

A festő, aki eltünteti a vászonnól a zajt,
avagy a mesterséges intelligenciáról közérthetően

Gáspár Merse Előd, fizikus

tudományos újságíró, társasjáték fejlesztő és bűvész

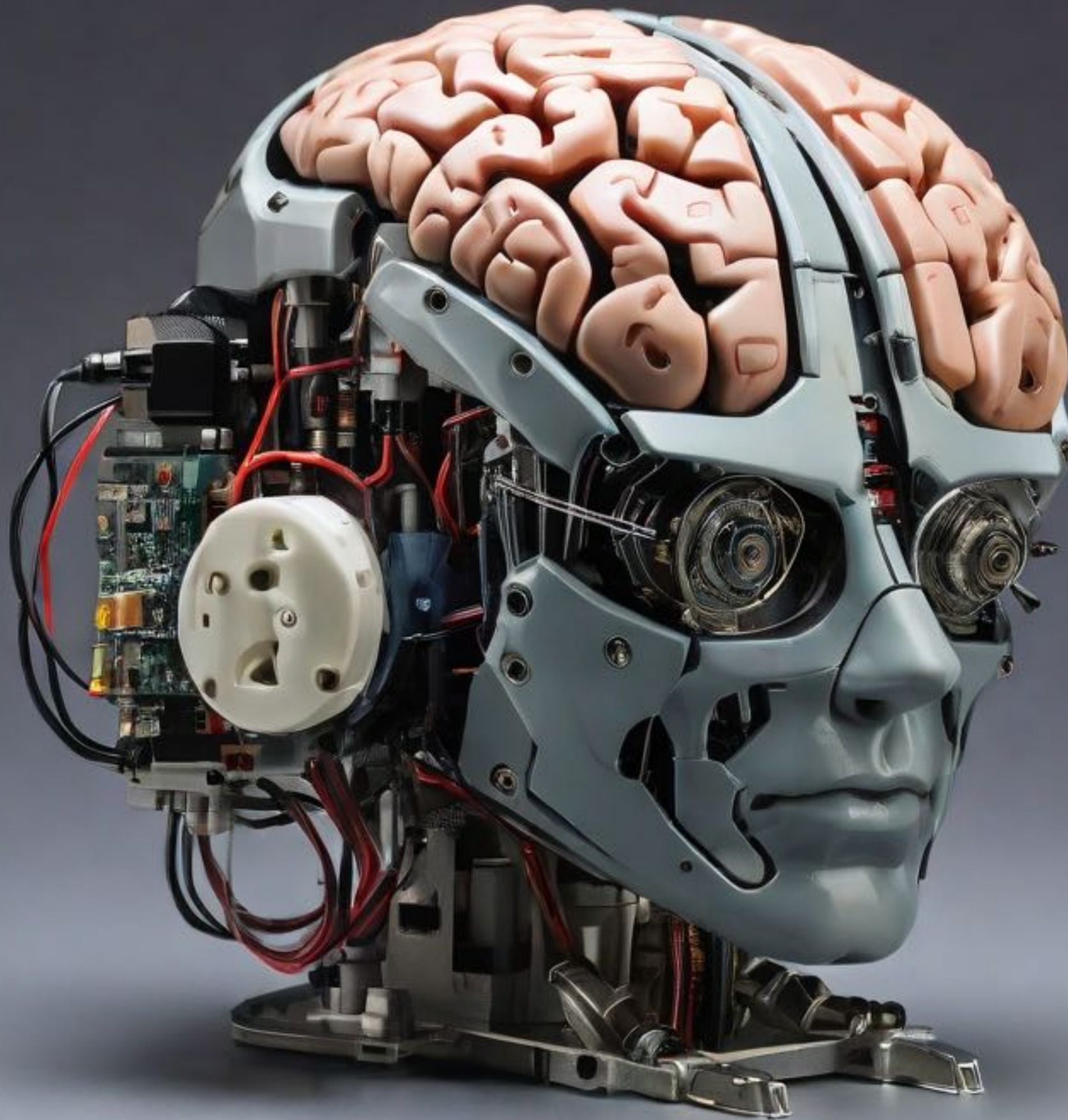
Közép-európai Egyetem Kognitív Tudomány Tanszék

korábbi munkahely: Wigner Fizikai Kutatóintézet

Komputációs Rendszerszintű Idegtudomány Kutatócsoport

2024. március 7.

Agy-számítógép analógia: Mivel foglalkozik egy elméleti idegtudós?

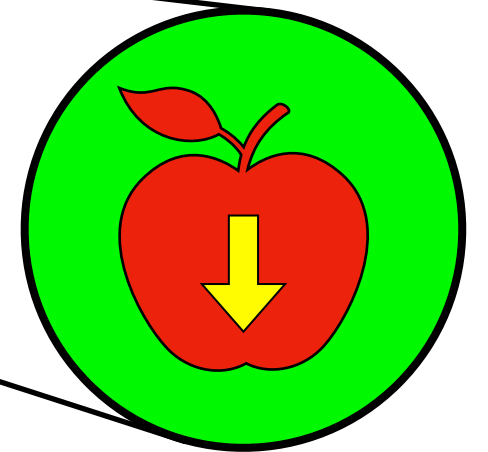
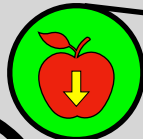
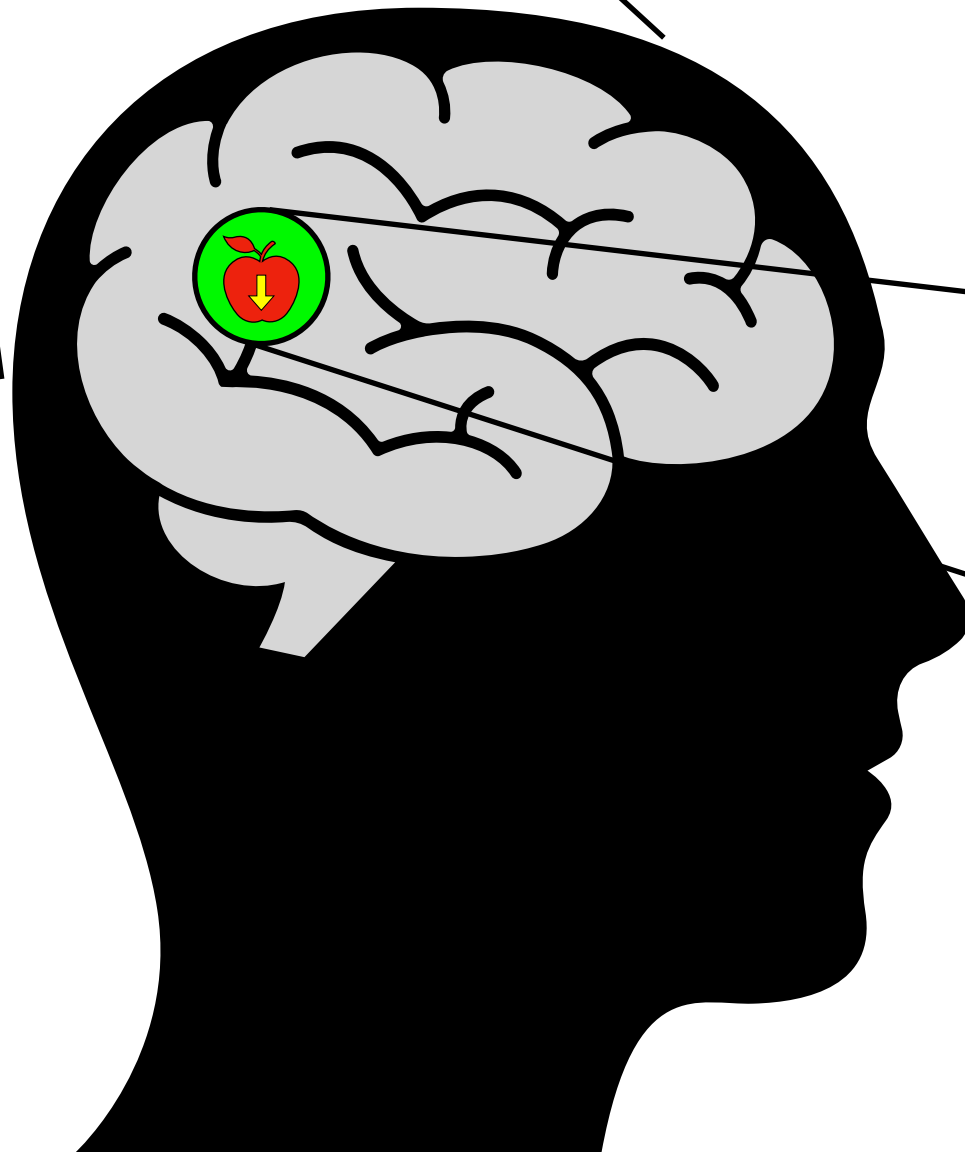




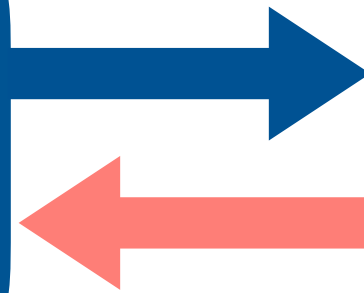
Duplacsavar:

Az agy modellezi a külvilágot.

A kognitív tudós modellezi az agyat.



Elméleti
idegtudomány



Mesterséges
intelligencia

Psychological Review

Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR
INFORMATION STORAGE AND ORGANIZATION
IN THE BRAIN¹

F. ROSENBLATT

Cornell Aeronautical Laboratory

Psychological Review

1981, Vol. 88, No. 2, 135–170

Toward a Modern Theory of Adaptive Networks:
Expectation and Prediction

Richard S. Sutton and Andrew G. Barto
Computer and Information Science Department
University of Massachusetts—Amherst

**Learning representations
by back-propagating errors**

David E. Rumelhart*, Geoffrey E. Hinton†
& Ronald J. Williams*

* **Institute for Cognitive Science**, C-015, University of California,
San Diego, La Jolla, California 92093, USA

† Department of Computer Science, Carnegie-Mellon University,
Pittsburgh, Philadelphia 15213, USA

COGNITIVE SCIENCE 9, 147–169 (1985)

**A Learning Algorithm for
Boltzmann Machines***

DAVID H. ACKLEY

GEOFFREY E. HINTON

*Computer Science Department
Carnegie-Mellon University*

TERRENCE J. SEJNOWSKI

*Biophysics Department
The Johns Hopkins University*

COGNITIVE SCIENCE 14, 179–211 (1990)

Finding Structure in Time

JEFFREY L. ELMAN

University of California, San Diego

Intelligencia

információ feldolgozási szempontból

- információ észlelése és értelmezése
 - összefüggések átlátása, tanulás, azaz tudásként való elraktározás
 - problémamegoldás: különböző környezetben/kontextusban a megfelelő tudás előhívása és alkalmazása
 - még magasabb szinten előre tervezés: döntéshozatal/akciótervezés, stratégia, stb.
-
- több szintje lehet: például emberi és állati intelligencia
 - különböző komponensei lehetnek: például IQ, EQ, RQ
 - belső támogatás: memória, számítási kapacitás, stb.
 - külső támogatás például az embernél: társadalom, írástudás, tudományok, stb.

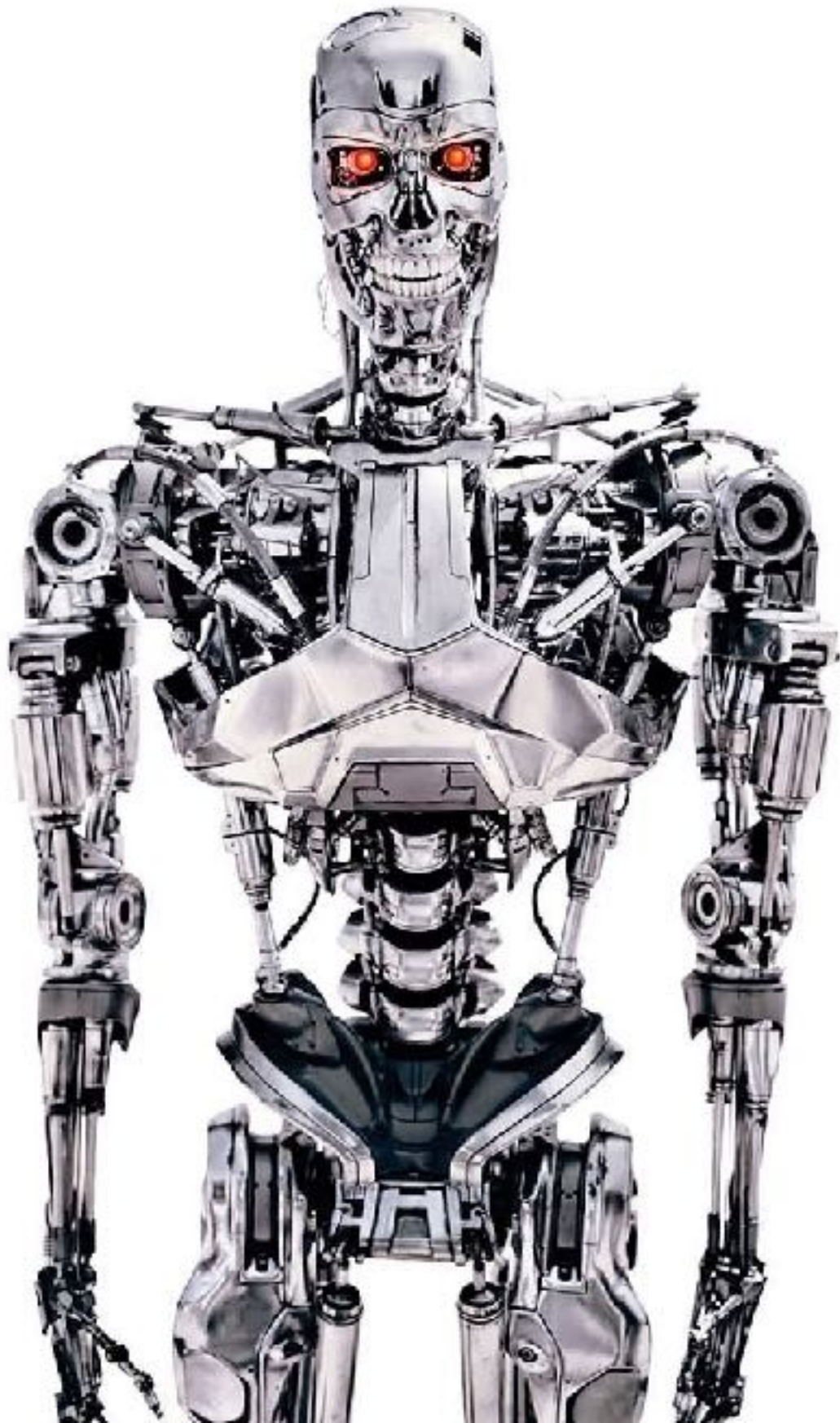


MI-vel kapcsolatos fogalmak

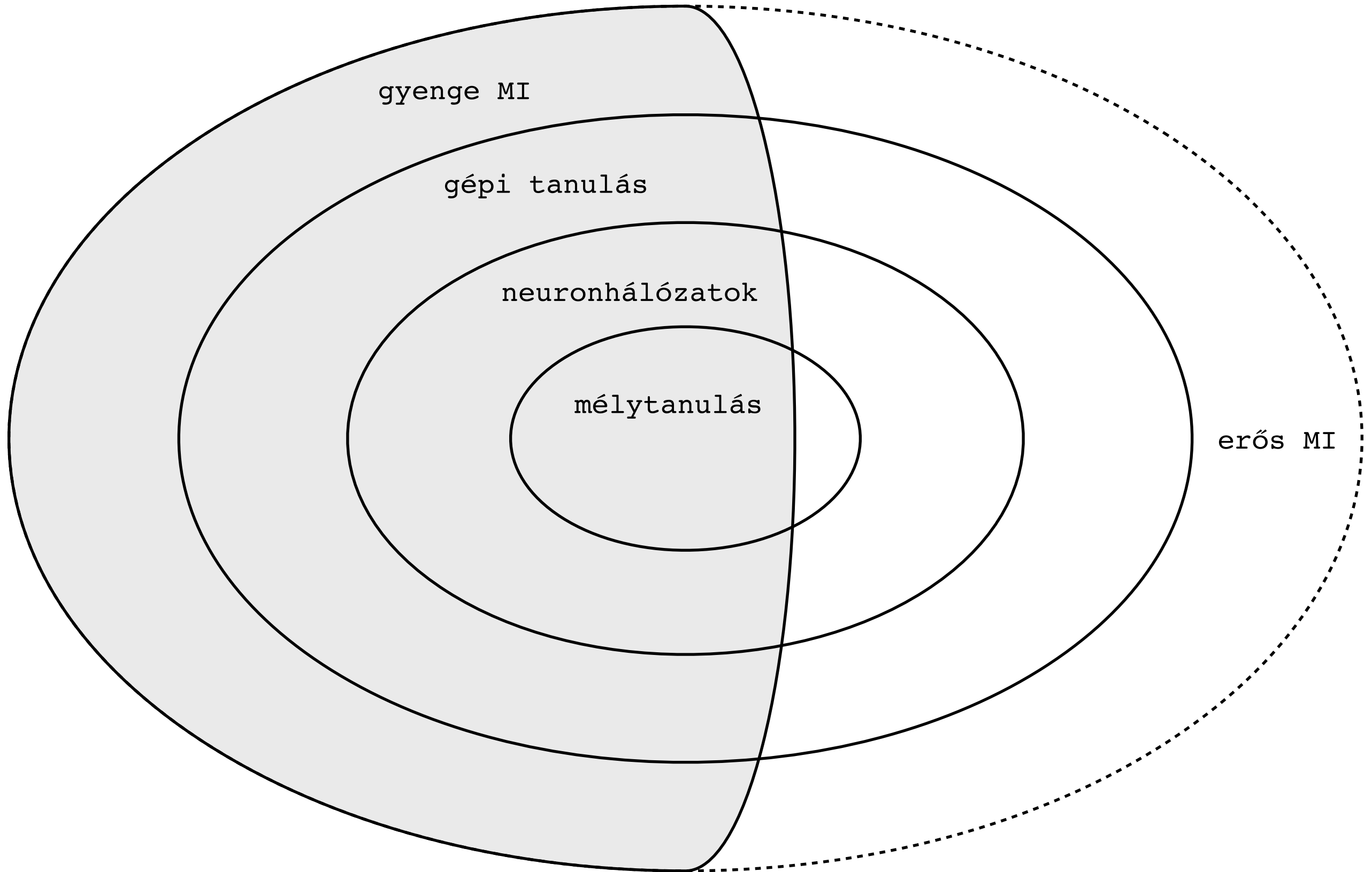
gyenge MI
specifikus MI
weak AI
narrow AI

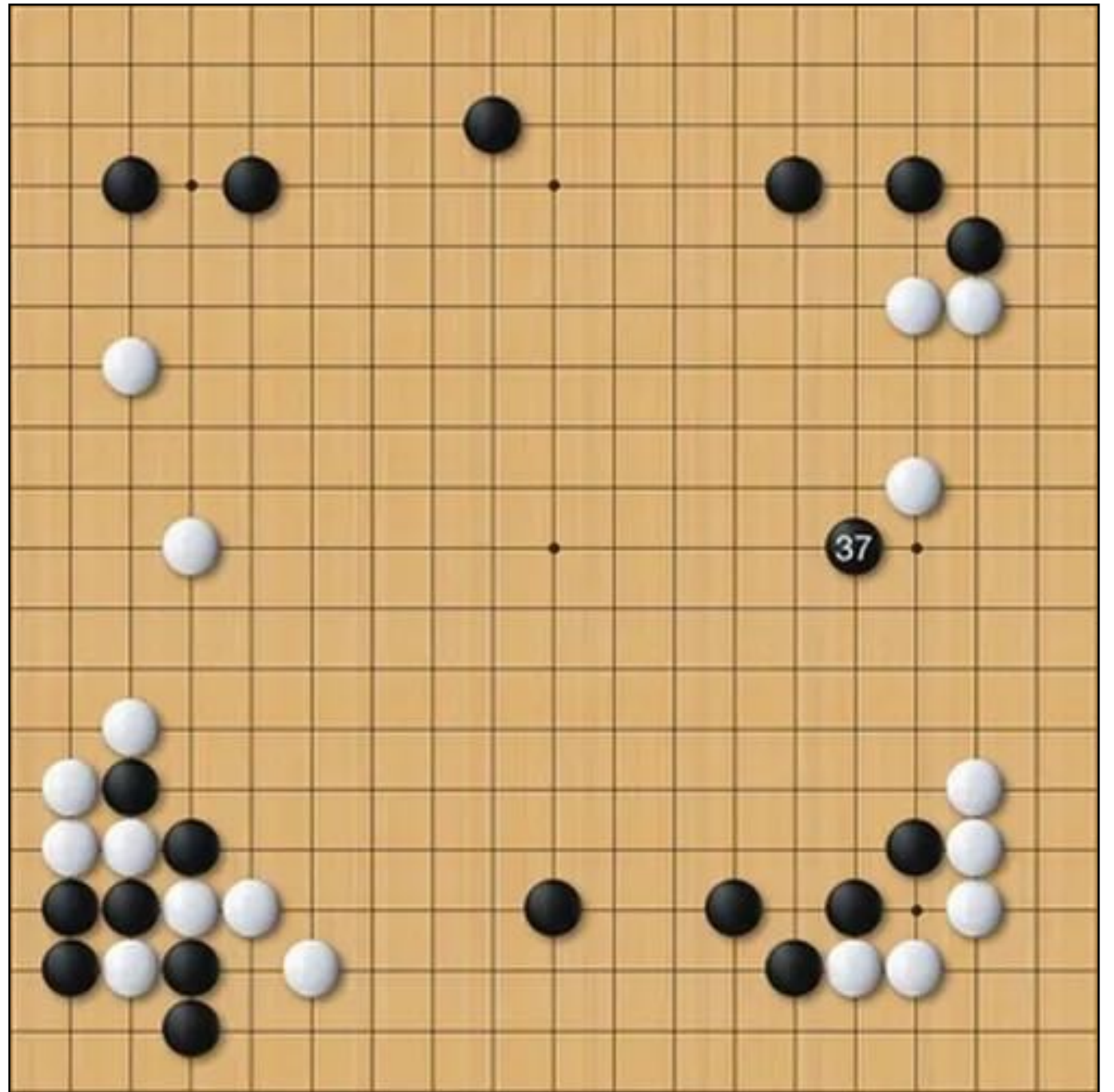
erős MI
általános MI
emberi szintű MI
strong AI
AGI

AGI = artificial general intelligence



MI-vel kapcsolatos fogalmak





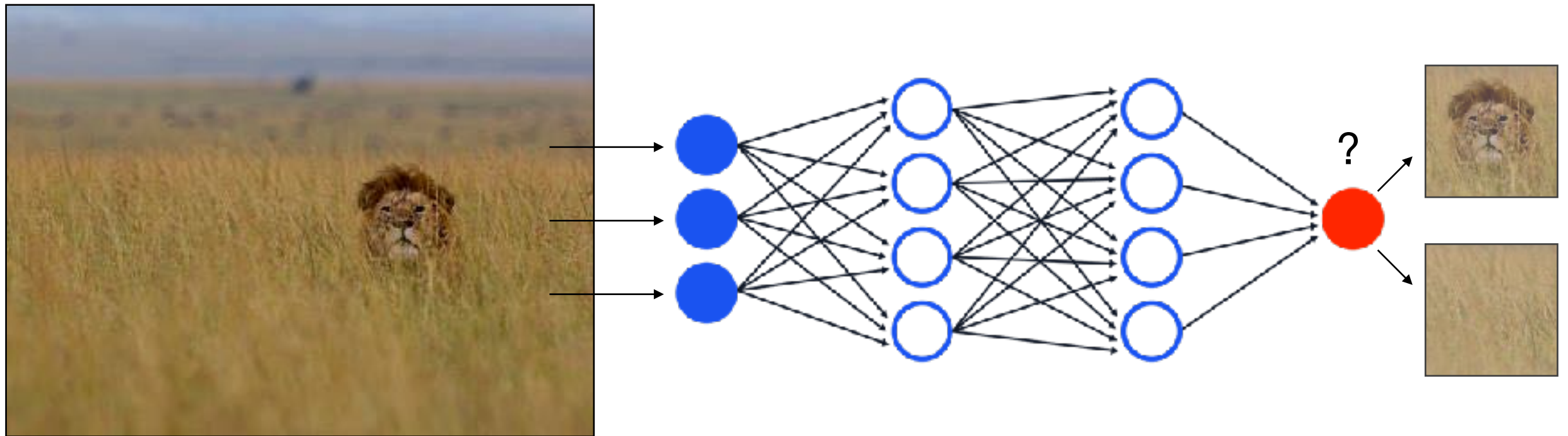
go



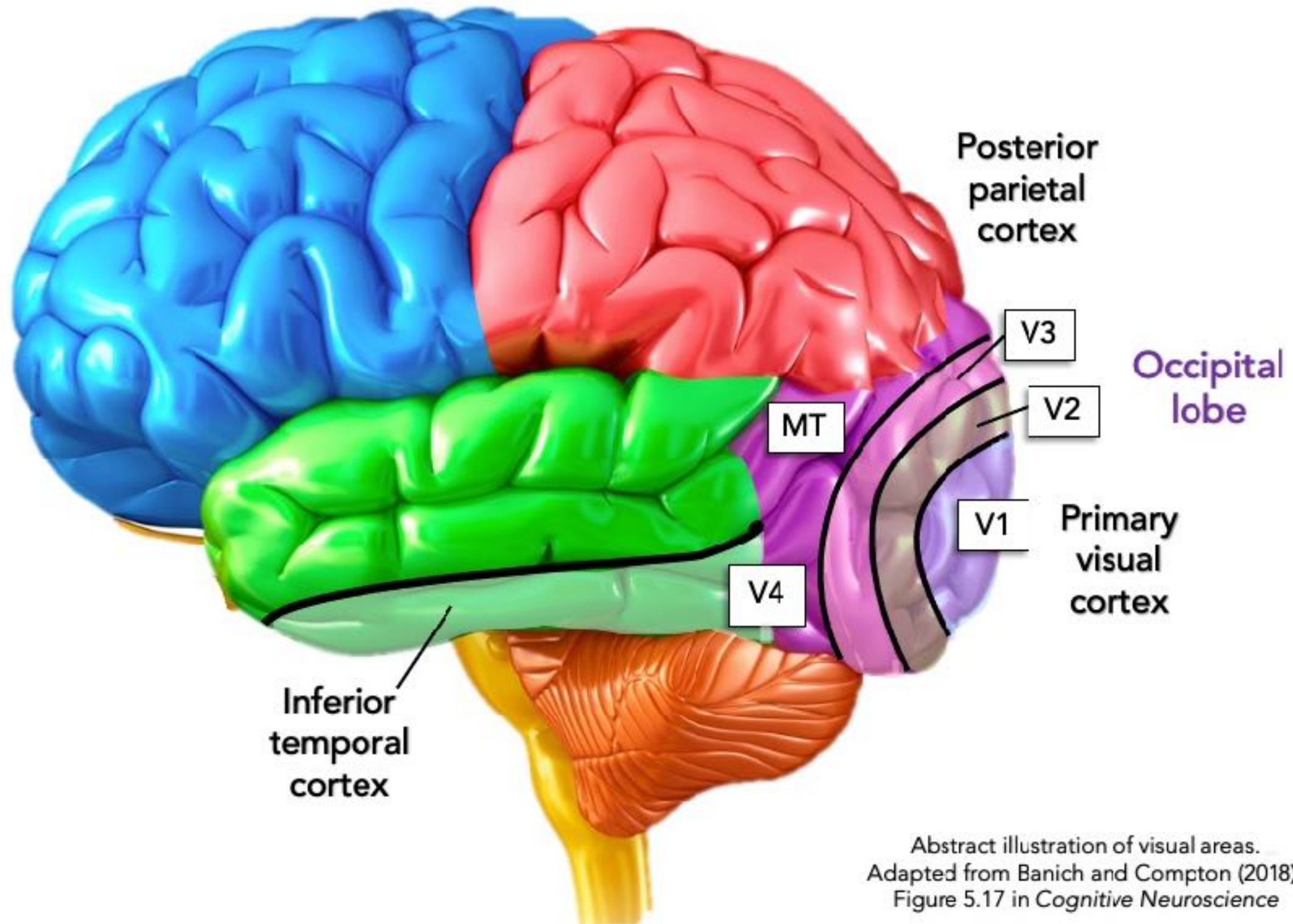
Neuronhálózatok

lényegében függvény:

bemenet → kimenet

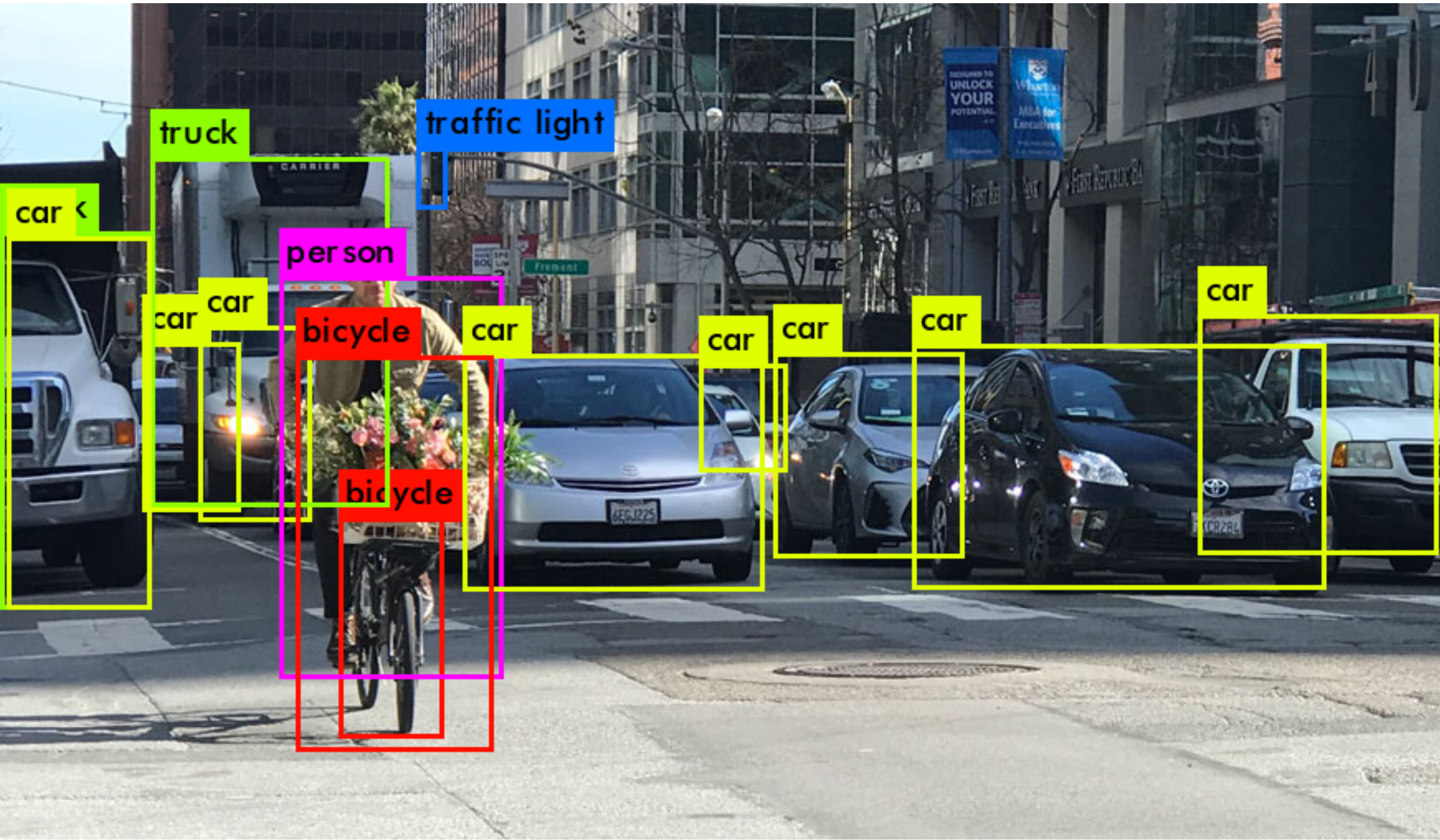


Rétegekbe szerveződés



Felügyelt gépi tanulás

címkézett adatok



Mély/többrétegű neuronháló

Backpropagation



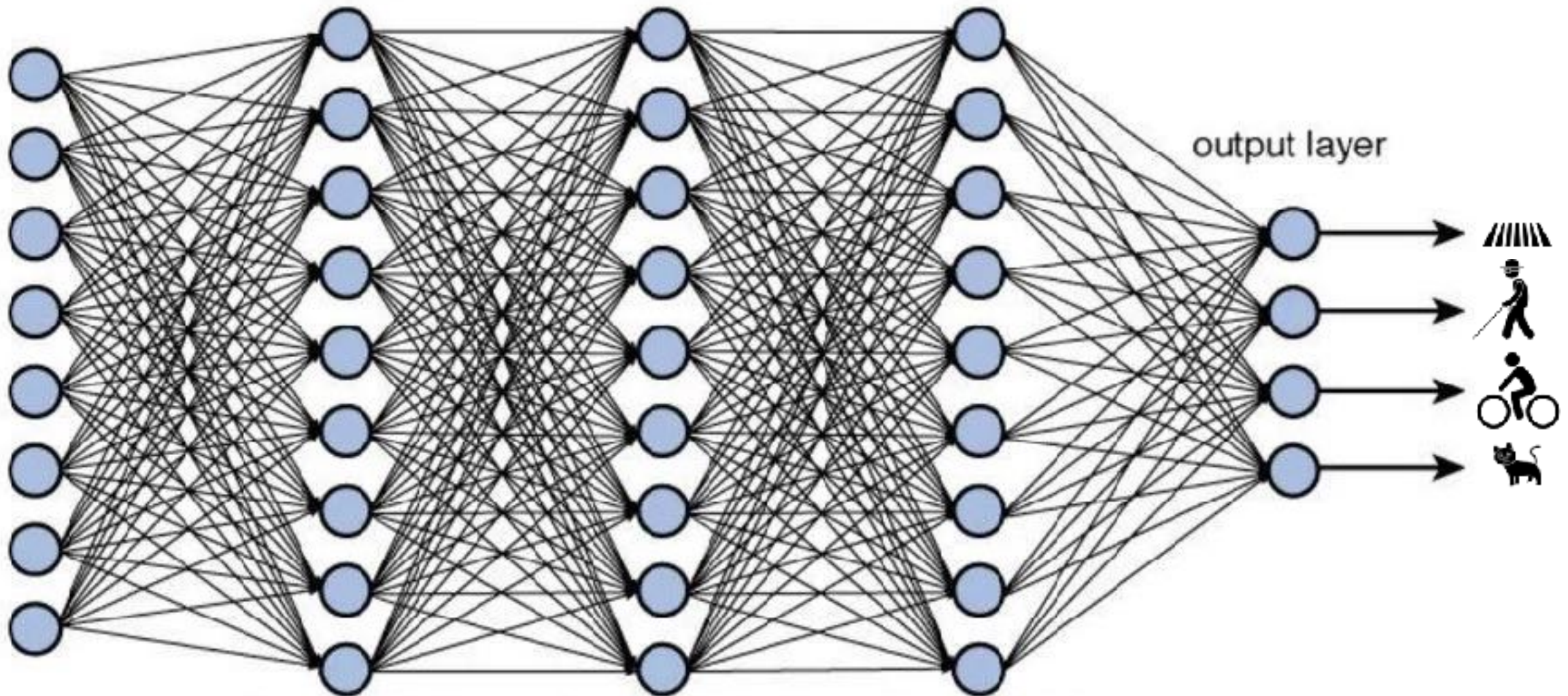
input layer

hidden layer 1

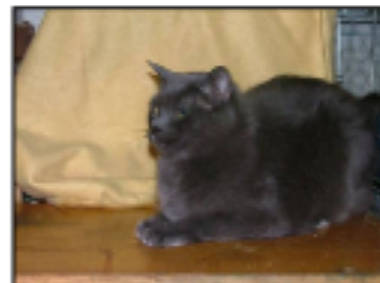
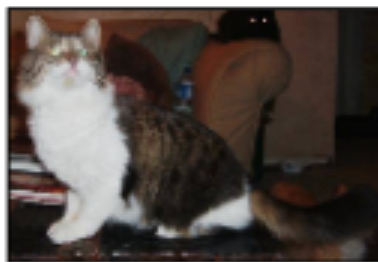
hidden layer 2

hidden layer 3

output layer



Tanító képek sok & reprezentatív

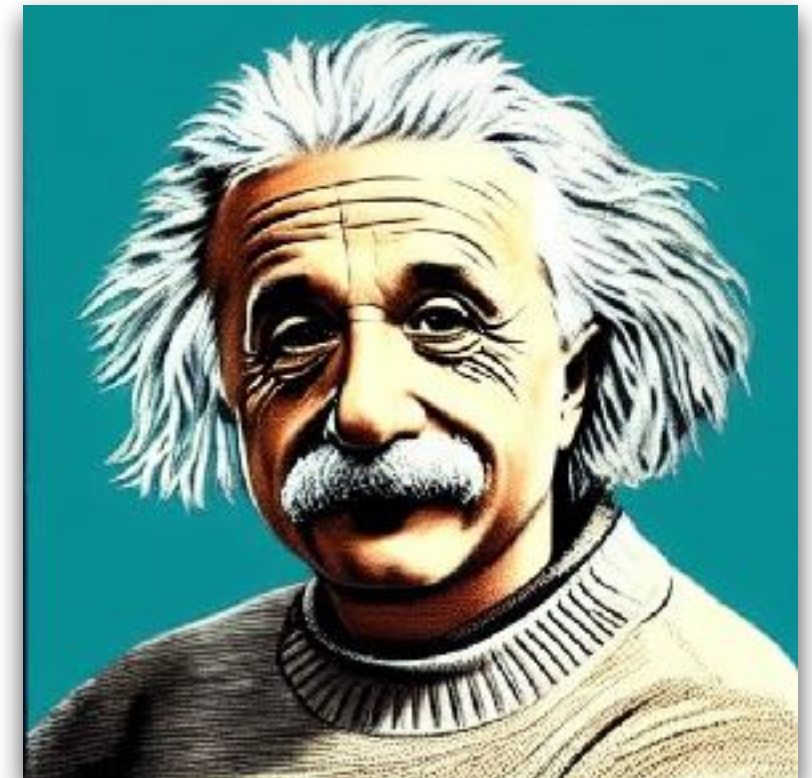
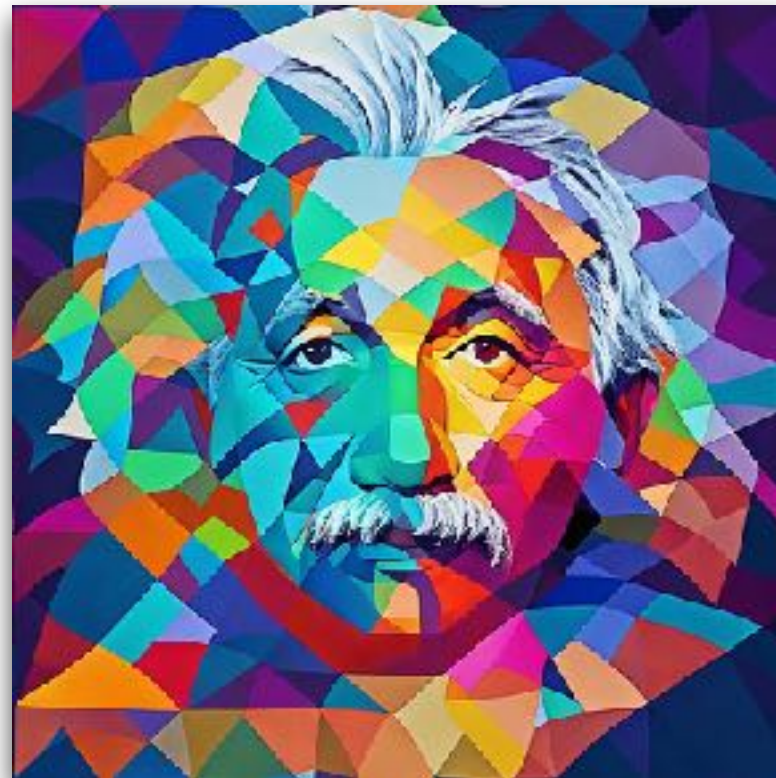
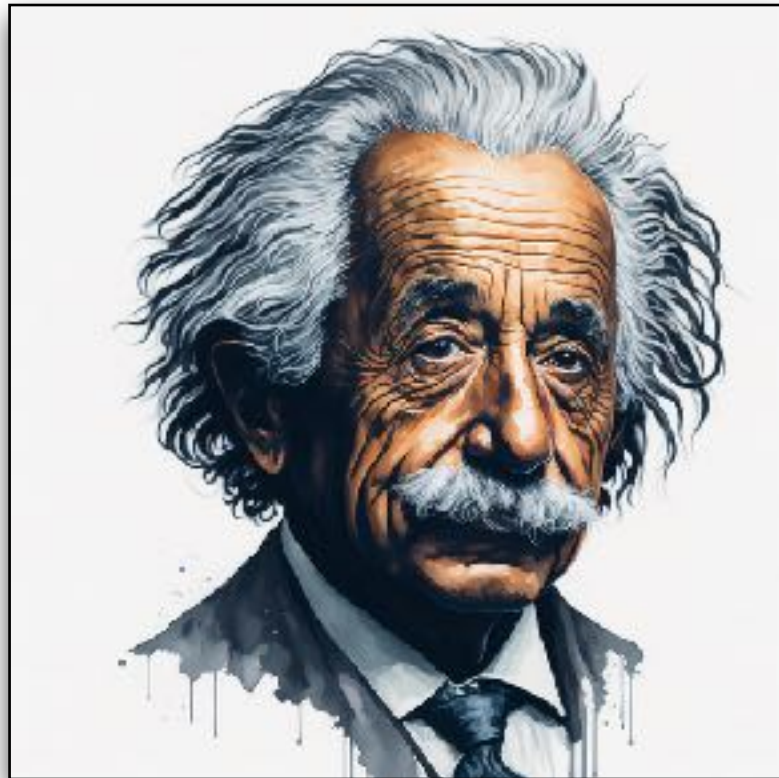
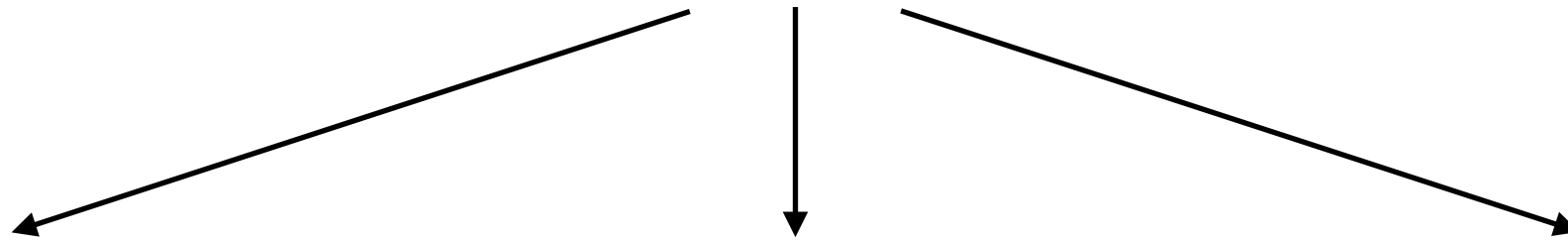


Generatív modellek

prompt → valami

sztochaszticitás

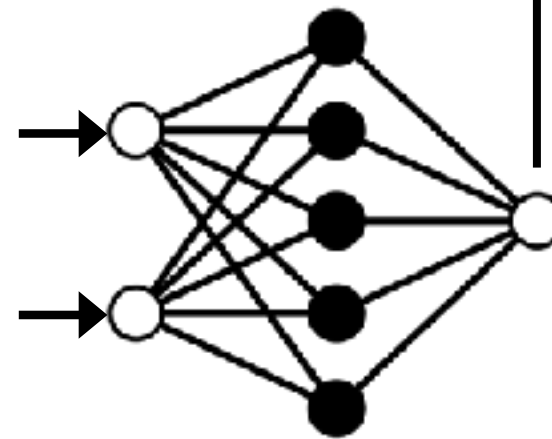
[portrait of Einstein]



Generatív modellek

semmi \rightarrow valami

GAN (Generative Adversarial Networks)



Generator

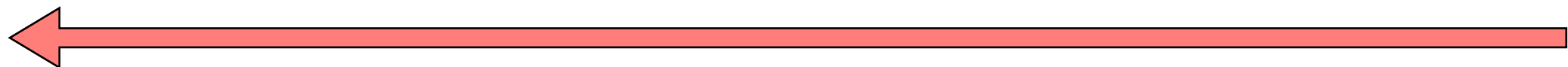
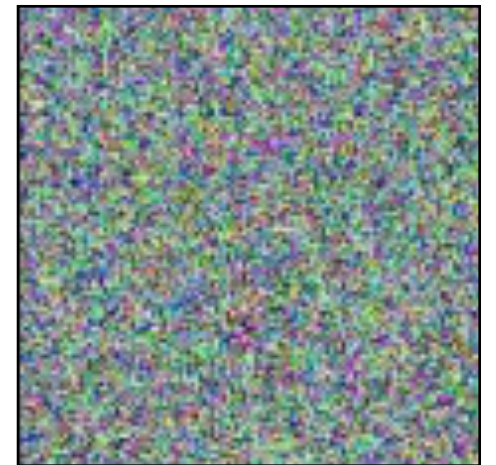
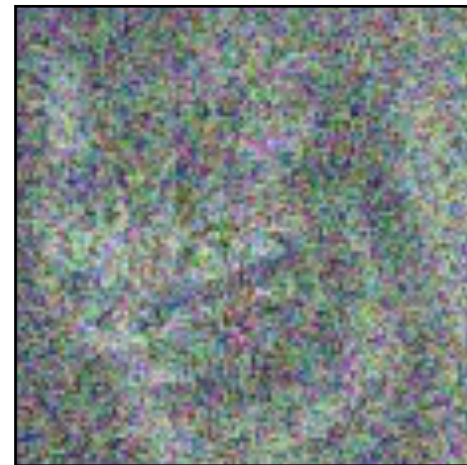
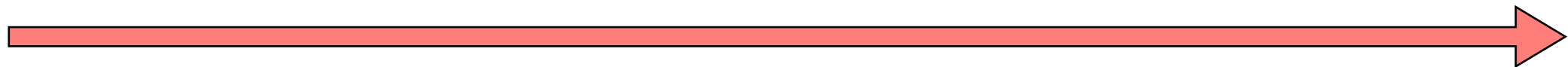
Discriminator

Model Name	Date opened to public	Original Authors	Licensing
DALL-E	January 2021	OpenAI	Proprietary
Imagen	Not public	Google	Proprietary
DALL-E2	July 2022	OpenAI	Proprietary
Craiyon (aka DALL-E Mini)	July 2022	Craiyon	Apache 2.0
Midjourney	July 2022	Midourney	Proprietary
Stable Diffusion	August 2022	LMU Munich, Runway, StabilityAI, EleutherAI, LAION	Creative ML OpenRAIL-M

Diffúziós modellek

festés = eltüntetjük a zajt a vászonról!

diffúzió: zaj lépésenkénti hozzáadása



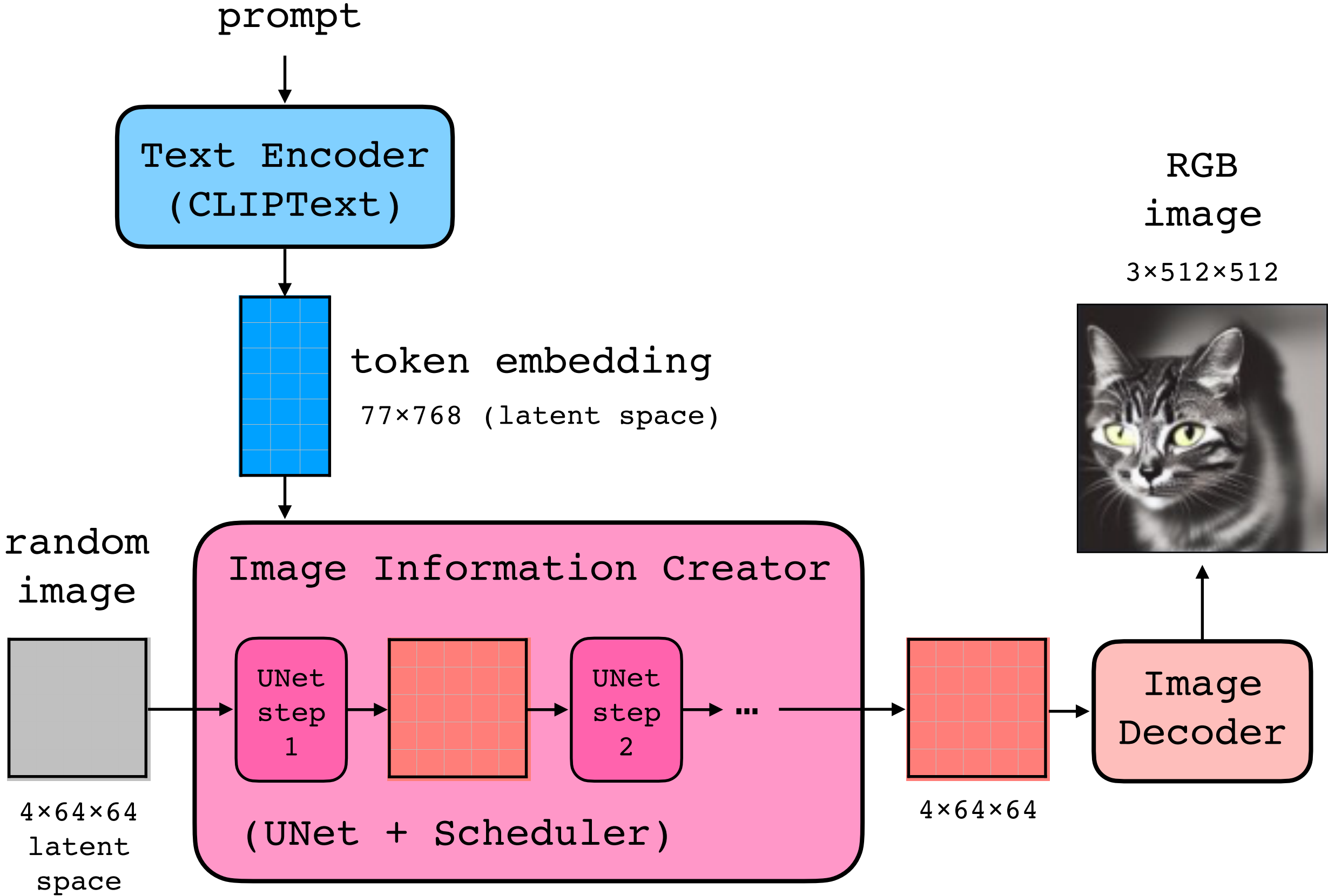
inverz diffúziós folyamat

Zajtalanítás/konvergencia

köztes lépések

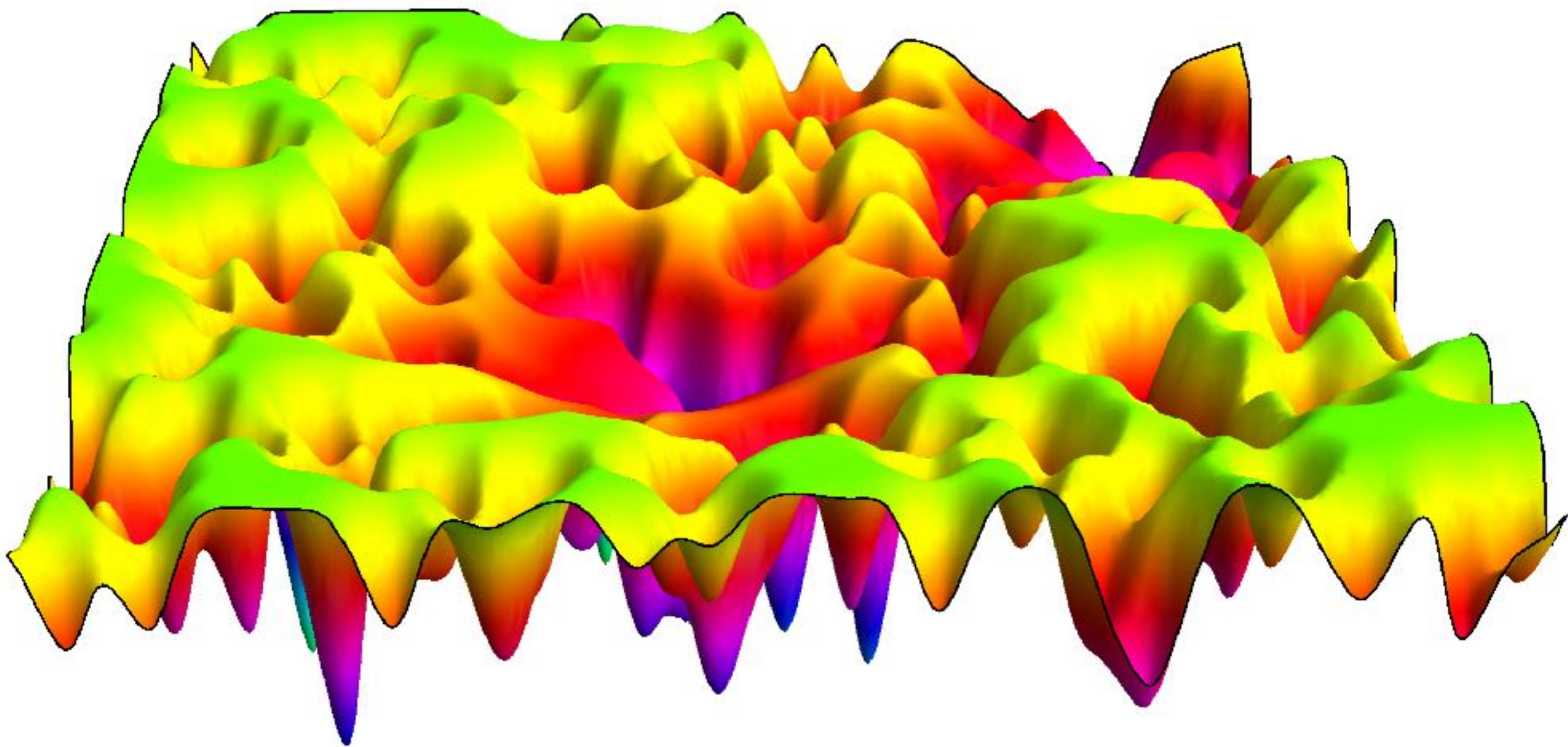


Stable Diffusion 1.5



Analógia

lokális minimum keresése



Karakterábrázolás

random seed használatával



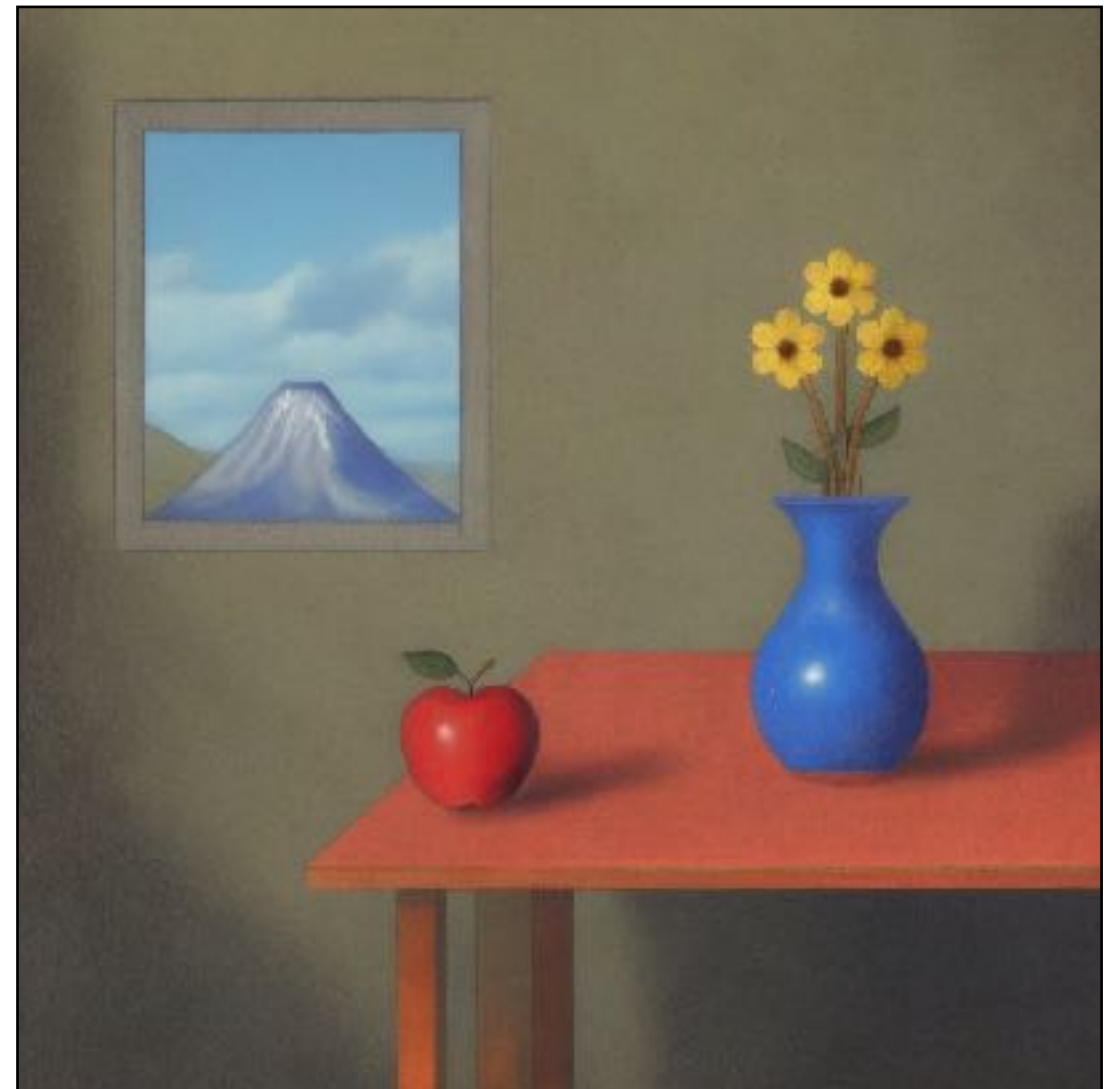
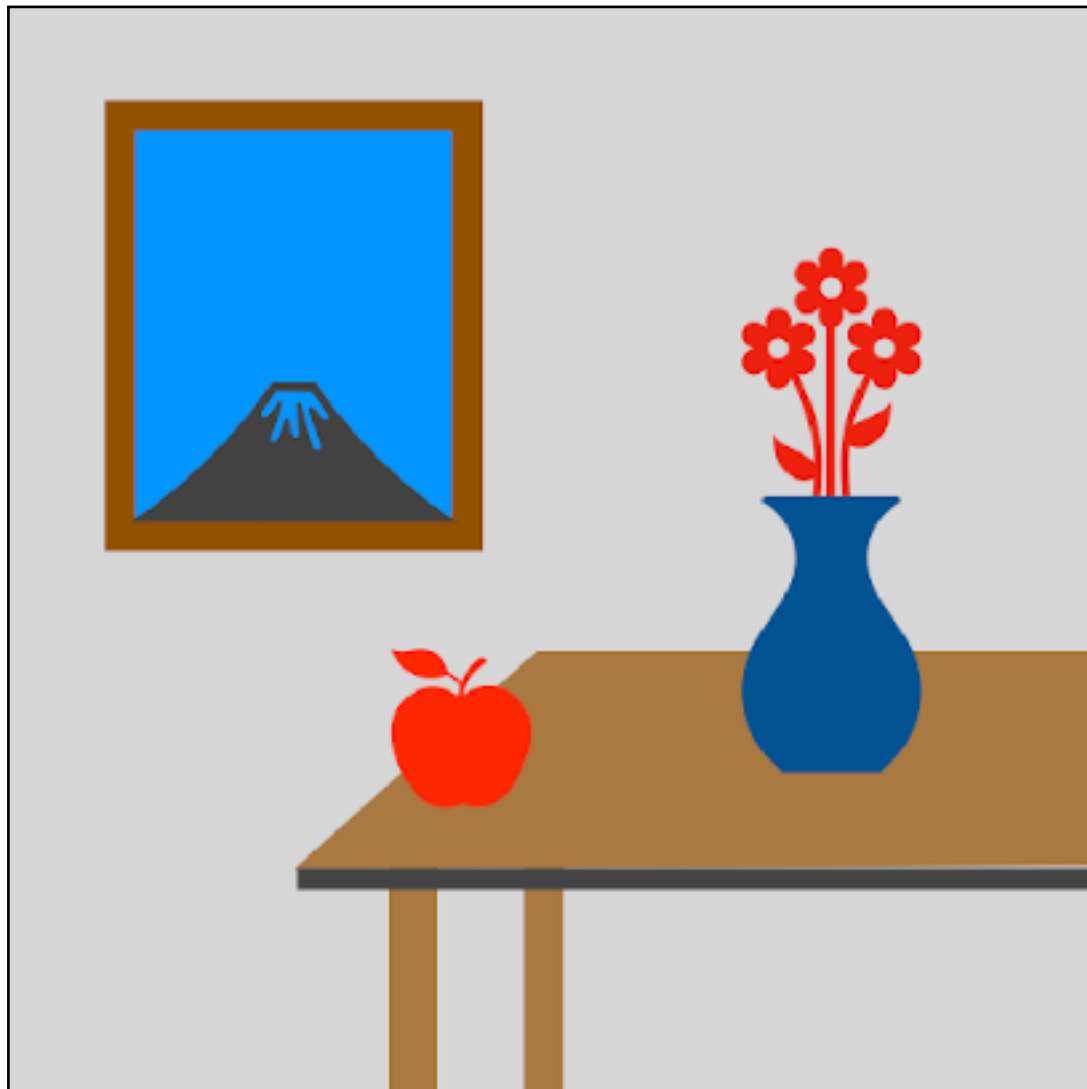
ControlNET: Edge to Image

fotórealisztikus stílus



ControlNET: Edge to Image

festmény stílus



ControlNET: Edge to Image

Példa: ujjlenyomatból labirintus

Prompt: hedge maze surrounded by green dense forest, top view of a labyrinth



ControlNET: Depth to Image

példa: Rubik-kockából csokitorta

Prompt: chocolate cake on a plate, realistic macro photo with blurred colour background



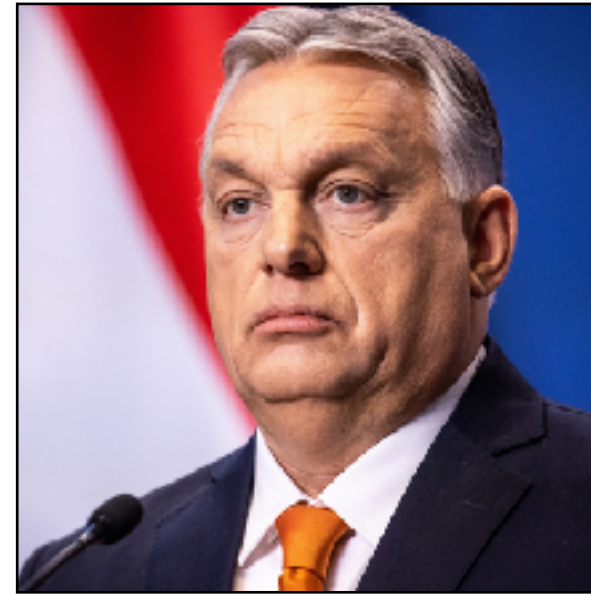
ControlNET: Pose to Image

Példa: egy sziluett is elég hozzá!



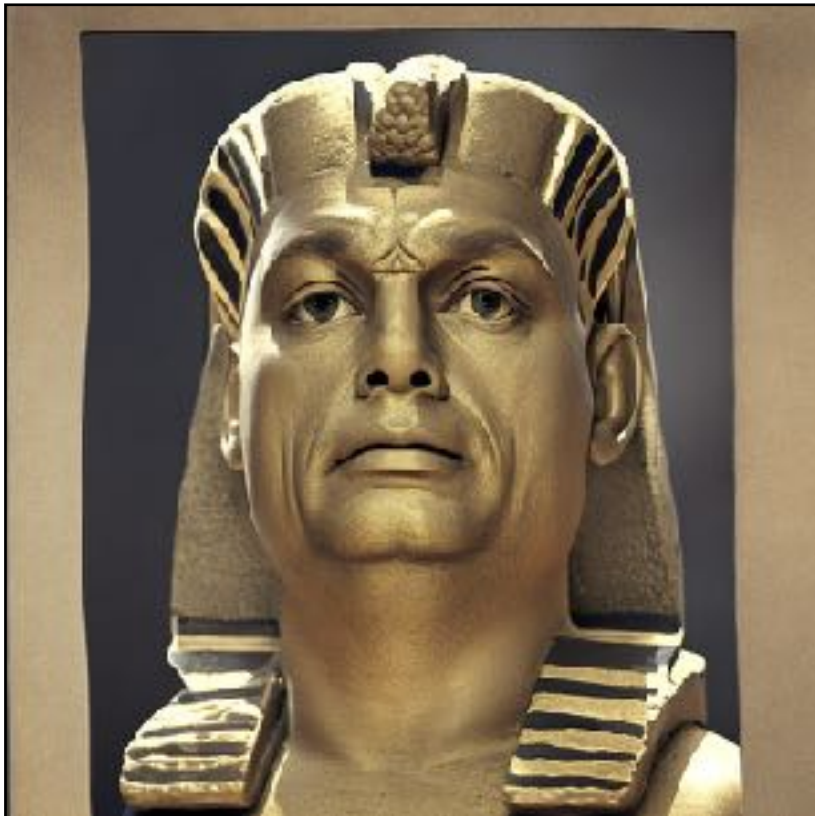
Saját modellek tréningezése

Már néhány kép elegendő!



8 db tréning kép

generált példák:



Háttér marad

kártya közepének törlése



Tévhitek és félreértések

1. Mit tud és mit nem az MI?
2. Mennyire kreatív az MI?

1. Mit tud és mit nem az MI?

Nincs enciklopédikus tudása!



Ez így nem túl informatív kódolás!

1 ← a

2 ← abajgat

3 ← abakusz

4 ← abál

5 ← abbahagy

6 ← abbamarad

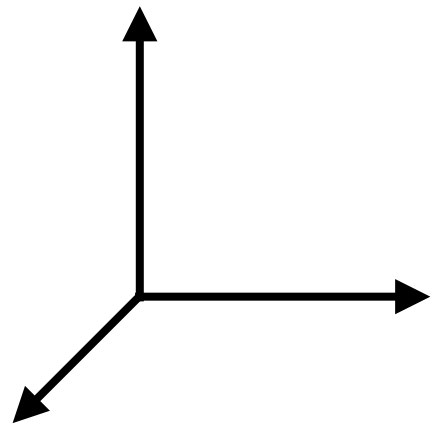
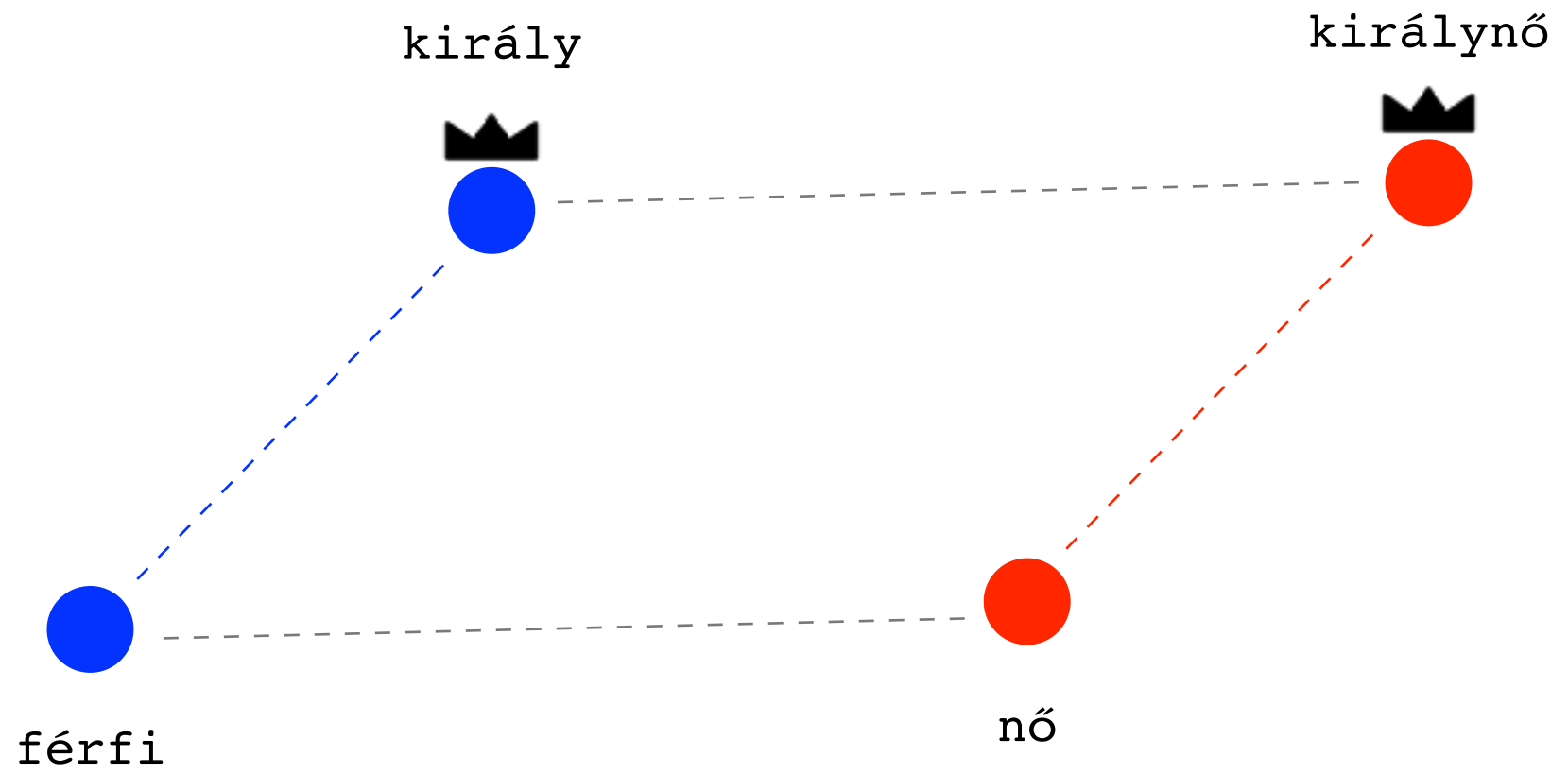
7 ← ábécé

8 ← ablak

9 ← ábra

Word embedding

csodás következmények



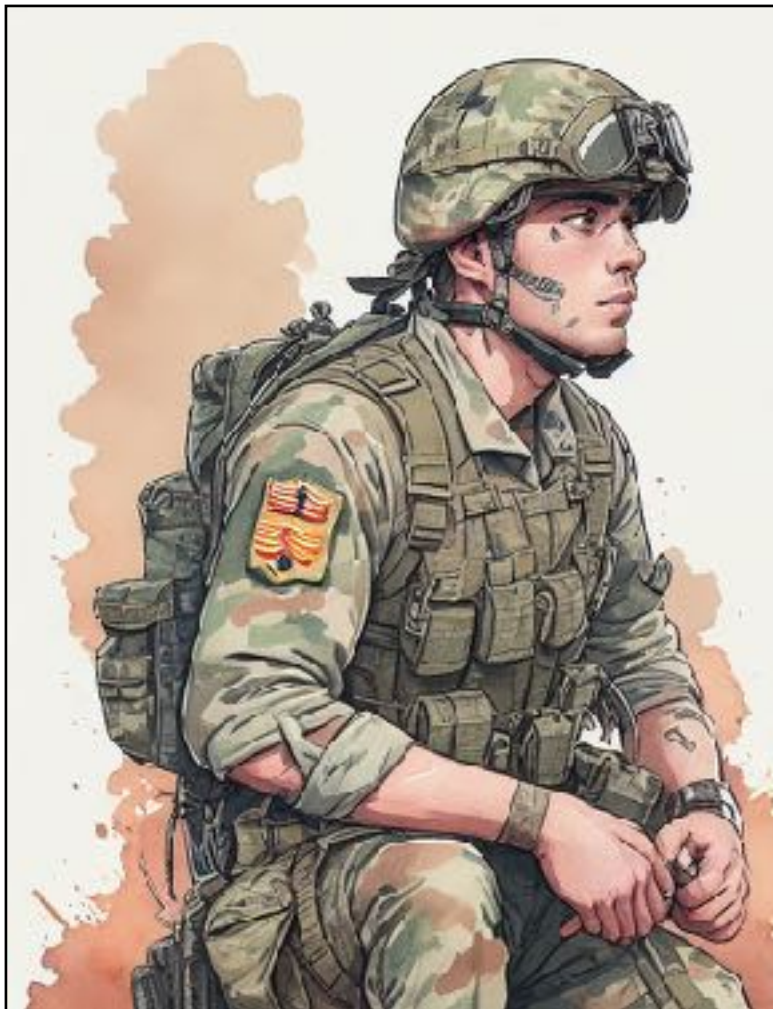
Kitérő: Mi is csak statisztikai alapon működünk?

1. példa: olvasás

Egy anligai etegyem ktuasátaí szenirt nem száímt, melyin serenrodbn vnanak a bteúk egy szbóan, az etegyeln ftonos dloog, hogy az esló és az ultosó bteúk a hölyeükn lneegyek. A tböbi bteú lheet tljees összevaisszásbagn, mgiés porbléma nlkéül oalvsaható a szevög.

Kitérő: Mi is csak statisztikai alapon működünk?

2. példa: vizuális részletek



2. Mennyire kreatív az MI?

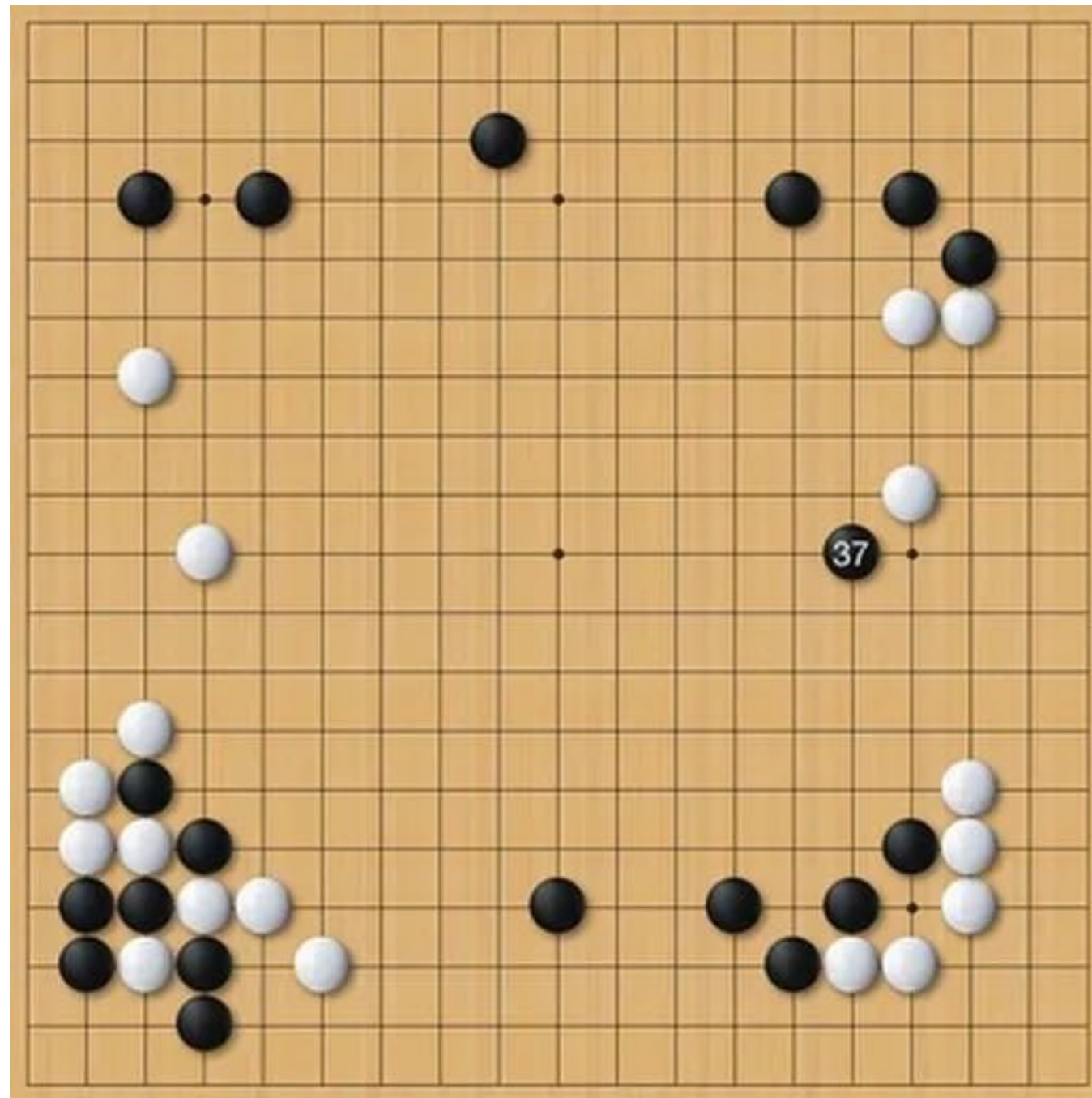


<https://www.youtube.com/live/qVIGJQlrywM&t=1108>

2. Mennyire kreatív az MI?



2. Mennyire kreatív az MI?



MI forradalom okai

Adatbőség: internet, okoseszközök, digitális lábnyomok, globális cégek, stb.

megnövekedett számítógépes kapacitás: elérhető árú GPU-k (párhuzamosítható algoritmusokkal)

algoritmusok fejlődése: mély neurális hálók, attention mechanizmus, stb.

hatékony nyelvi kommunikáció: nagy nyelvi modellek, hétköznapi nyelv használata, stb.

Kitekintés a közeljövőbe

Adattisztaság

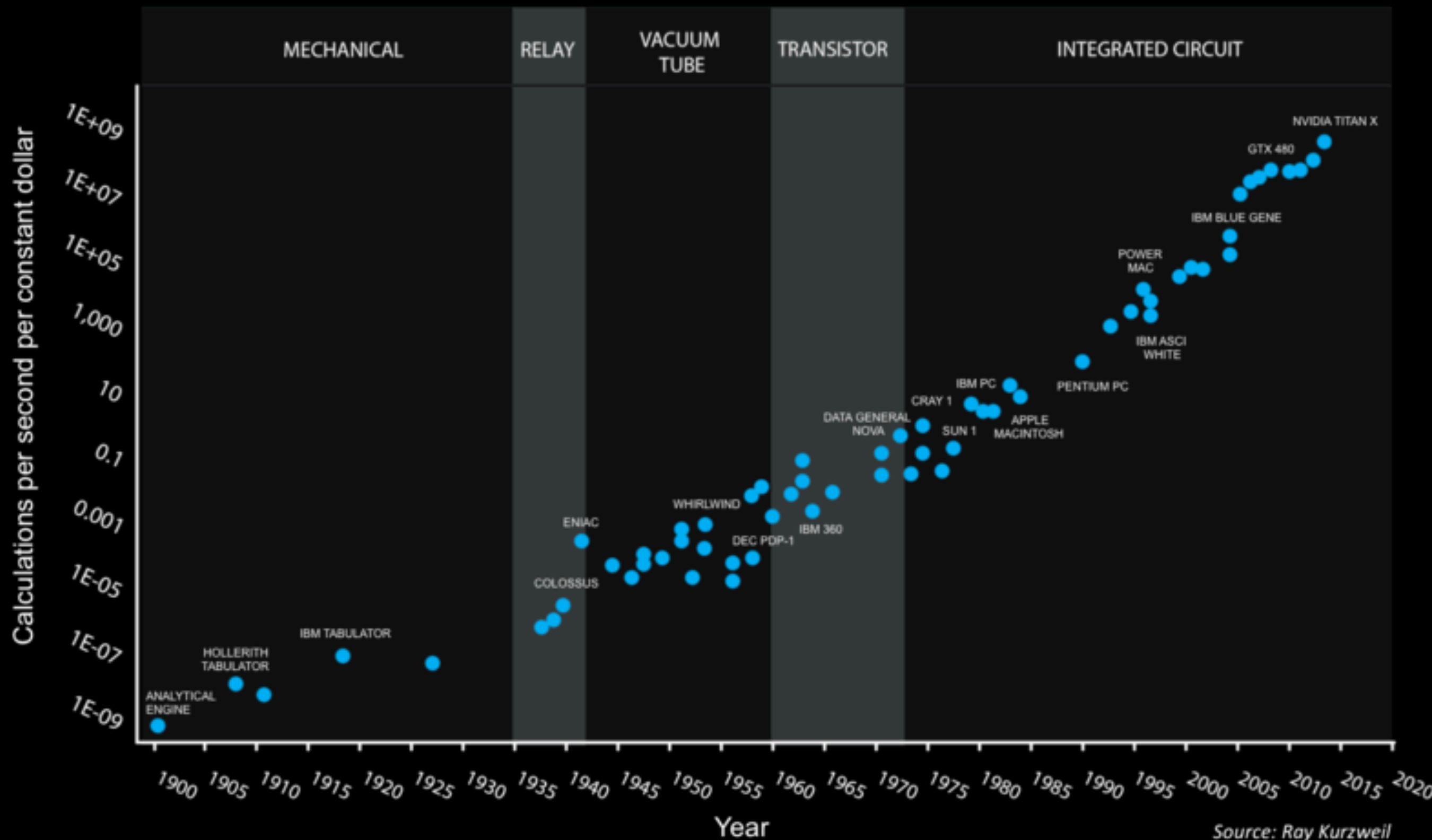
Hitelesség és megbízhatóság

Személyre szabott MI

Társadalmak, demokráciák, befolyásolás

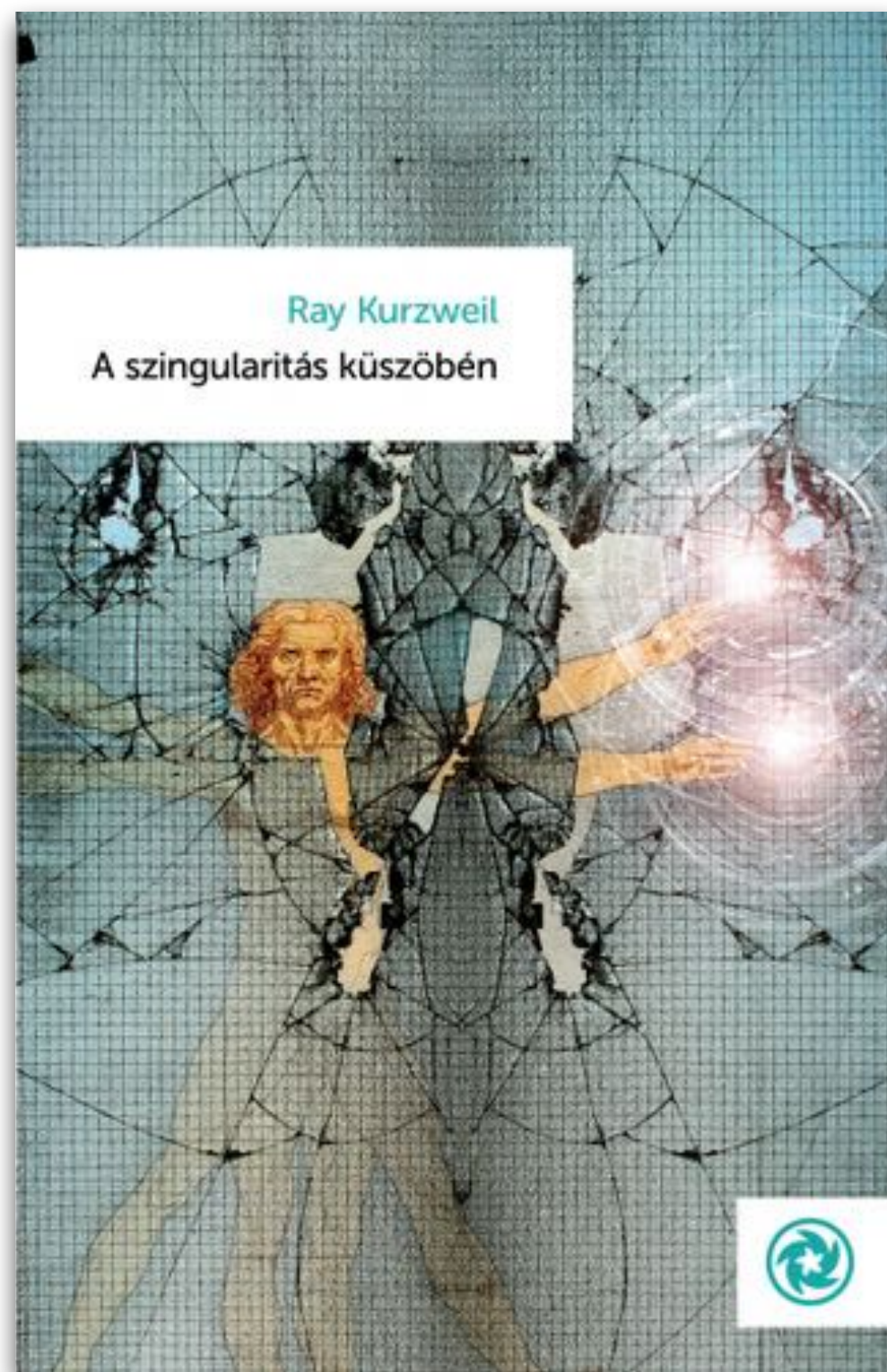
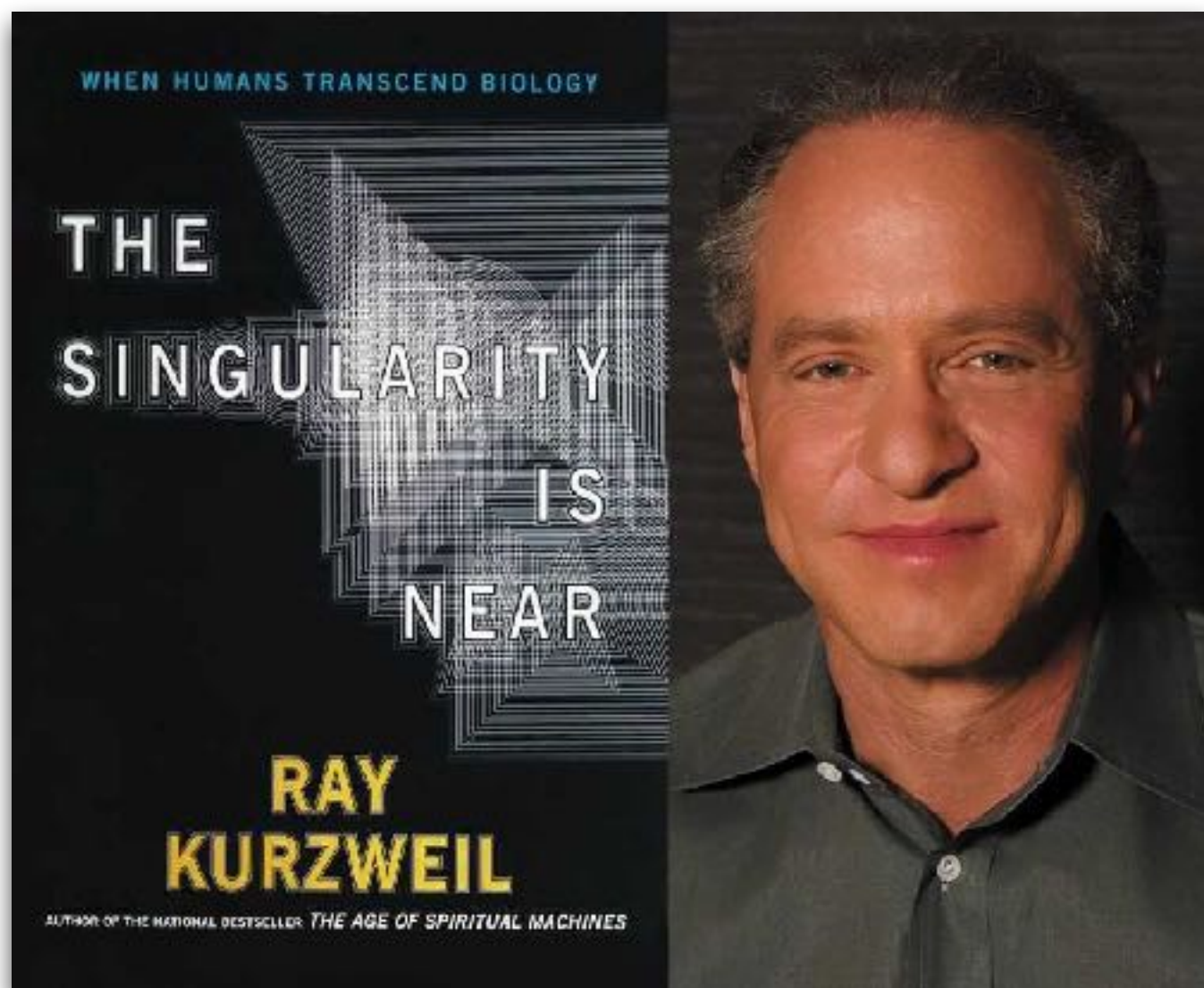
Kvantumkomputerek és neuroinformatika

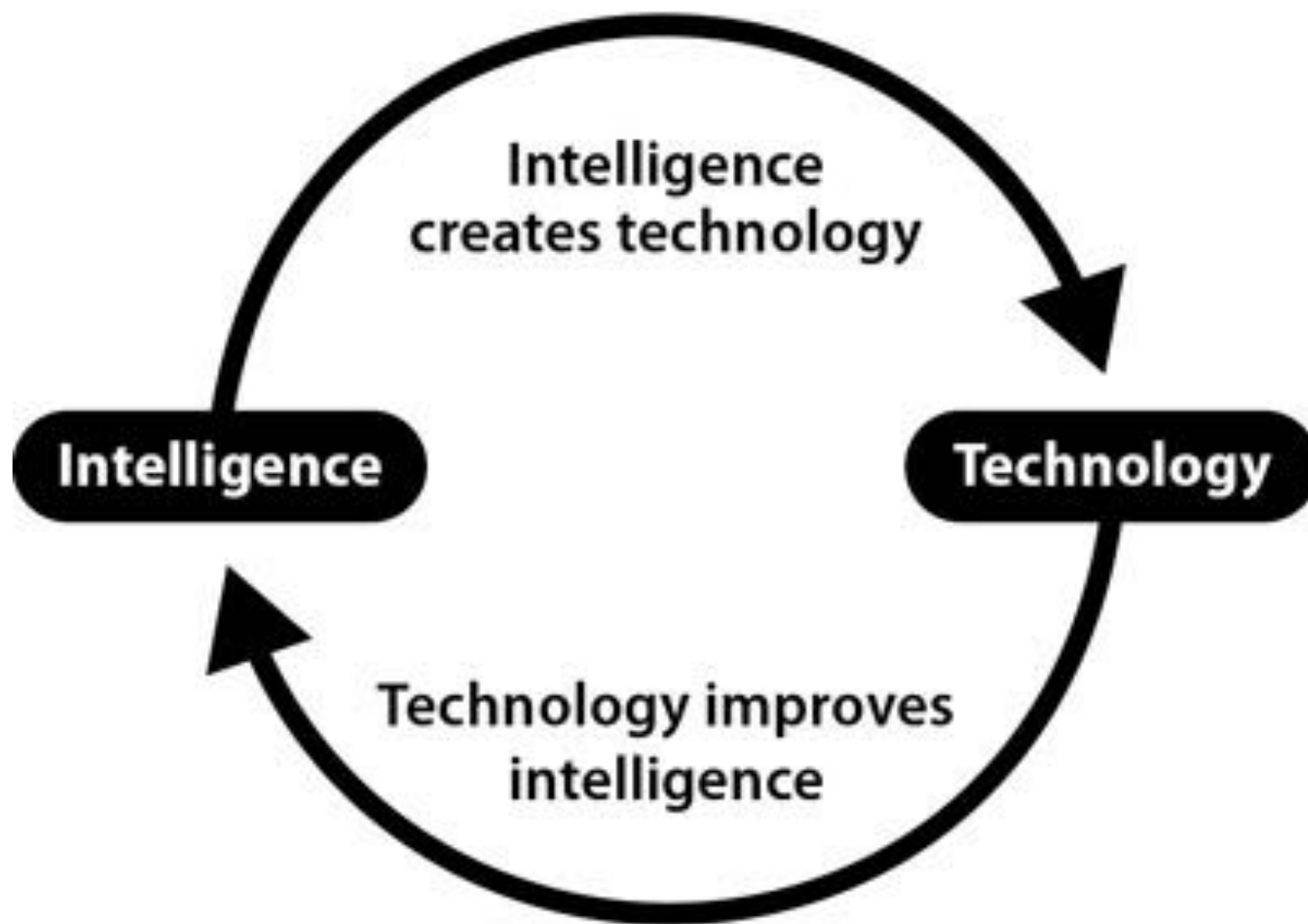
120 Years of Moore's Law



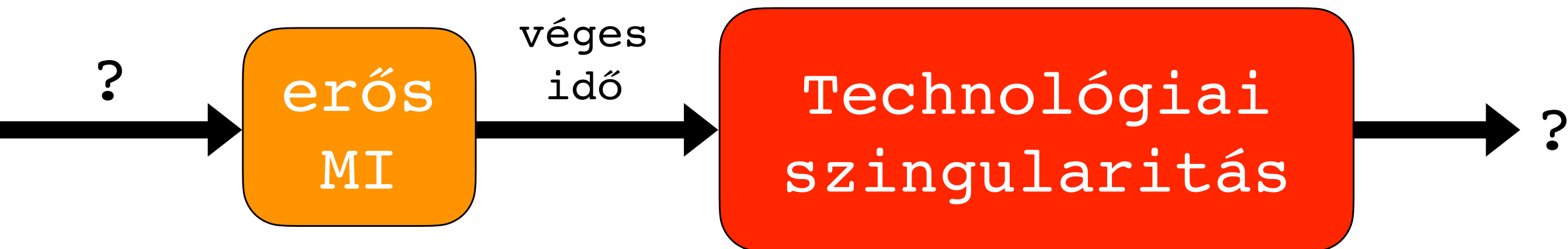
Source: Ray Kurzweil

Kurzweil jóslata a technológiai szingularitásra 2045





The Singularity Feedback Loop

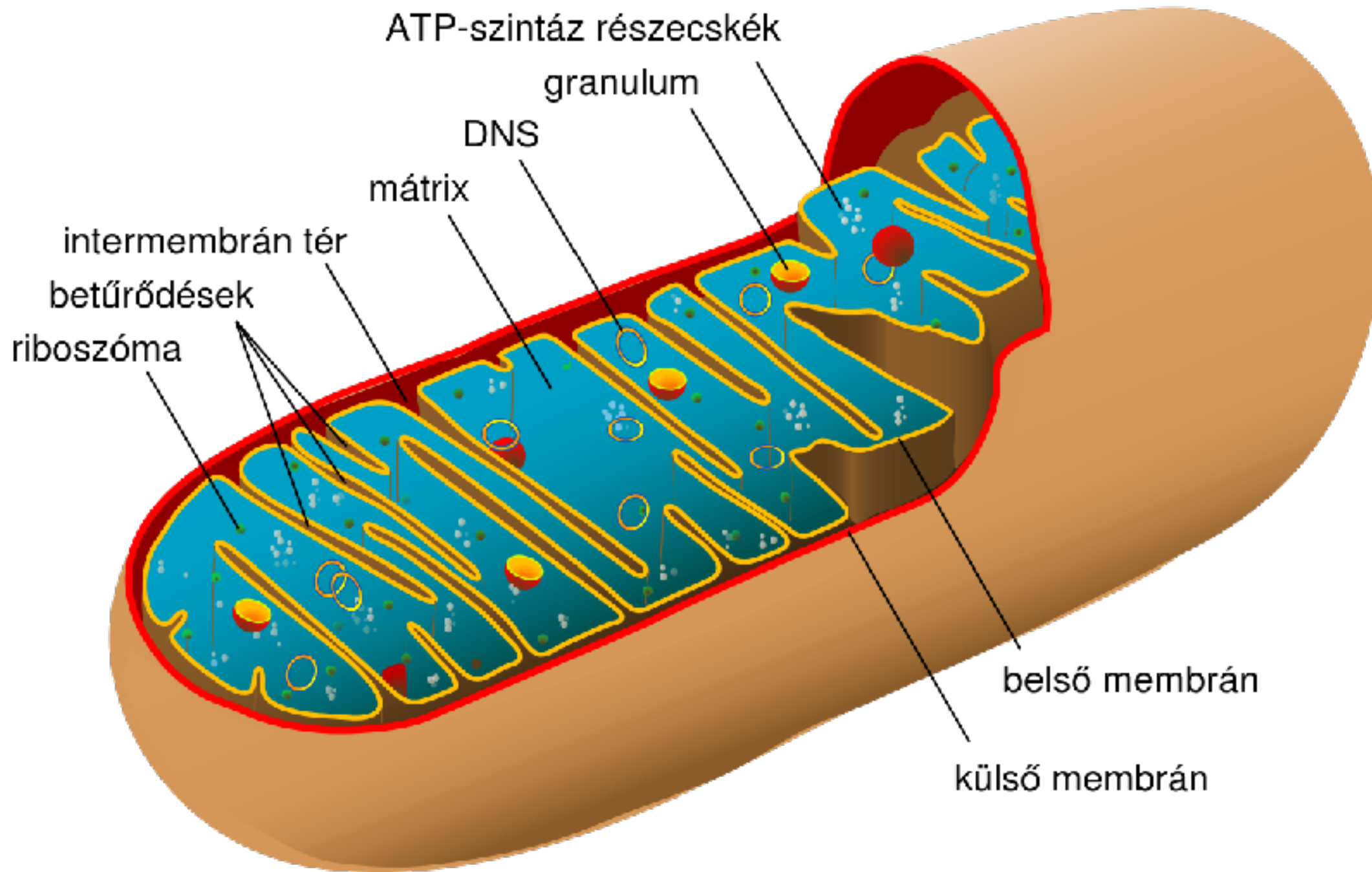


MI a tudományban

- programozás segítése
- gazdaság:
automatizáció, optimalizáció
- anyagtudomány
- idegtudomány segítése
- elméleti tudományok, alapkutató: AlphaTensor, AlphaGeometry, stb.

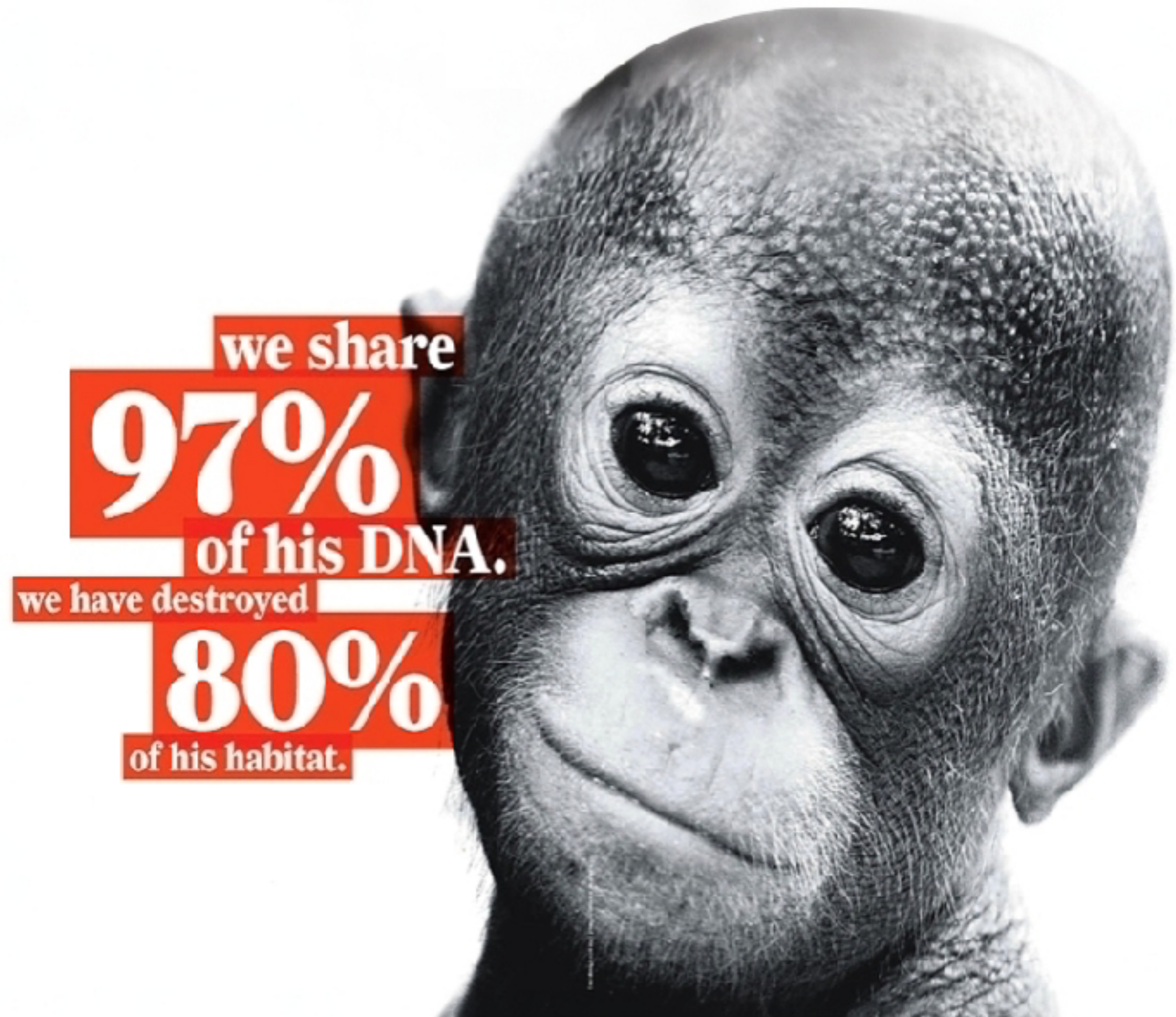
Lehetséges forgatókönyvek I.

Szimbiózis: Mitokondrium analógia



Lehetséges forgatókönyvek II.

Irrelevánsak leszünk számára



we share

97%

of his DNA.

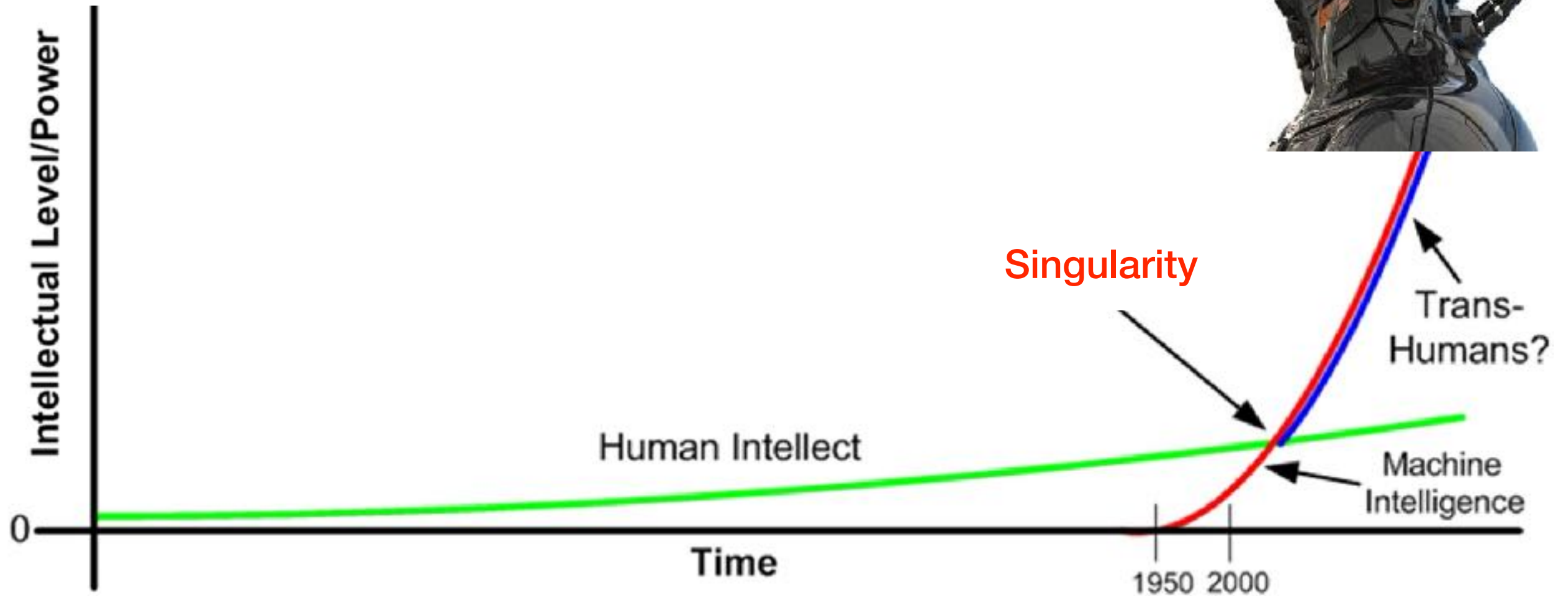
we have destroyed

80%

of his habitat.

Lehetséges forgatókönyvek III.

Mi változunk át fokozatosan





És még ezt is jó lenne elkerülni!



Melyiket rajzolta: művész, gyerek, majom, sertés szájjal, vagy a mesterséges intelligencia?

