge in addition to free text
- Build new architectures for M.L. capable of learning external software tool use
- Build new architectures for M.L. capable of performing search / planning
- Build symbolic rules already into a neural switch level of M.L. systems

Explainable A.I. ehk selgitatav tehismõistus

Masinõppe süsteemide tehtud otsuste taga on suur hägune võrgustik seoseid, mitte inim-arusaadavad reeglid.

Otsuste põhjendamiseks oleks lisaks hägusele võrgustikule vaja ehitada mõistetavad reeglid, mille abil otsust põhjendada.

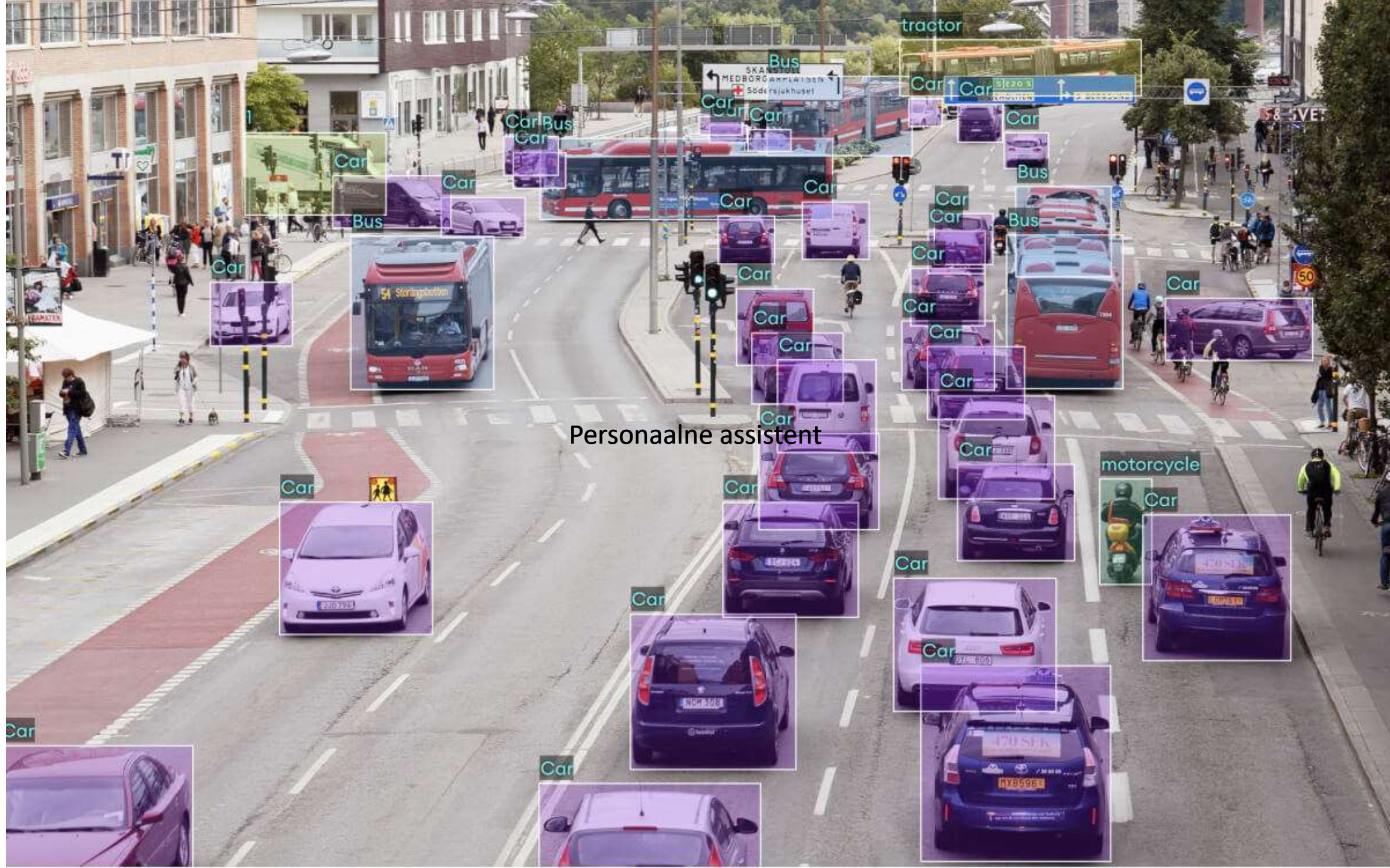
Hetkel puhtalt akadeemiline uurimistöö, praktilisi süsteeme ei osata teha.

*Hübriidsüsteemi näide:
isesõitvad autod*

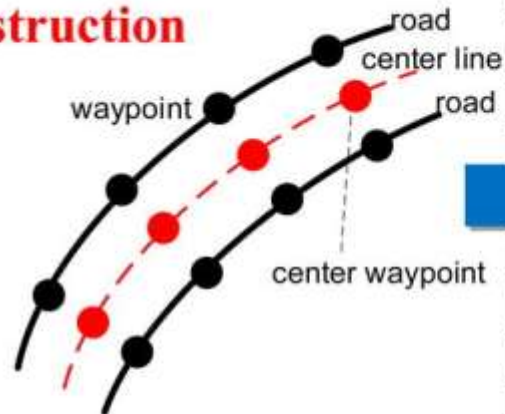
Pildil: 2005,
Stanfordi Stanley võidab
DARPA Grand Challenge II



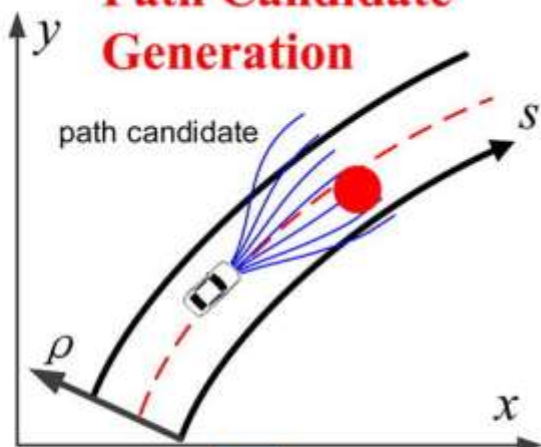
Personaalne assistent



Center Line Construction

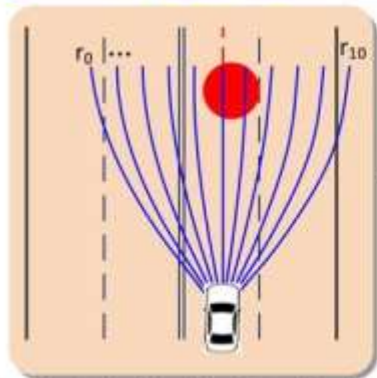


Path Candidate Generation

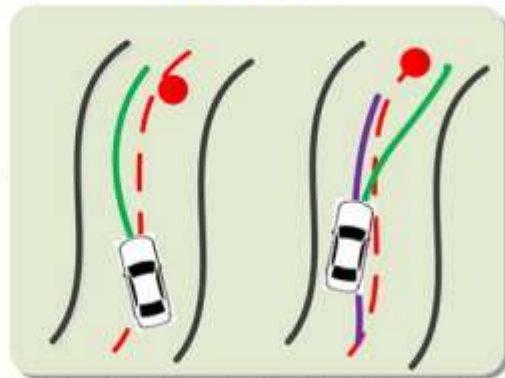


Path Selection

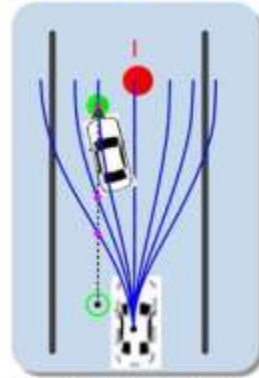
$$f = \omega_s f_s(r_i) + \omega_c f_c(r_i) + \omega_d f_d(r_i, a(r_i))$$



+

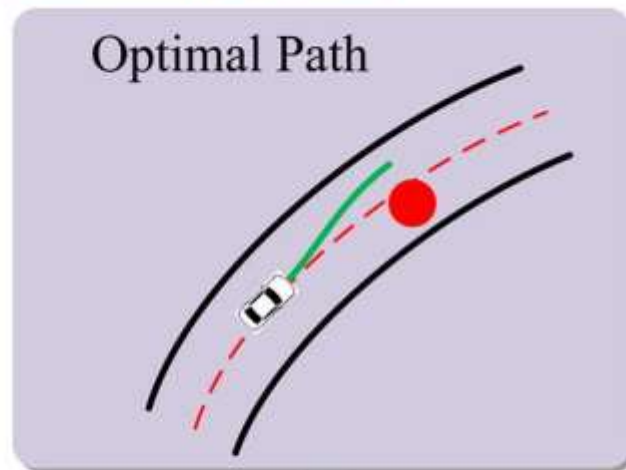


+

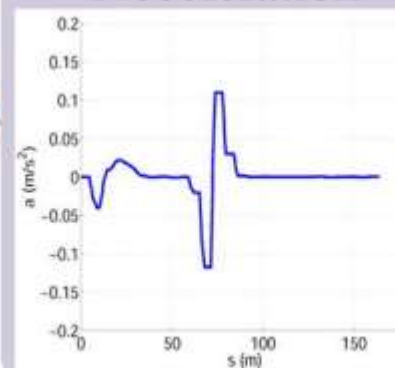


Path Planning Results

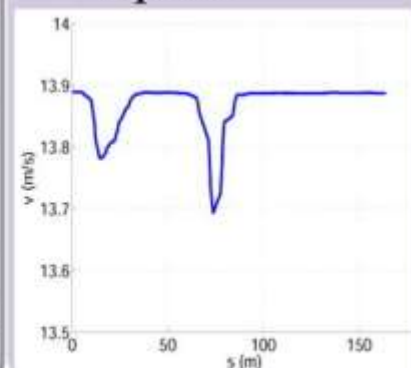
Optimal Path



Appropriate Acceleration



Appropriate Speed





15
WHEN
CHILDREN
ARE PRESENT

Avenue

May stop quickly

Rotini

cruise

23810T3

Rotini

36
TERRACE
University & 15th

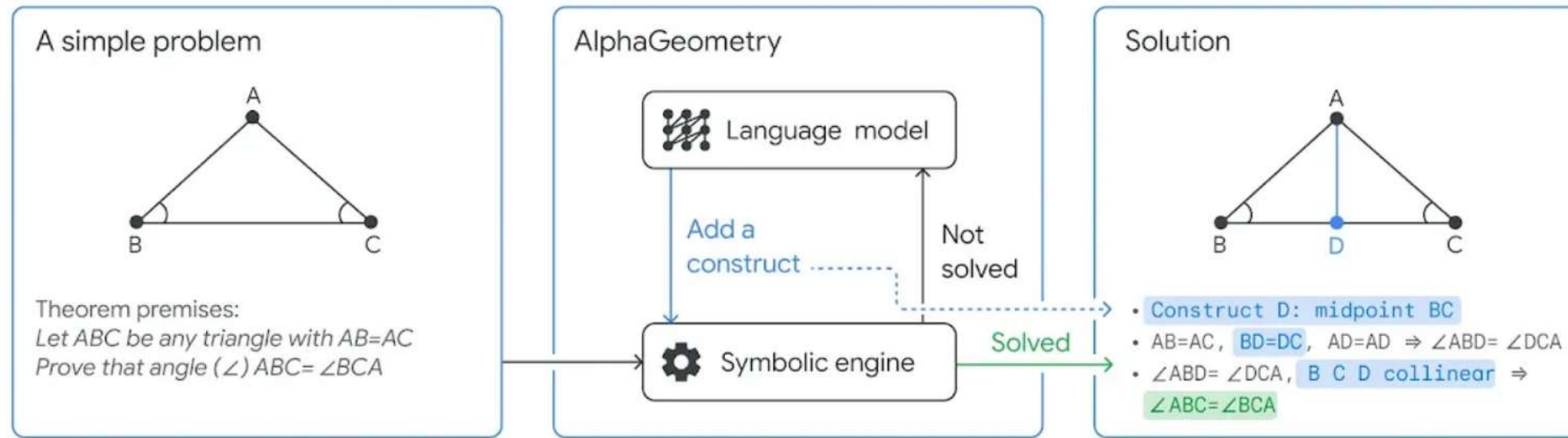
AlphaGo ja AlphaZero: hübriidsüsteemid

DeepMind / Google projekt masinõppeks Go ja male jaoks.

Iseenda vastu mängides õpib seisumustrite järgi seisu headust kiirelt ja hästi hindama.

Hea seisuhindamine on kombineeritud **Monte-Carlo otsinguga**: vaatab juhusliku valikuga läbi hulga võimalikke käikude jadasid mängu lõpuni välja.

Google DeepMind AlphaGeometry: autoformalization + symbolic reasoning engine



Solves complex geometry problems at a level approaching a high-school Olympiad gold-medalist

Minu fookus **neurosümboolse** tehisintellekti valdkonnas:

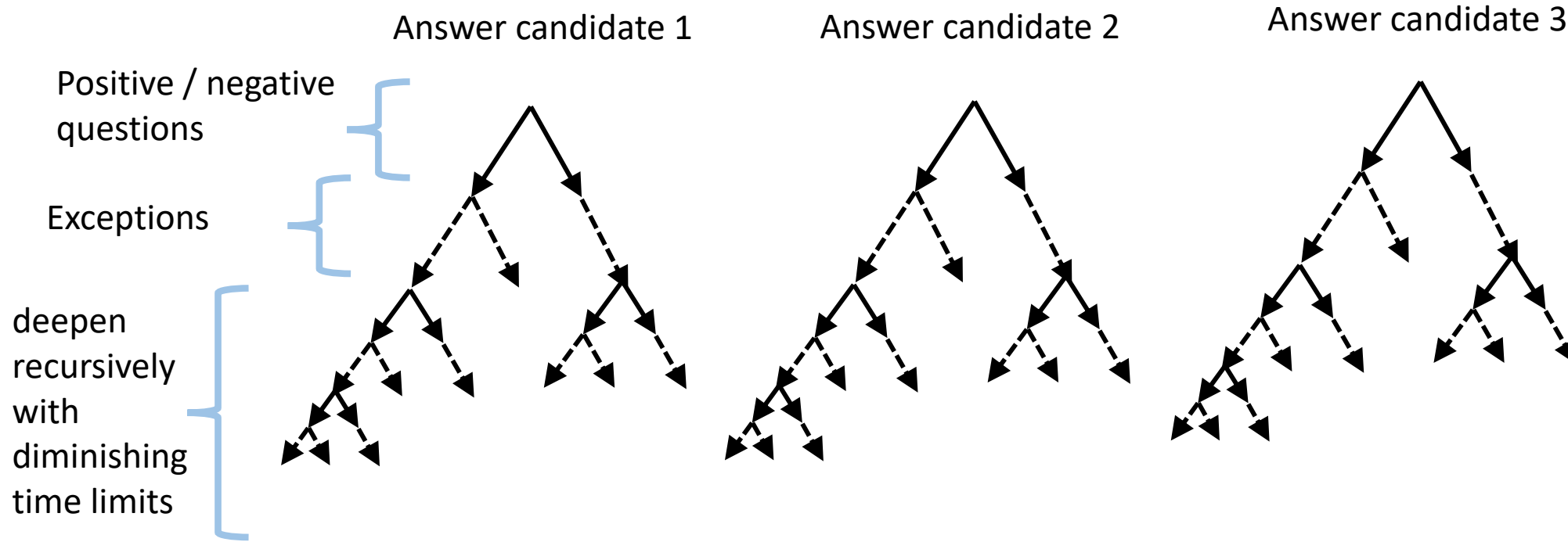
Inimkeelse teksti automaatne formaliseerimine ja korrektsete vastuste leidmine
ja/või keelemudeli antud vastuse valideerimine.

Suurte keelemudelite kasutamine inimkeelse teksti formaliseerimiseks

Mary's car is clean. Mary has a car?

```
["and", ["exists", "X",  
        ["and", ["isa", "car", "X"],  
                ["has", "Mary", "X"],  
                ["has property", "X", "clean"]]]],  
["question",  
 ["exists", "Y",  
  ["and", ["isa", "car", "Y"],  
          ["has", "Mary", "Y"]]]]]].
```

GK loogikajäreldaja rakendamine erandite ja tõenäosustega rikastatud loogikas



Almost all birds can fly. Penguins cannot fly. John is bird. Mike is a penguin. Who can fly?

Robotid

*Robots: a **very** hard area!*



Most robots contain fairly little A.I.

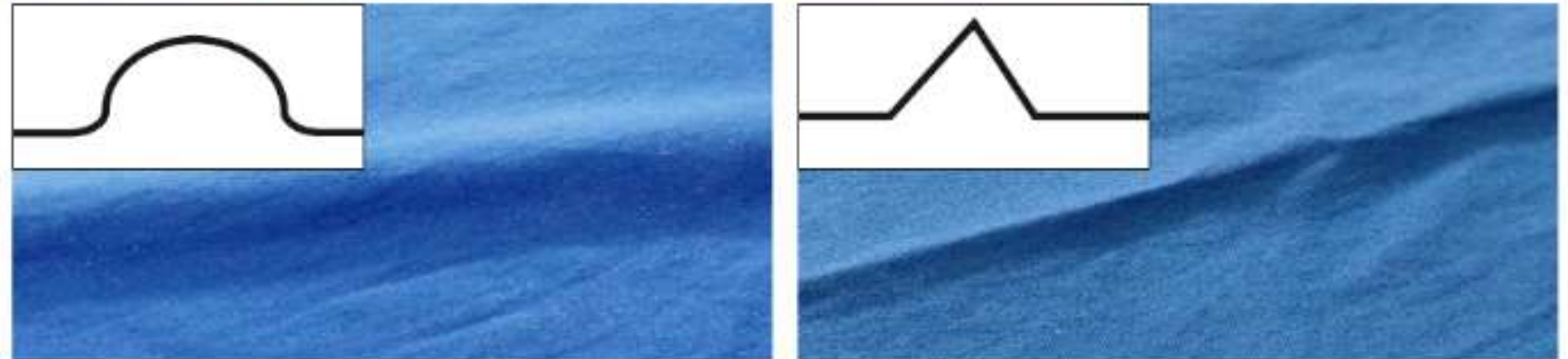


Industrial robots are **mostly hand-coded**,
with A.I. relegated to detection etc



Example:

No real success in developing robots able to iron ordinary clothes.



Example:

Bringing dirty dishes from the table to the dishwasher does not seem realistic in visible future.





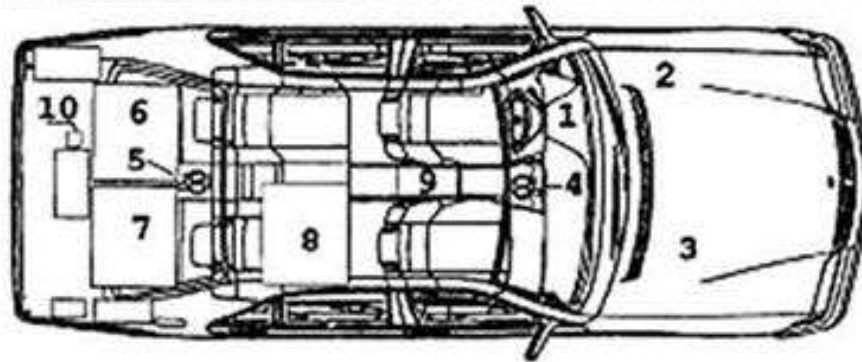
@sarahstacks_



TikTok
@espn

6.127 Second Cycle
Second Fastest Time in US

1980- ja 1990-d:
Isesõitev auto
Saksamaal:
tuhanded
kilomeetrid
kiirteel



- | | |
|------------------------------|--|
| 1 Torque motor for steering | 6, 8 Transputer system, image processing |
| 2 brake system | 7 processors for gaze & locomotion control |
| 3 electric throttle control | 8 user interface |
| 4 front platform } for 2 CCD | 9 linear accelerometers |
| 5 rear platform } cameras | 10 angular rate sensors |

https://en.wikipedia.org/wiki/History_of_self-driving_cars

2005: Stanfordi
Stanley võidab
DARPA Grand
Challenge II



Darpa urban challenge 2007

96 km linnas, aeg 4:10



Tesla

„Traffic-Aware Cruise Control **cannot detect all objects and may not brake/decelerate for stationary vehicles**, especially in situations when you are driving over 50 mph (80 km/h) and a vehicle you are following moves out of your driving path and a stationary vehicle or object is in front of you instead.”

Volvo

„In other words, **your Volvo won't brake to avoid hitting a stopped car that suddenly appears up ahead**“.



WHEN
CHILDREN
ARE PRESENT

Avenue

38
TERRACE
UNDERSTANDING & JOURNAL

Rotini

cruise

May stop quickly

23810T3

Rotini

Filosoofia ja prognoosid

*Tugev ehk „päris“ AI, mida meil ei ole
(ehk AGI: artificial general intelligence):*

Masin on intellektilt inimloomusele lähedane või ületab seda, suudab täita tüüpilisi inimülesandeid, kasutada laia spektrit taustateadmisi ja omab mingil määral eneseteadvust.

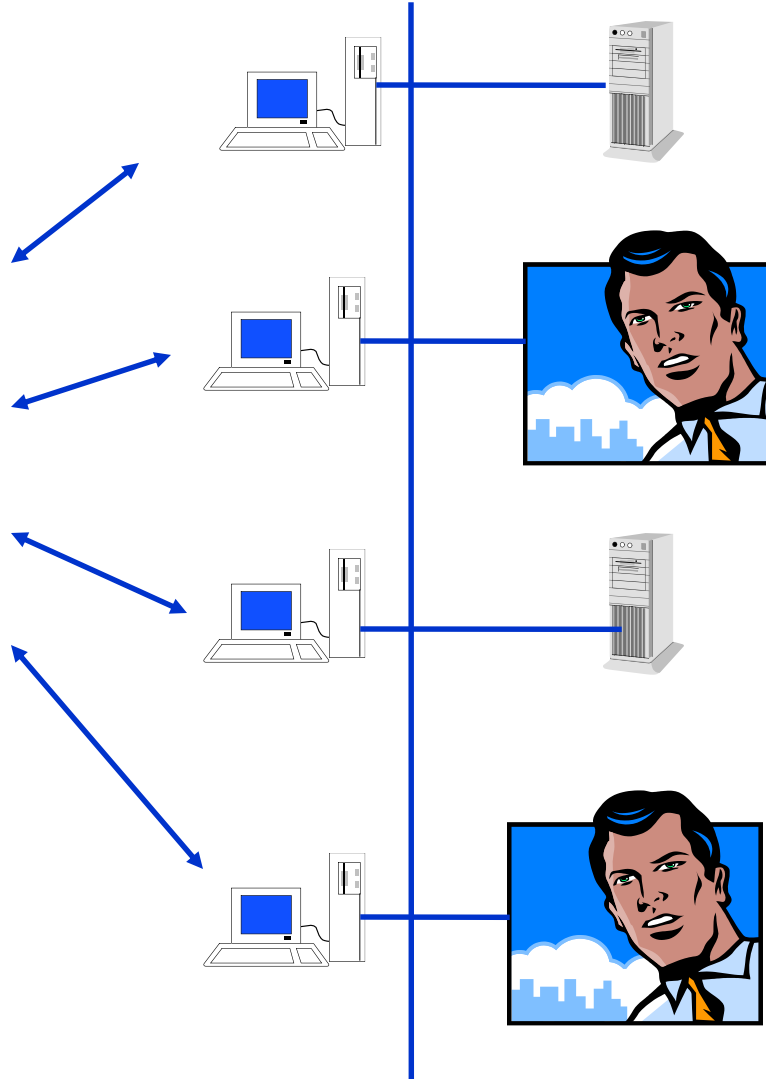
Nõrk AI ehk see, mis tegelikult on:

Tarkvara kasutamine spetsiifiliste probleemide või mõtlemisülesannete lahendamiseks, mis ei vaja laia spektrit inimvõimeid.

Turingi test: filosoofiline idee/ettepanek 50-datest



Mõistata, kas
chati-ekraani
taga on inimene
või programm?



Turing:

Kui katsetajad
ei suuda ära
arvata (st
ära-arvamise
sagedus on
50% ja 50%
eksitakse), siis
on jutlev
masin
päriselt
intelligentne.



**IF ANYONE
BUILDS IT,
EVERYONE
DIES**

**WHY
SUPERHUMAN AI
WOULD
KILL US ALL**

**ELIEZER
YUDKOWSKY &
NATE SOARES**

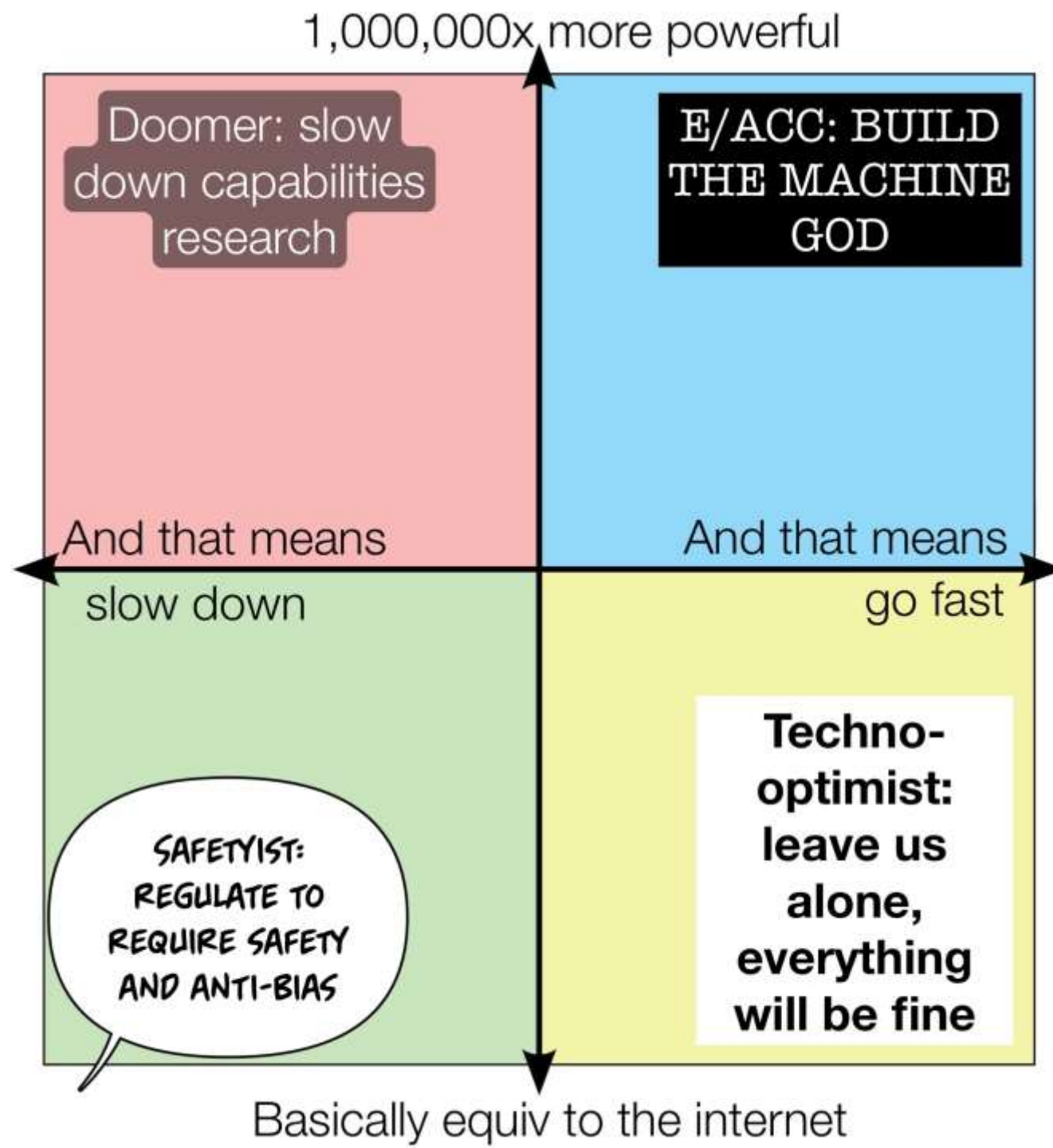
WHY AI WILL SAVE THE WORLD

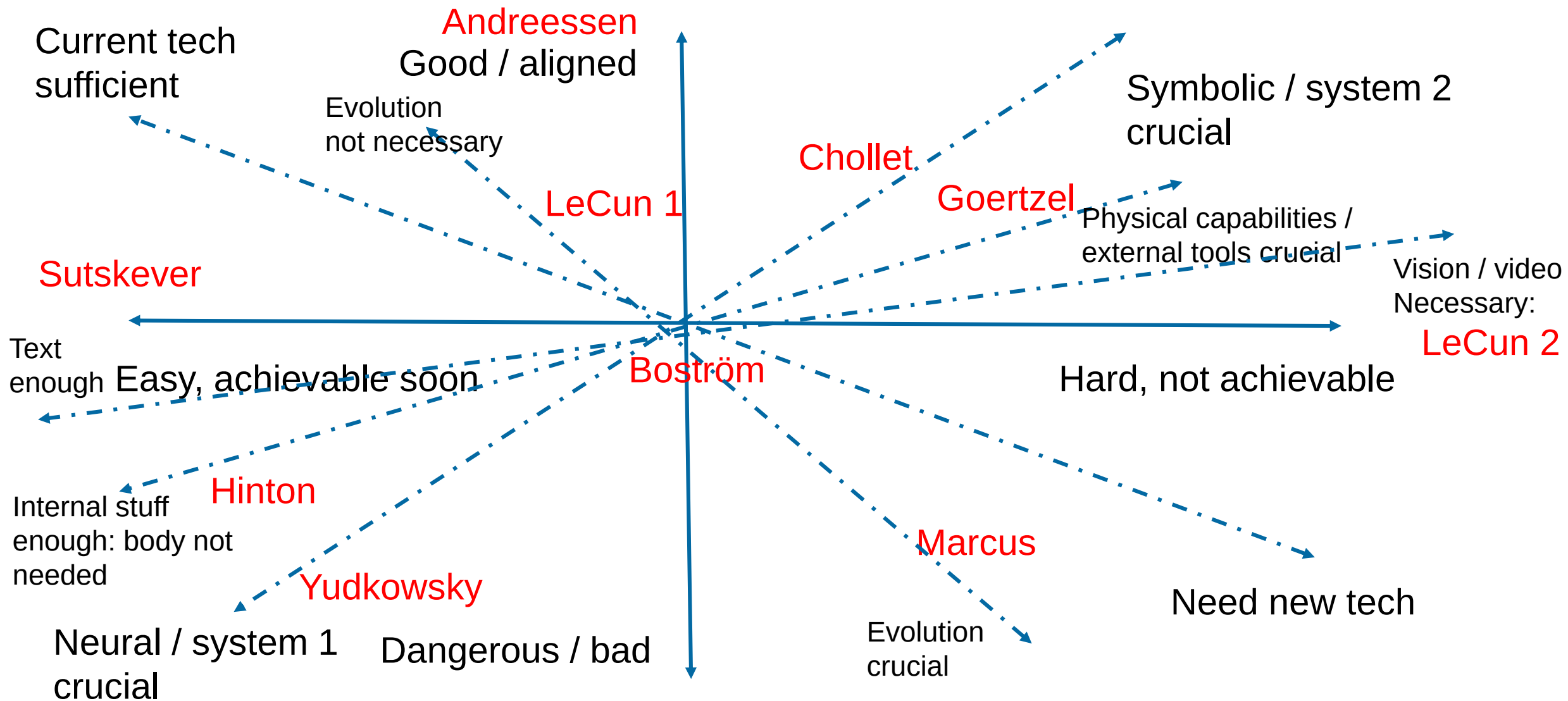
MARC
ANDREESSEN



The AI Race Will **Determine** How We All Live







ON BULLSHIT

Harry G. Frankfurt

Bullshit is speech intended to persuade without regard for truth.

The liar cares about the truth and attempts to hide it; the bullshitter doesn't care whether what they say is true or false.





The bullshitter is neither on the side of the true or the side of the false. His eye is not on the facts at all. He does not reject the authority of the truth, as the liar does, and oppose himself to it. He pays no attention to it at all. By virtue of this, bullshit is a greater enemy of the truth than lies are.

— *Harry Frankfurt* —

AZ QUOTES



European Union A.I. Act: 12 July 2024

The Act assigns applications of AI to several risk categories.

Applications and systems that create an **unacceptable** risk are banned (manipulative / deceptive, exploiting vulnerabilities, inferring emotions in organisations, unrestricted collection of facial databases, social scoring of the type used in China, etc).

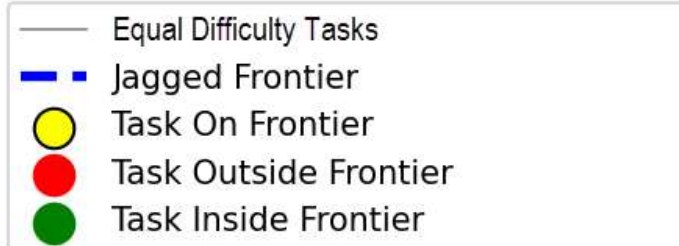
High-risk applications (profiling individuals, remote biometric identification systems, CV-scanning tool that ranks job applicants, determining access, admission or assignment to educational and vocational training institutions, applying the law to disputes, etc) are subject to specific legal requirements

Limited risk AI systems are subject to lighter transparency obligations: developers and deployers must ensure that end-users are aware that they are interacting with AI (chatbots and deepfakes).

Minimal risk is unregulated (including AI applications like AI enabled video games and spam filters).

Tasks AI Can NOT Do

Tasks AI Can Do



Jagged Frontier of AI Task Capabilities

Task Outside Frontier

Task Inside Frontier

Task Outside Frontier

Task On Frontier

Task Inside Frontier

Moravec'i paradoks:

Inimeste jaoks rasked ülesanded on masina jaoks tihtipeale lihtsad.

Inimeste jaoks lihtsad ülesanded on masina jaoks tihtipeale äärmiselt keerulised.

Amara printsiip:

Meil on kalduvus lähiperspektiivis tehnoloogia mõju üle hinnata ja pikemas perspektiivis alahinnata.

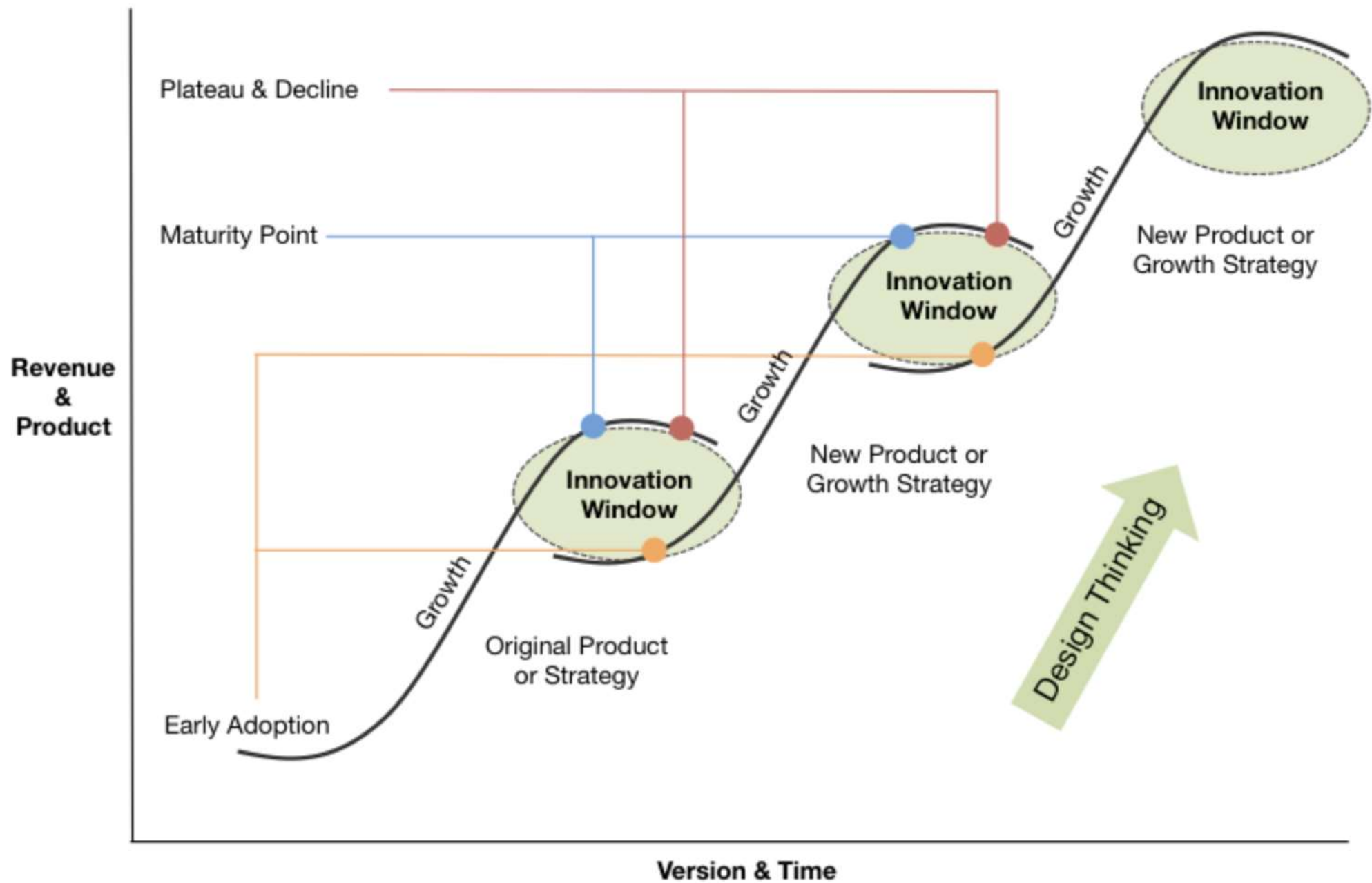
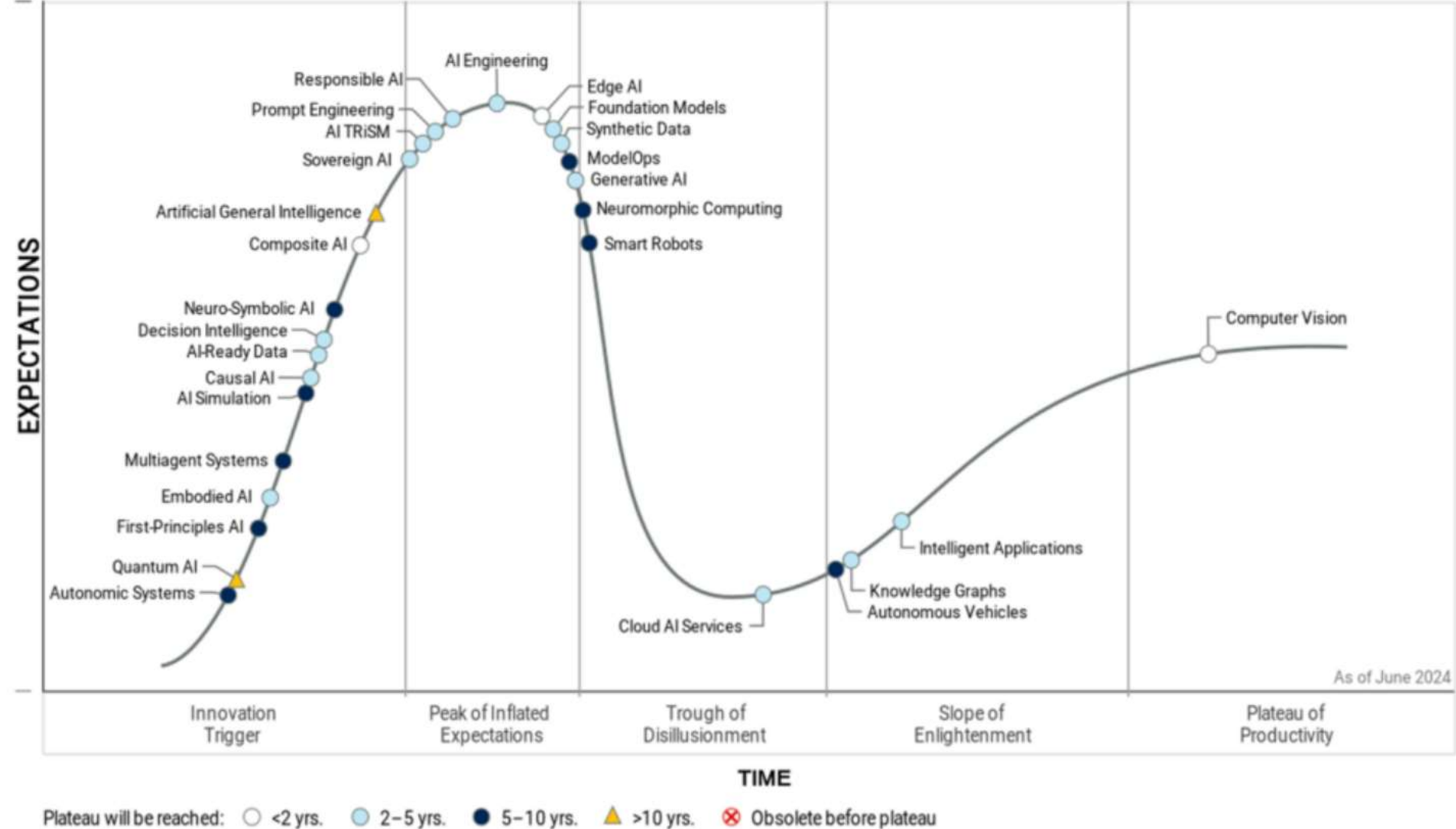
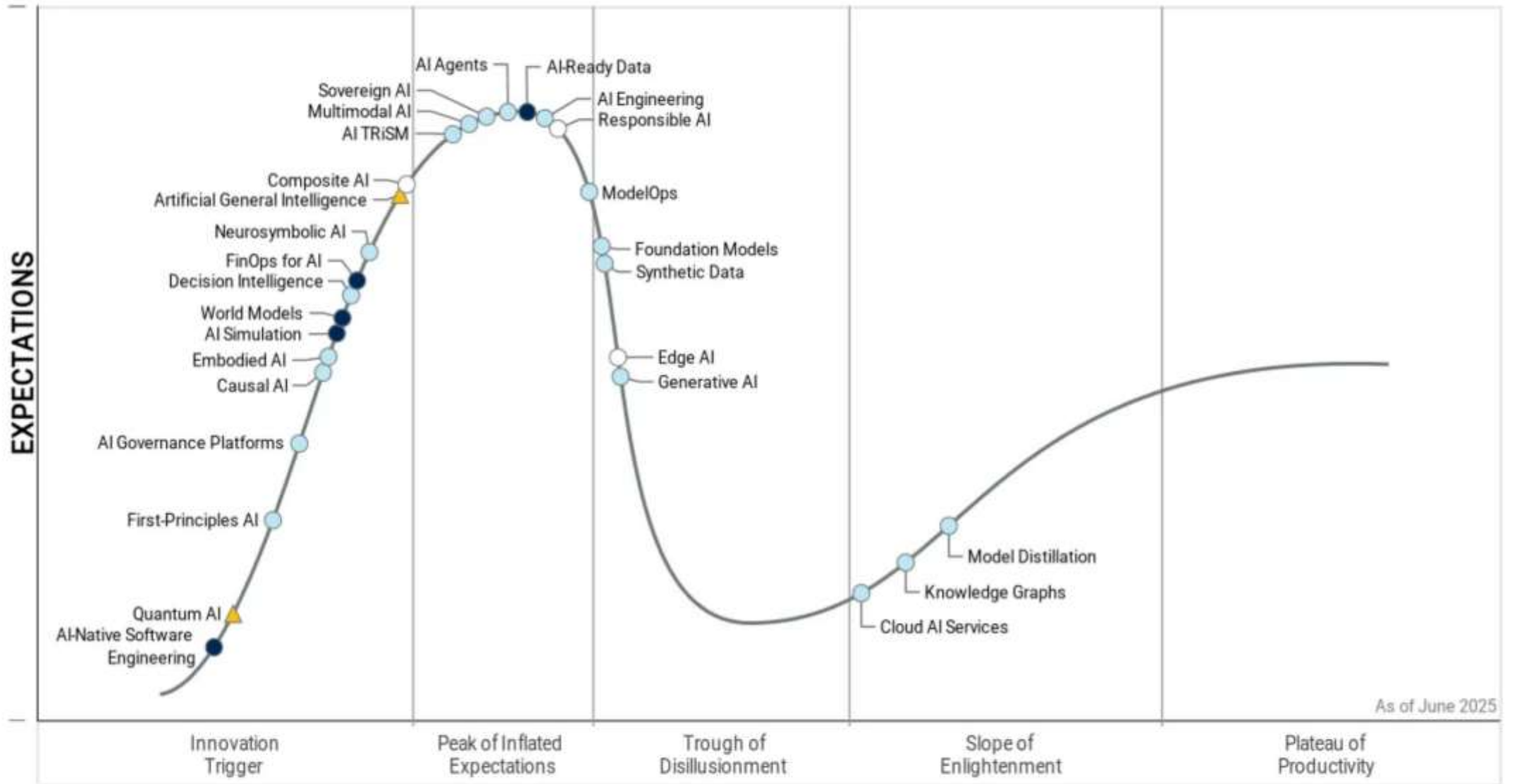


Figure 1: Hype Cycle for Artificial Intelligence, 2024

Hype Cycle for Artificial Intelligence, 2024



Hype Cycle for Artificial Intelligence, 2025



Plateau will be reached: ○ <2 yrs. ● 2-5 yrs. ● 5-10 yrs. ▲ >10 yrs. ✗ Obsolete before plateau

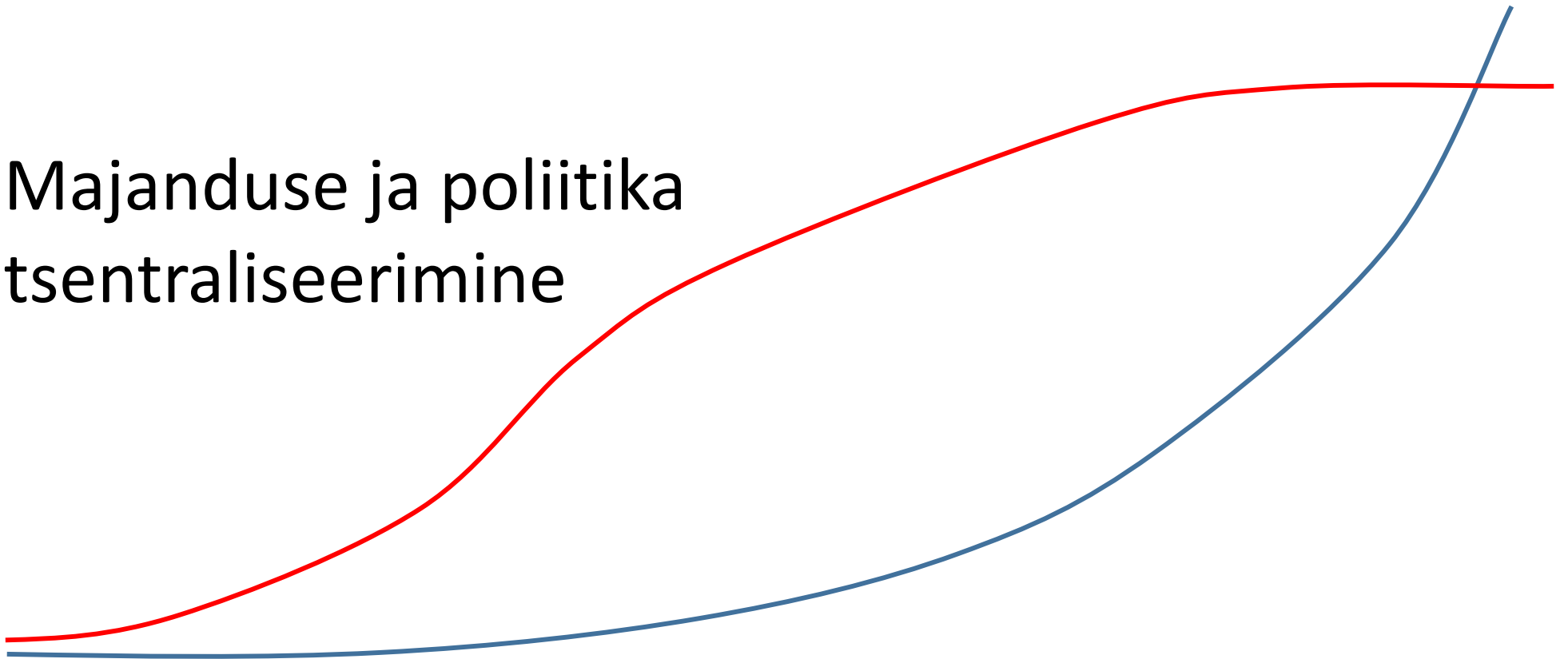
Erinevad suunad, eesmärgid, tehnoloogiad:

- Tsentraliseerimine
- Automatiseerimine
- Tehisintellekt



Vastastikune mõju
ja võimendus

Majanduse ja poliitika
tsentraliseerimine



Tööde automatiseerimine: AI ja
robotika

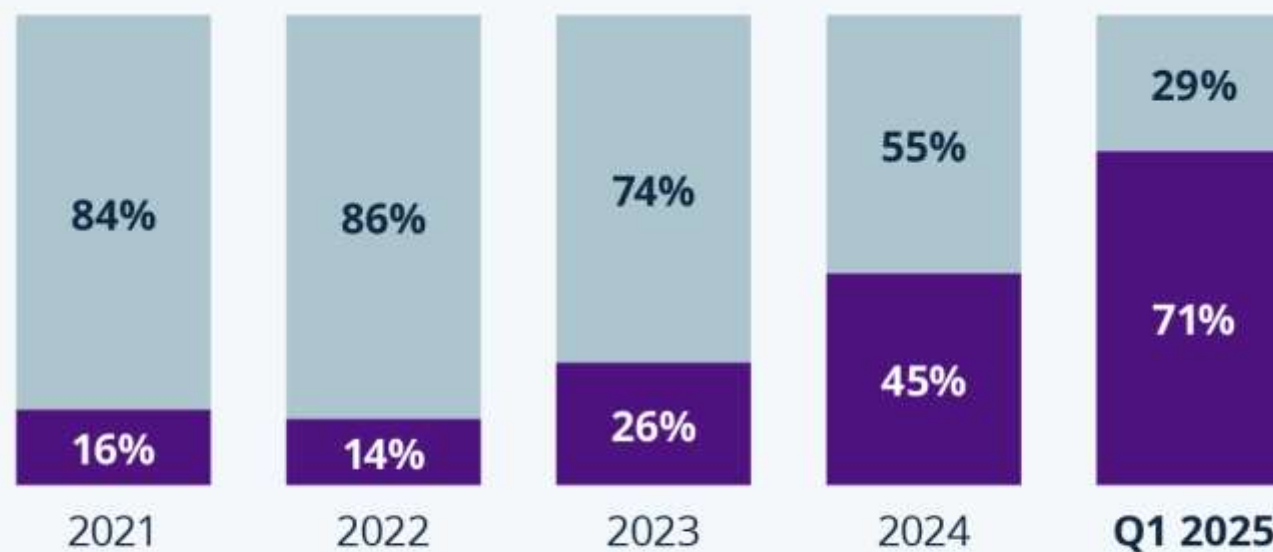
Keelemudelite äri ja kasutuse hetkeseisust

AI Sucks Up a Growing Chunk of VC Funding in the U.S.

Share of equity investments into U.S. VC-backed companies that are (not) AI-related*



■ AI-related investments ■ Non-AI investments



* Sources of cash investments include VC firms, corporate investors, other private equity firms and individuals.

Source: Crunchbase via EY



Humlum & Vestergaard, May 2025

“Large Language Models, Small Labor Market Effects”

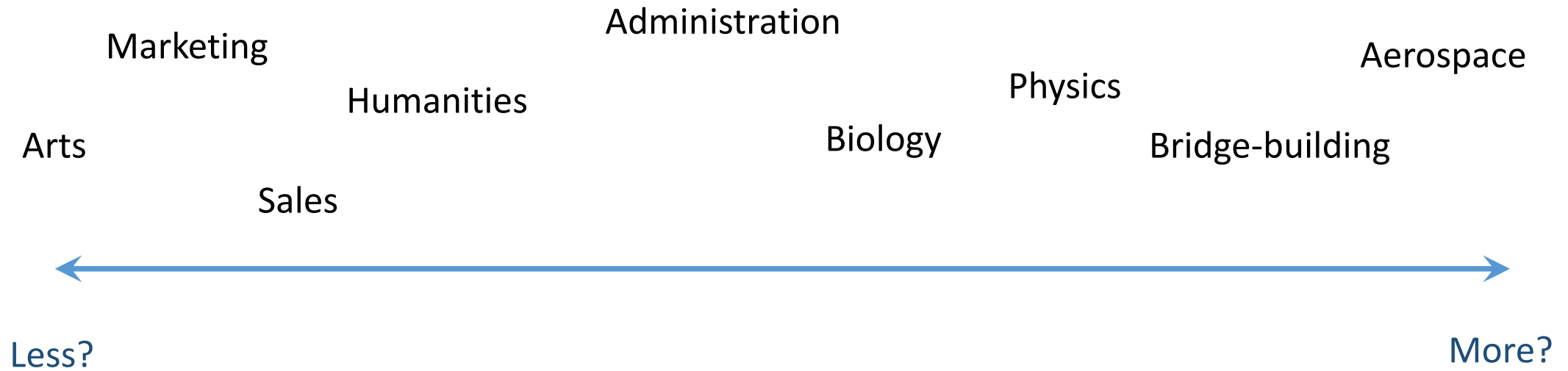
We examine the labor market effects of AI chatbots using two large-scale adoption surveys (late 2023 and 2024) covering 11 exposed occupations (25,000 workers, 7,000 workplaces), linked to matched employer-employee data in Denmark.

AI chatbots have had no significant impact on earnings or recorded hours in any occupation, with confidence intervals ruling out effects larger than 1%. Modest productivity gains (average time savings of 3%), combined with weak wage pass-through, help explain these limited labor market effects.

MIT study: The GenAI Divide: State of AI in Business 2025

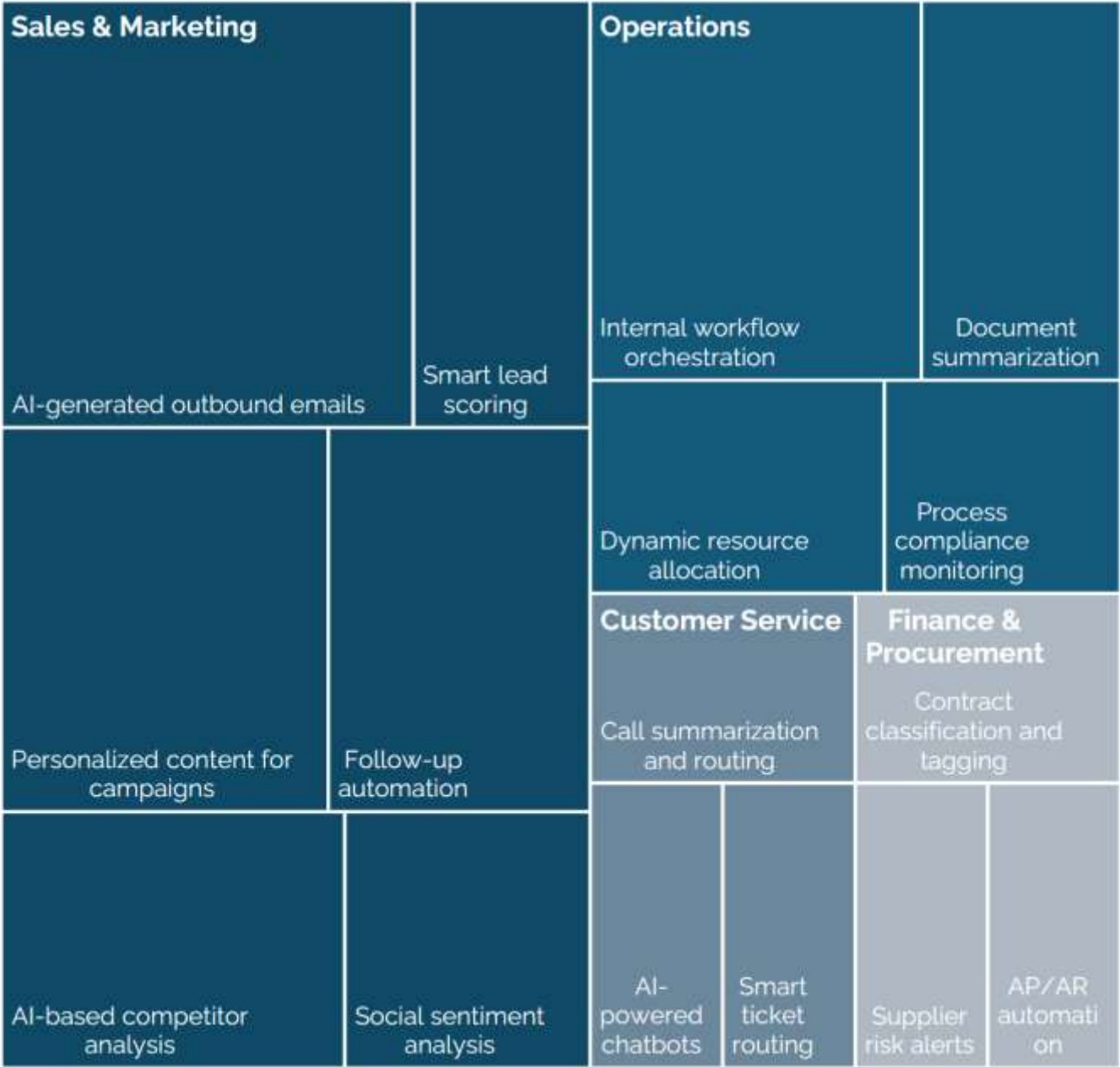
- Only 5% of organizations are seeing transformative returns. The vast majority—95%—report zero impact on profit and loss statements from formal AI investments.
- The core barrier to scaling is not infrastructure, regulation, or talent. It is learning. Most GenAI systems do not retain feedback, adapt to context, or improve over time.
- At the same time, employees in over 90% of companies regularly use personal AI tools for work.

Where does truth matter?



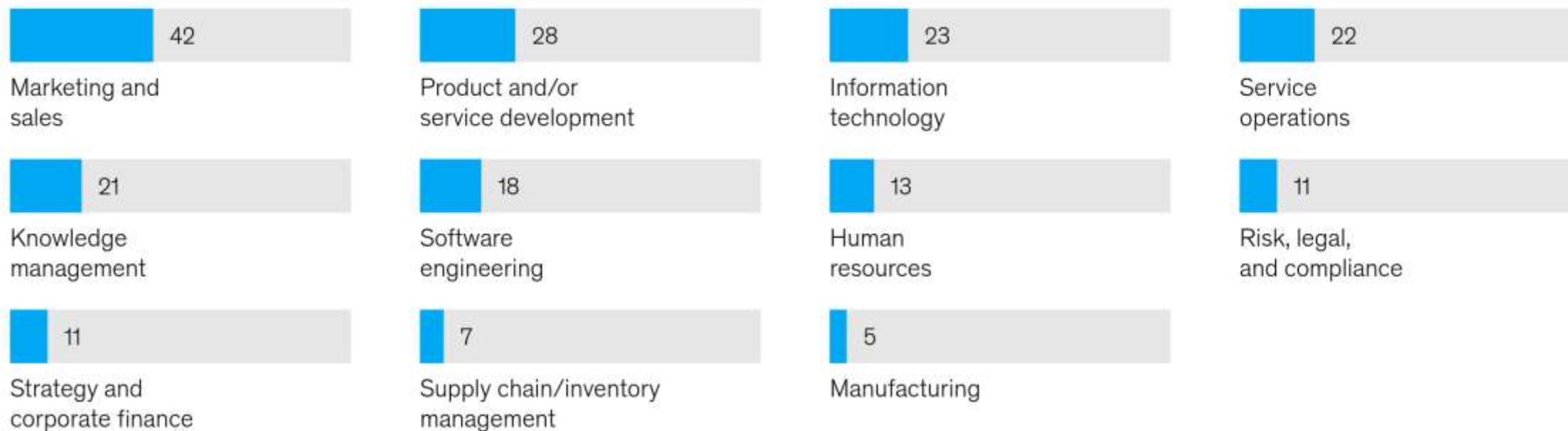
Industry	Key Signals
Technology	New challengers gaining ground (e.g., Cursor vs Copilot); shifts in workflows
Media & Telecom	Rise of AI-native content; shifting ad dynamics; incumbents still growing
Professional Services	Efficiency gains; client delivery remains largely unchanged
Healthcare & Pharma	Documentation/transcription pilots; clinical models unchanged
Consumer & Retail	Support automation; limited impact on loyalty or leaders
Financial Services	Backend automation; customer relationships stable
Advanced Industries	Maintenance pilots; no major supply chain shifts
Energy & Materials	Near-zero adoption; minimal experimentation

Exhibit: GenAI Investment Distribution by Function



McKinsey report, March 2025

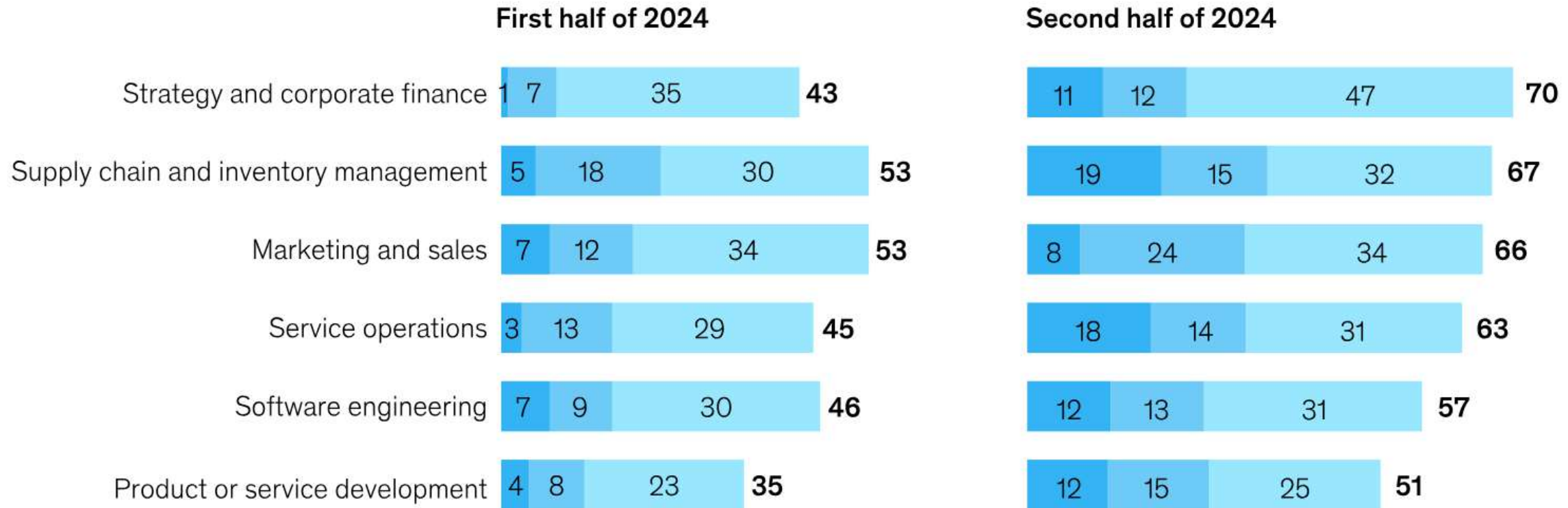
Use of gen AI by business function, % of respondents



Organizations increasingly see gen AI's effects on revenues in the business units using the technology.

Revenue increase within business units from gen AI use, past 12 months, by function,¹ % of respondents

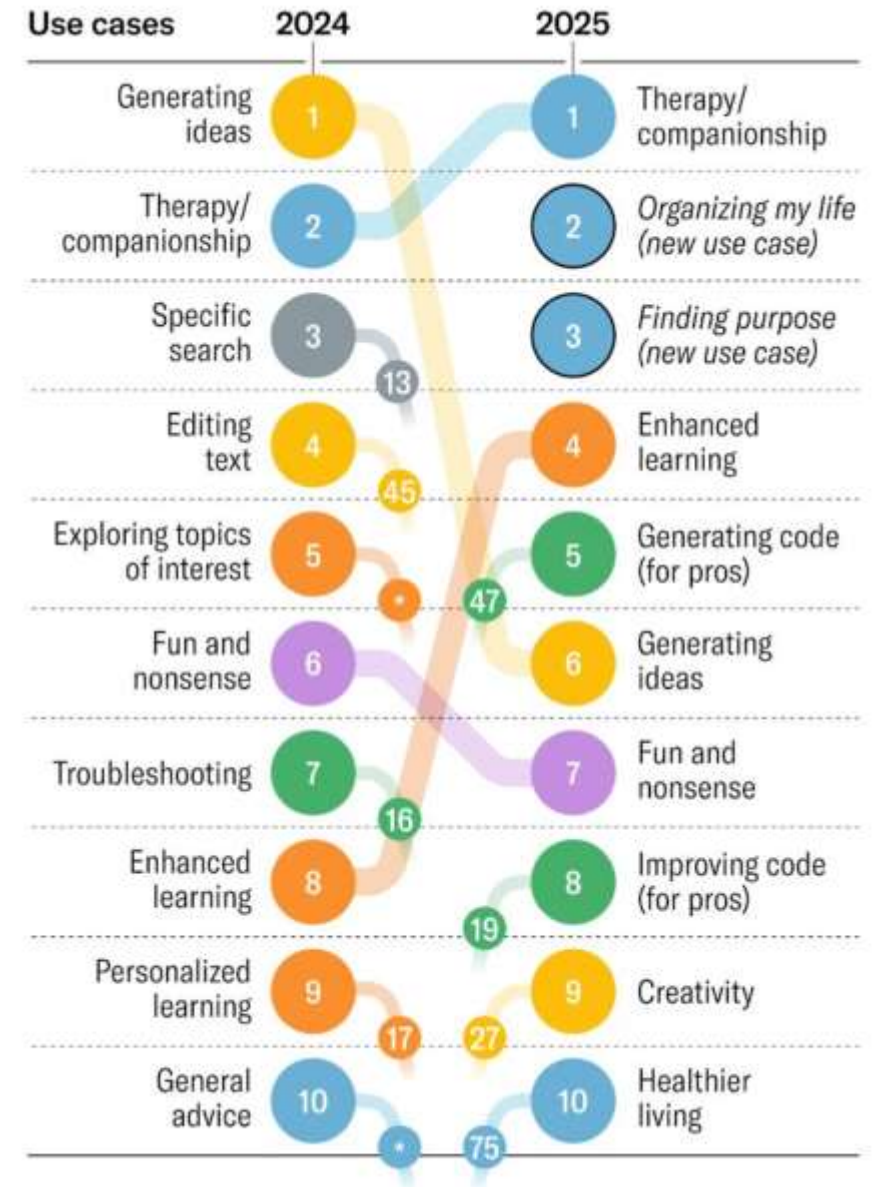
■ Increase by >10% ■ Increase by 6–10% ■ Increase by ≤5%



¹Questions were asked only of respondents who said their organizations regularly use gen AI in a given function. Respondents who said “no change,” “decreased revenue,” “don't know,” and “not applicable,” as well as business functions that are cost centers, are not shown. Segments may not sum to the total shown, because of rounding. The first 2024 survey was in the field from Feb 22 to Mar 5, and the second was fielded from July 16 to July 31.

“How People Are Really Using Gen AI in 2025”:

Marc Zao-Sanders (2025)



Oletused

- A.I. areneb ennustamatutes suundades, aga mitte väga kiirelt.
- Pea kõik “tava-tööd” jäävad nähtavas tulevikus alles.
- Vajalikest A.I. / IT oskustest on väga puudu.
- A.I. võimestab tarkvara arendajaid.
- Nõudlus tarkvara arenduse vastu kasvab.
- Õpetamise struktuuri variatiivsust tasub suurendada.

Näiteks: käivitame integreeritud erialasid lisaks hüperspetsialiseeritutele.

Näiteks: suurendame A.I. doktorantuuri.