







Artificial Intelligence has been inspired by models of the brain

- Neural Networks
- Deep Convolution
- Reward Reinforcement



Al's Technological Inflection Points

AI Concerns & 'Truthiness'

The Free Energy Principle (the low road)



Al's Technological Inflection Points

AI Concerns & 'Truthiness'

The Free Energy Principle (the low road)

"A LOGICAL CALCULUS OF IDEAS IMMANENT IN NERVOUS ACTIVITY"

1943, McCullough & Pitts. 1958 The Perceptron was implemented by Frank Rosenblatt, funded by the Office of Naval Research.

July 13, 1958



The New York Times Archives

Electronic 'Brain' Teaches Itself

The Navy last week demonstrated the embryo of an electronic computer named the Perceptron which, when completed in about a year, is expected to be the first non-living mechanism able to "perceive, recognize and identify its surroundings without human training or control." VIEW FULL ARTICLE IN TIMESMACHINE »

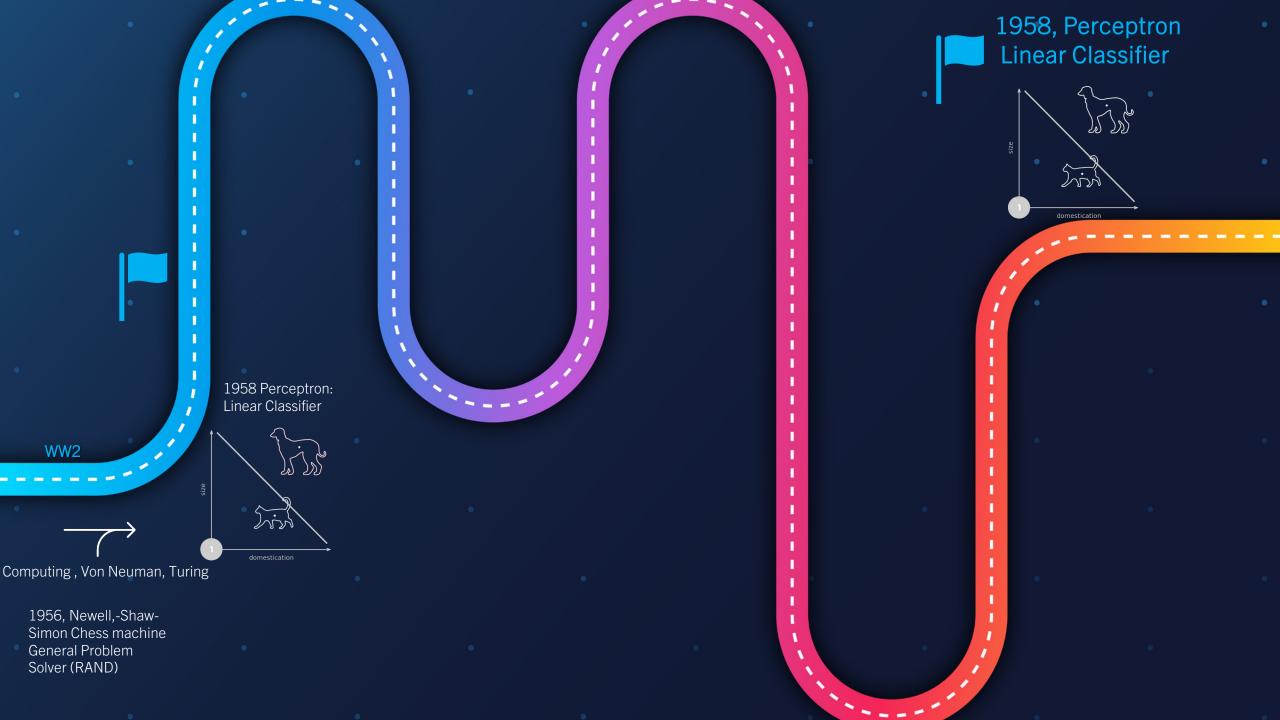


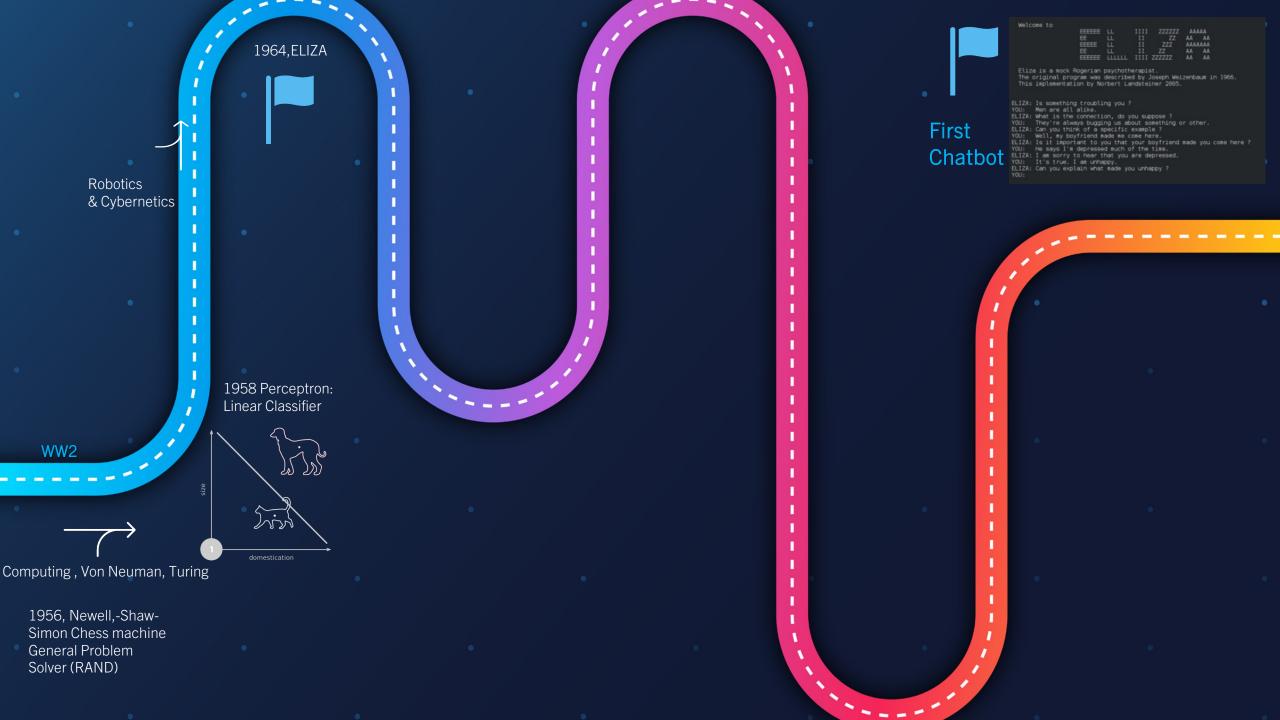


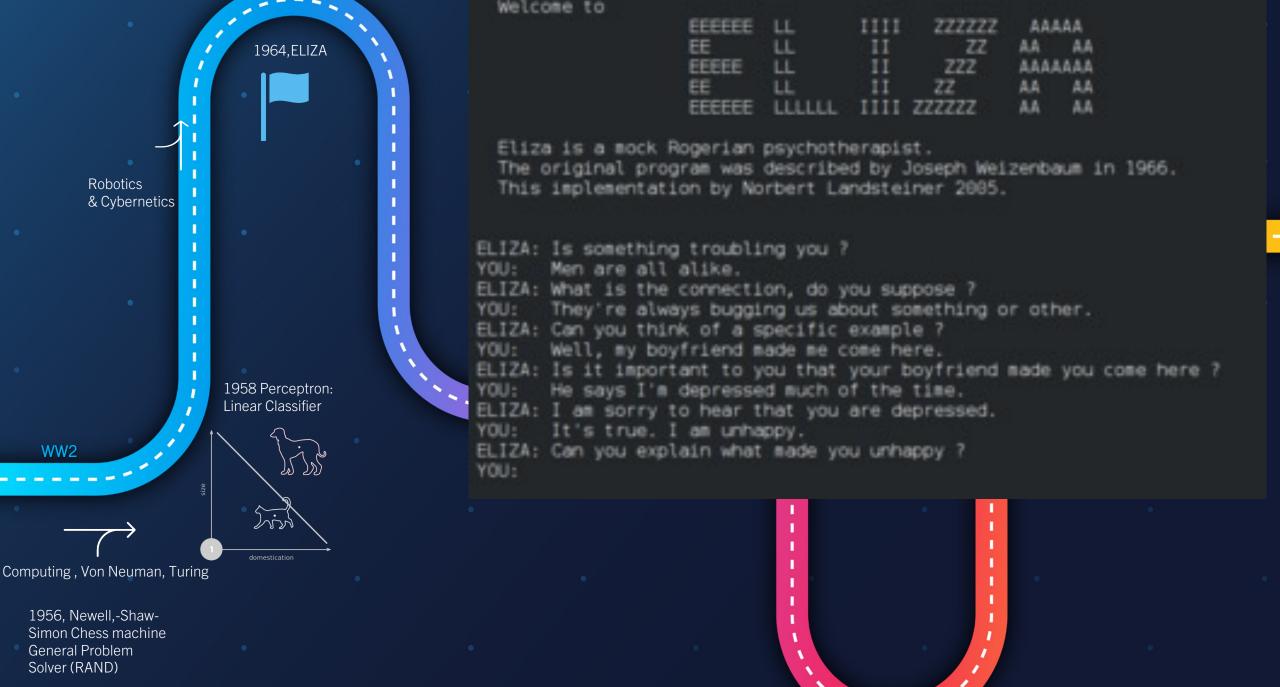


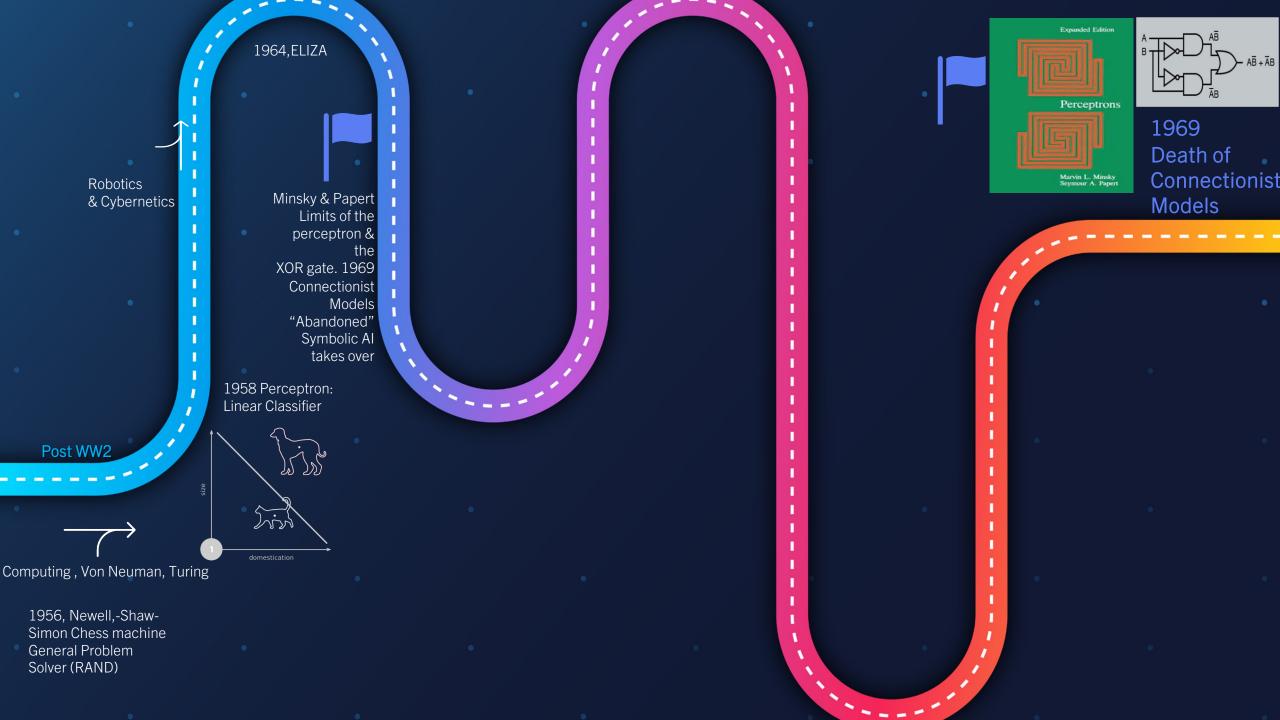


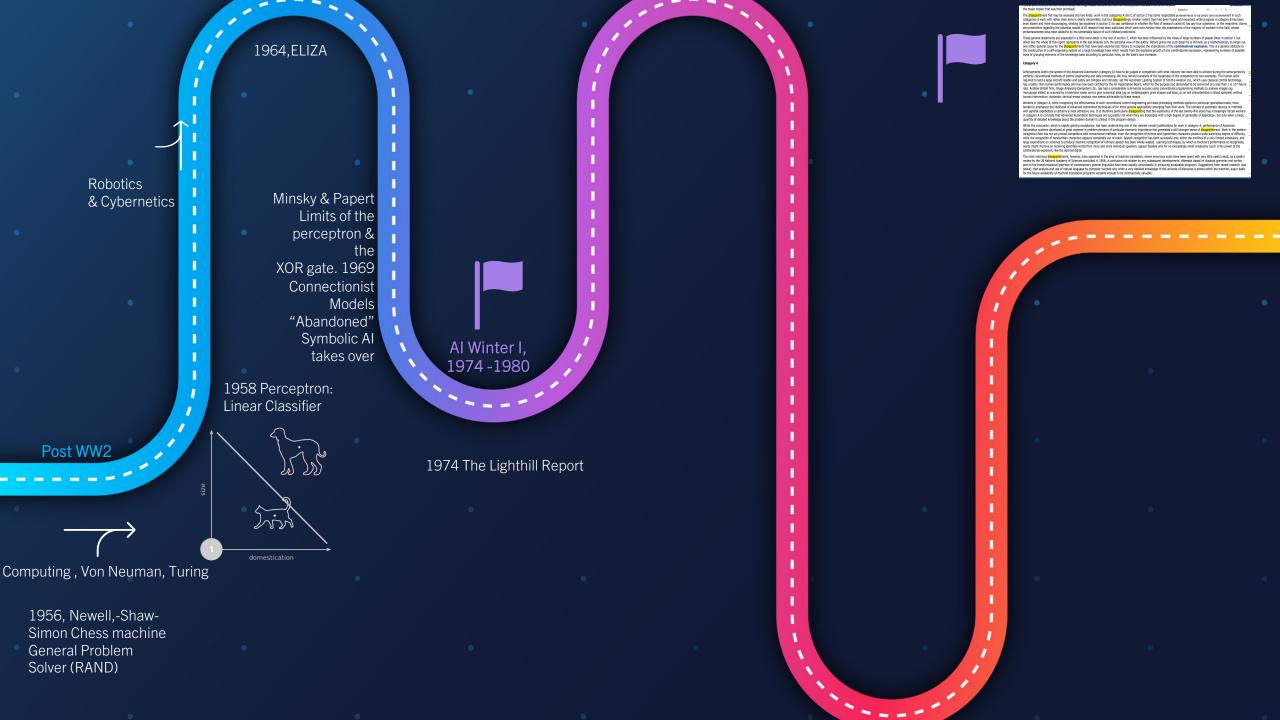


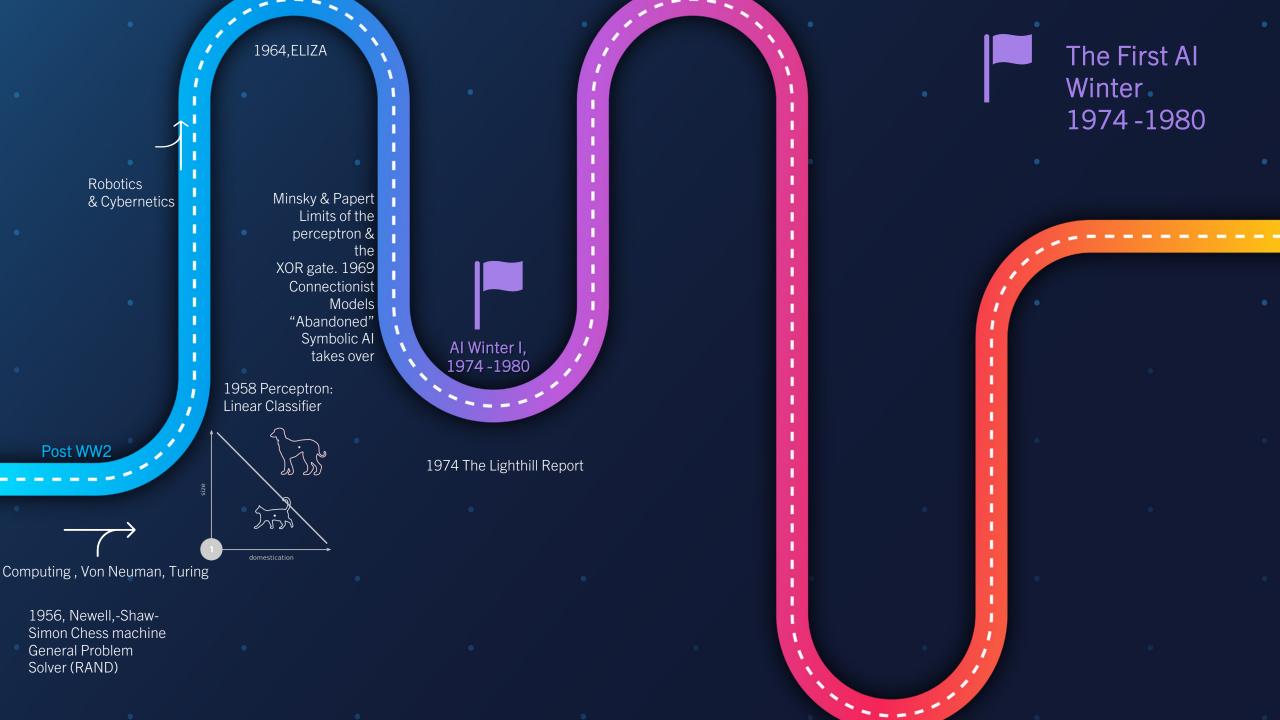


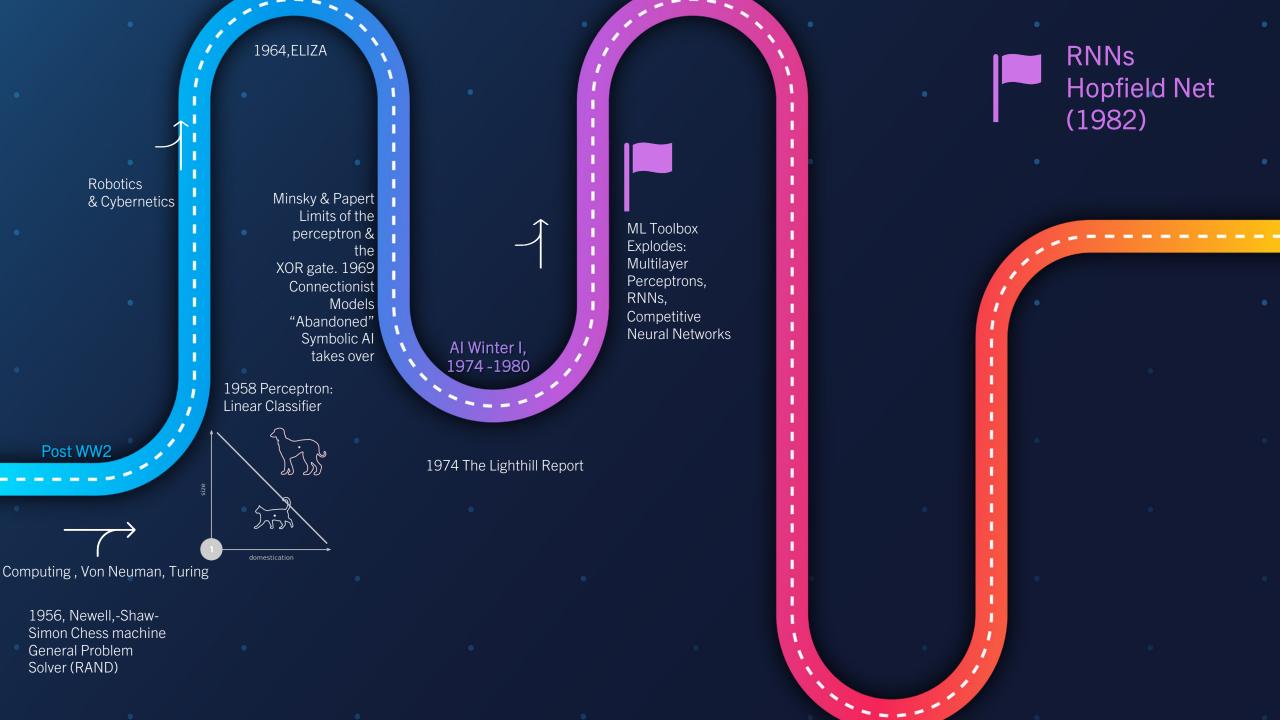


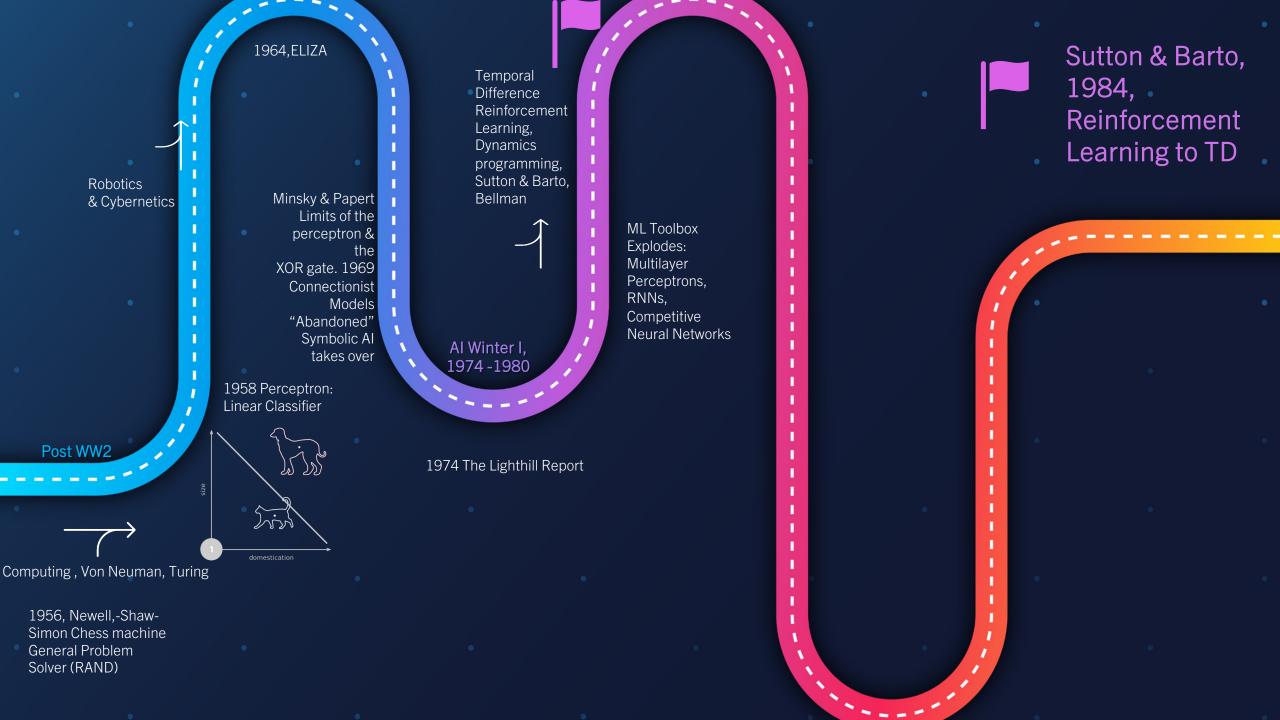


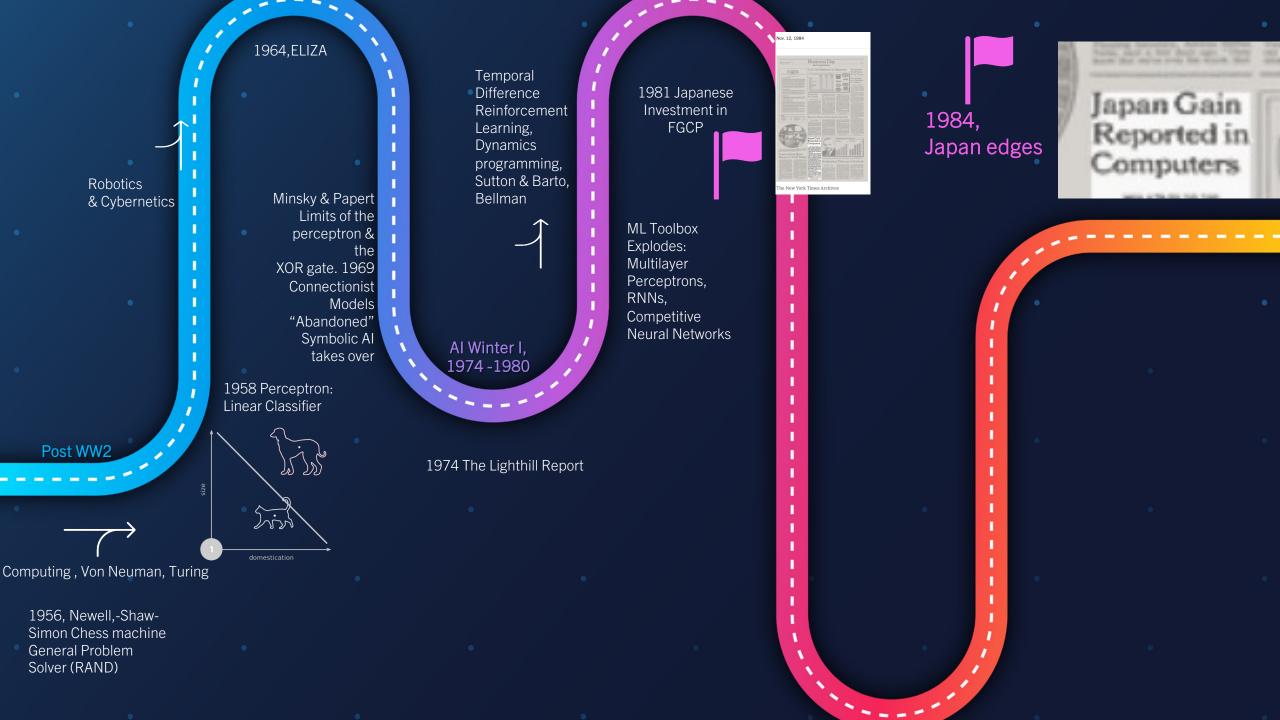


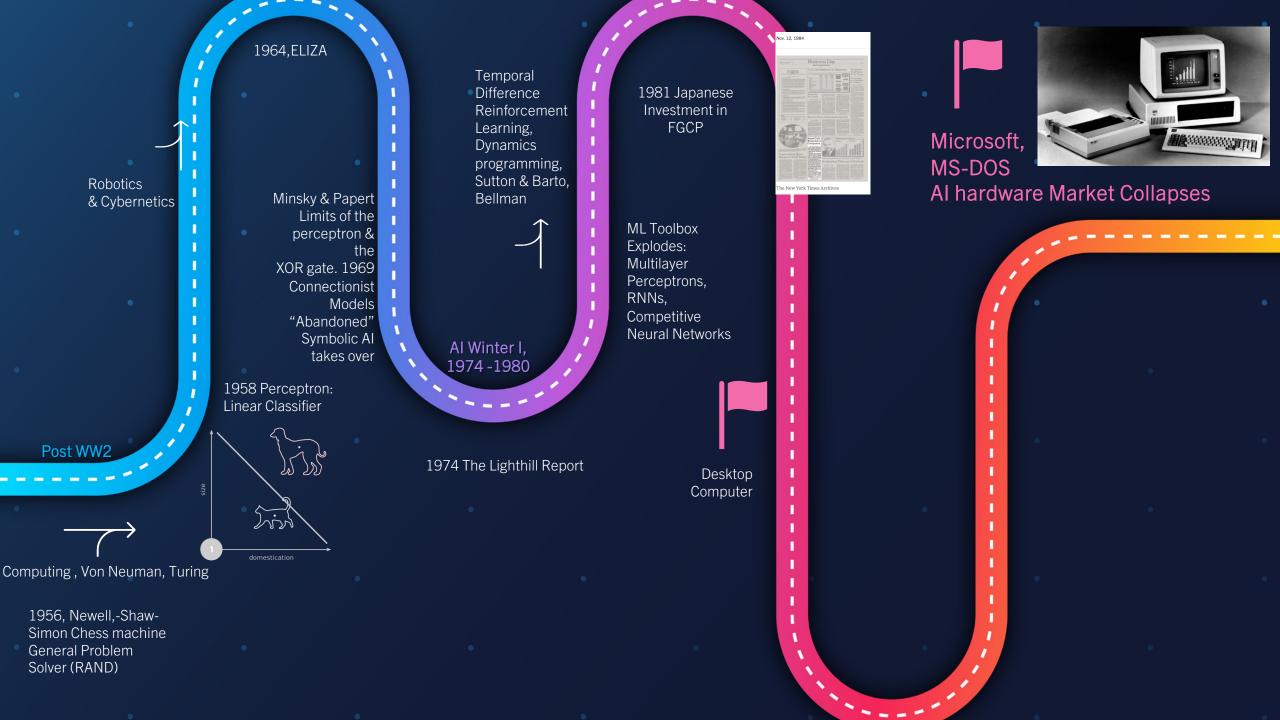


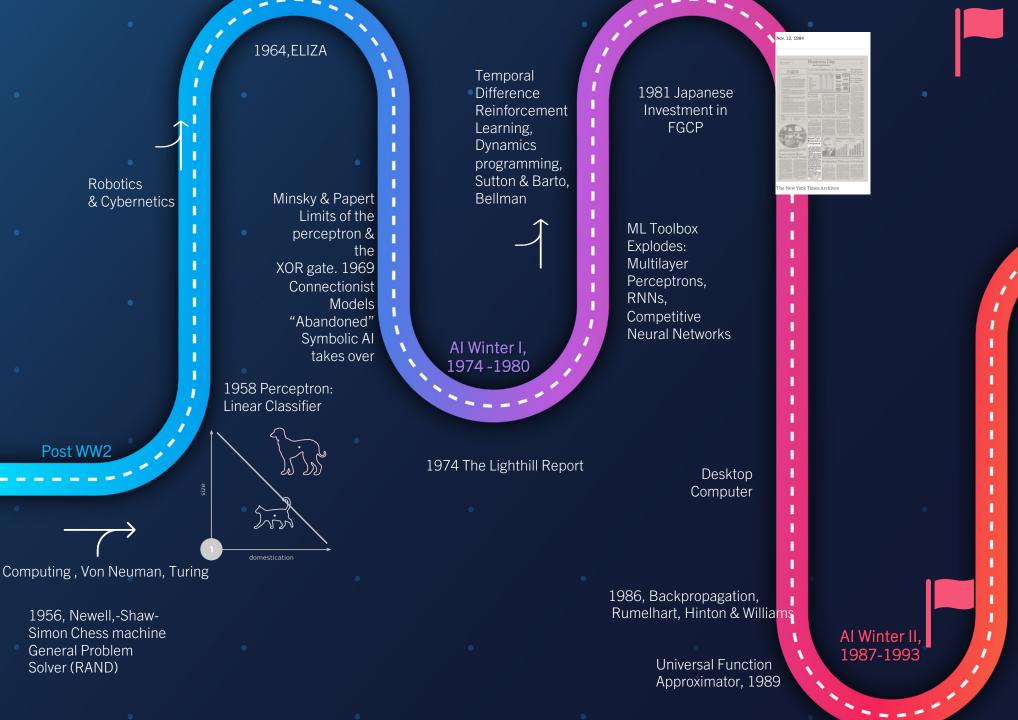




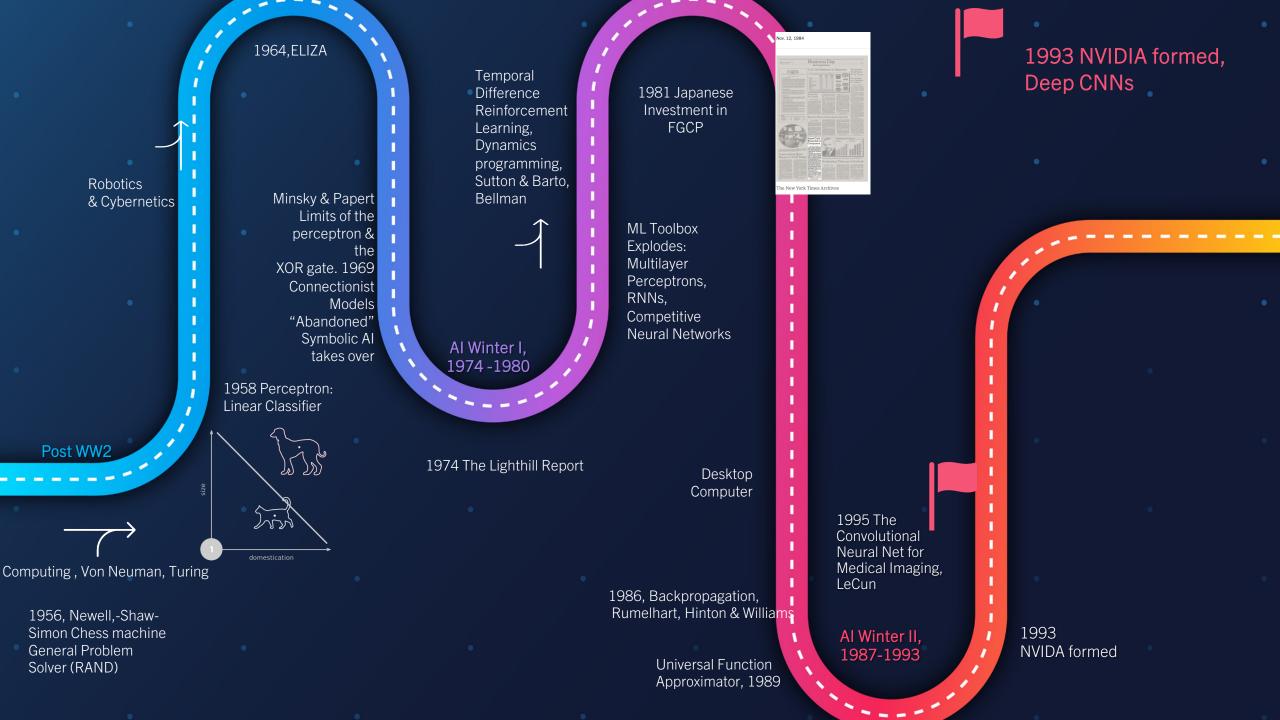


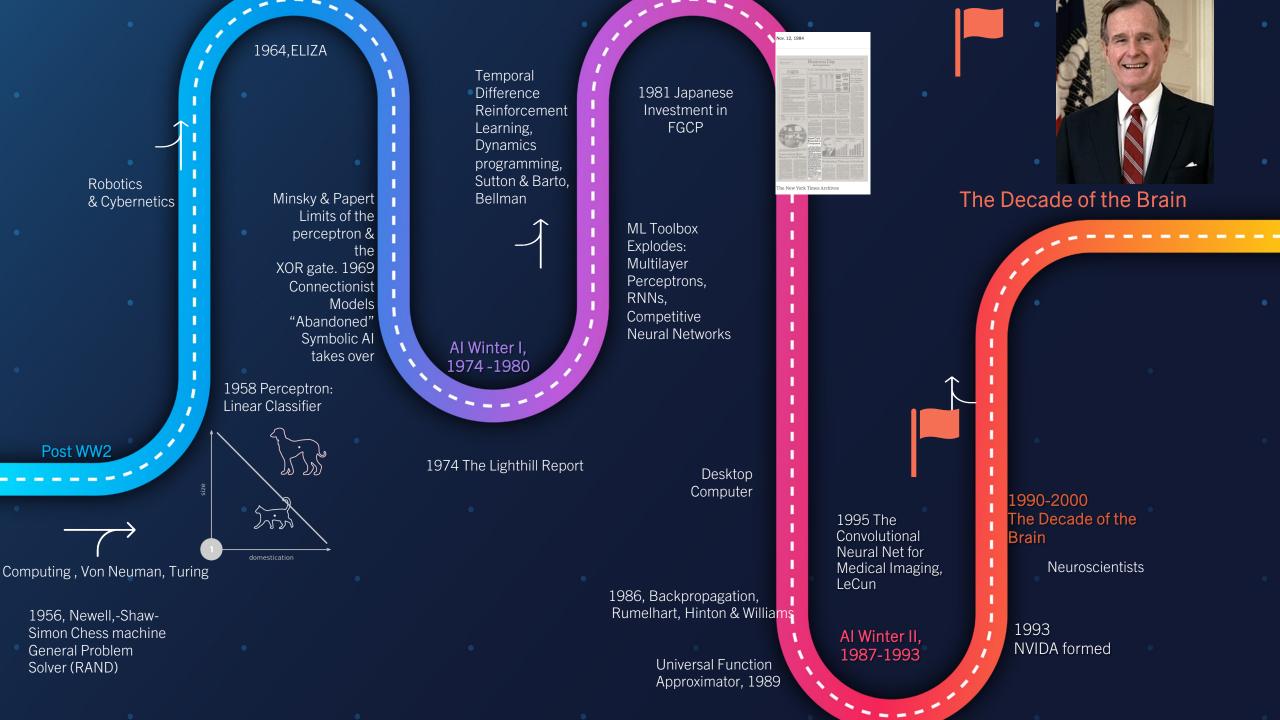


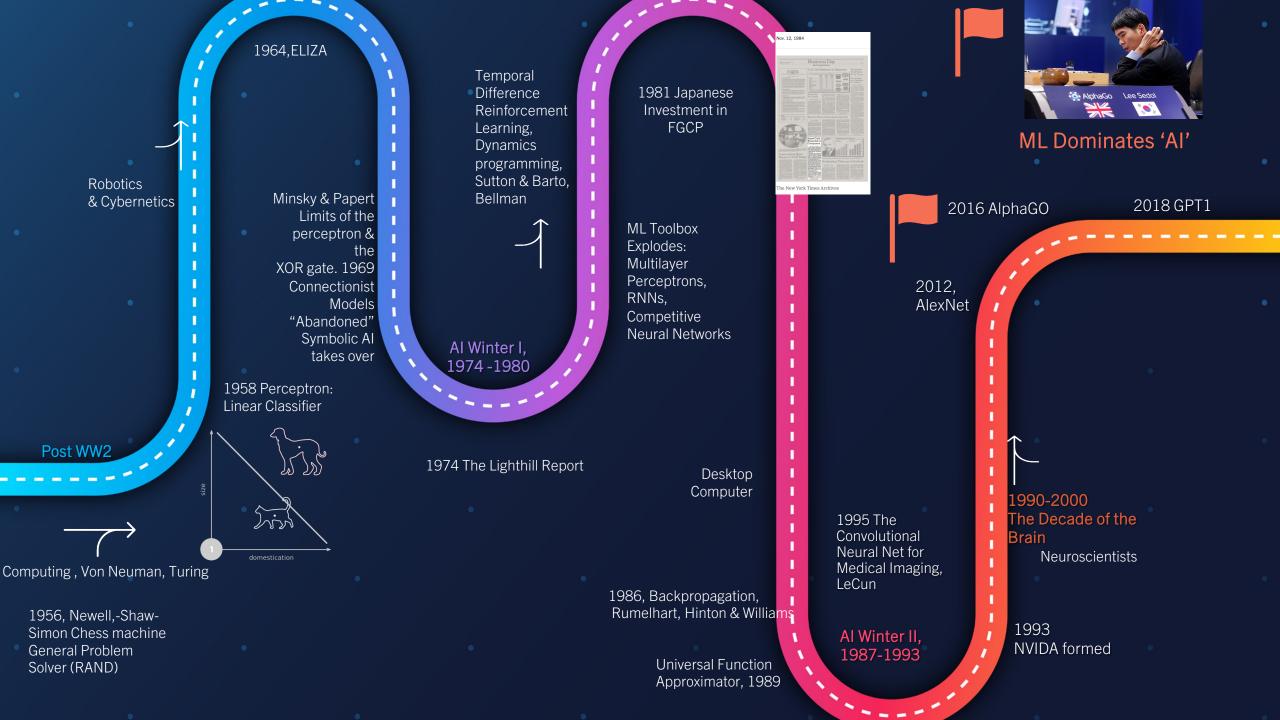


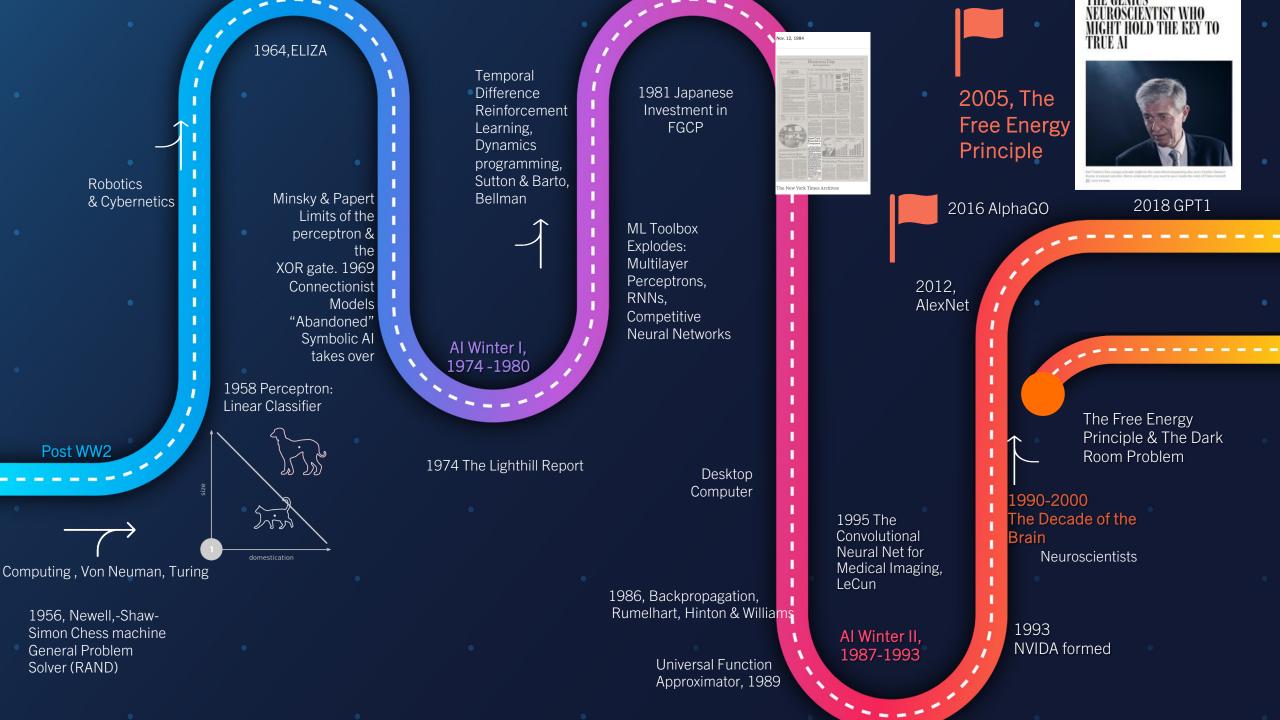


1987-1993 Second Al Winter & Al Breakthroughs





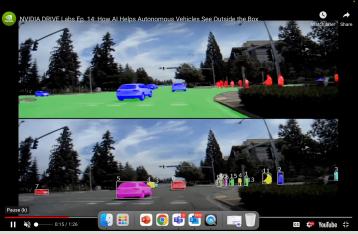


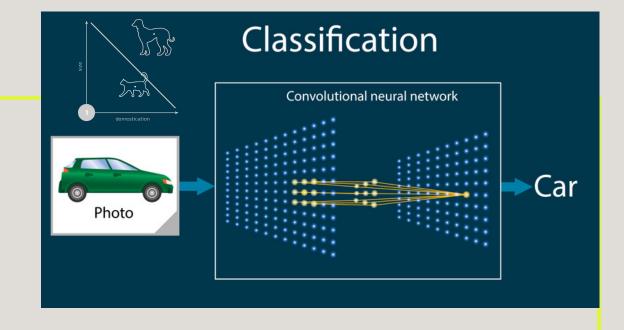


ML DOMINATES AI

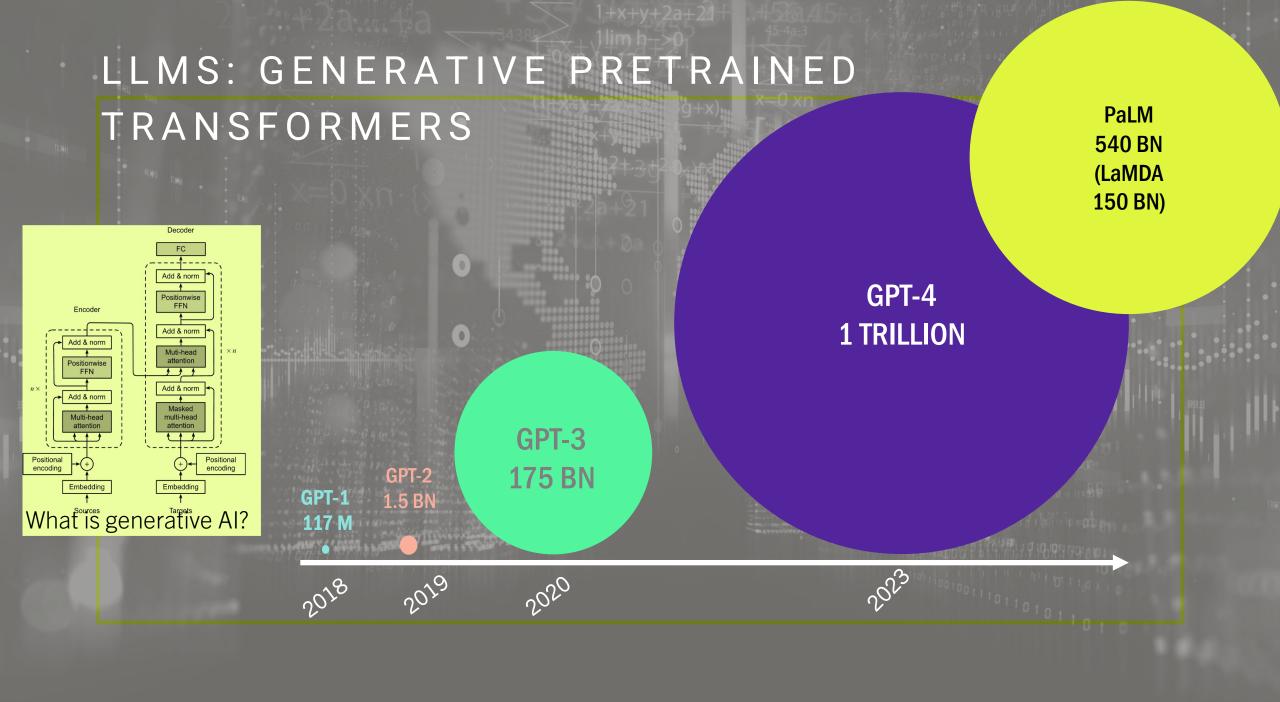
- Backprop (the chain rule)
- Bigger Models 8 layers
- Compute
- Data Supervised learning
- AlphaGo Zero trained using RL

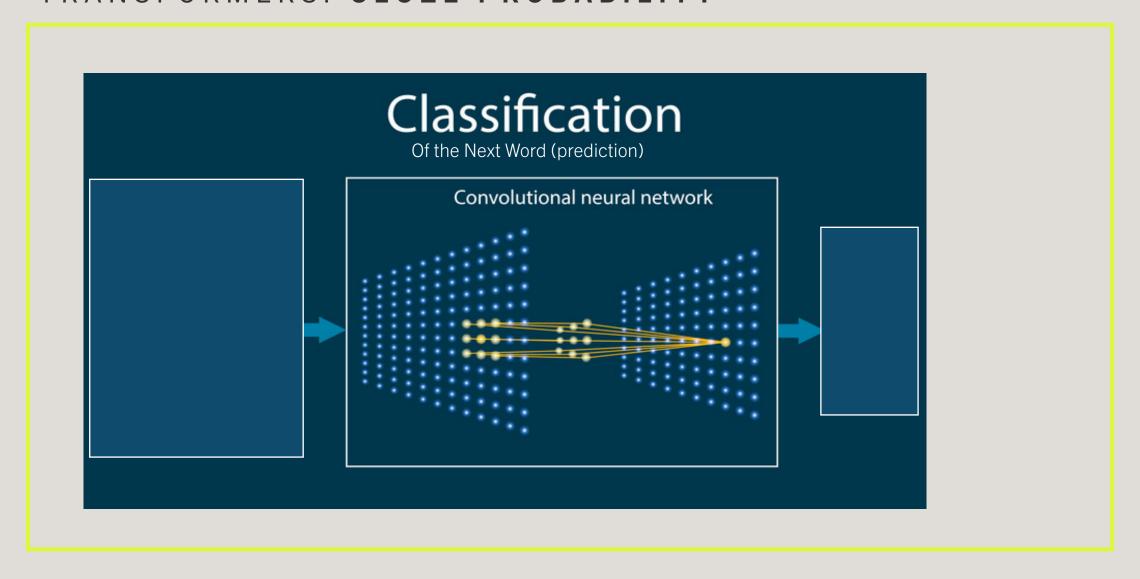


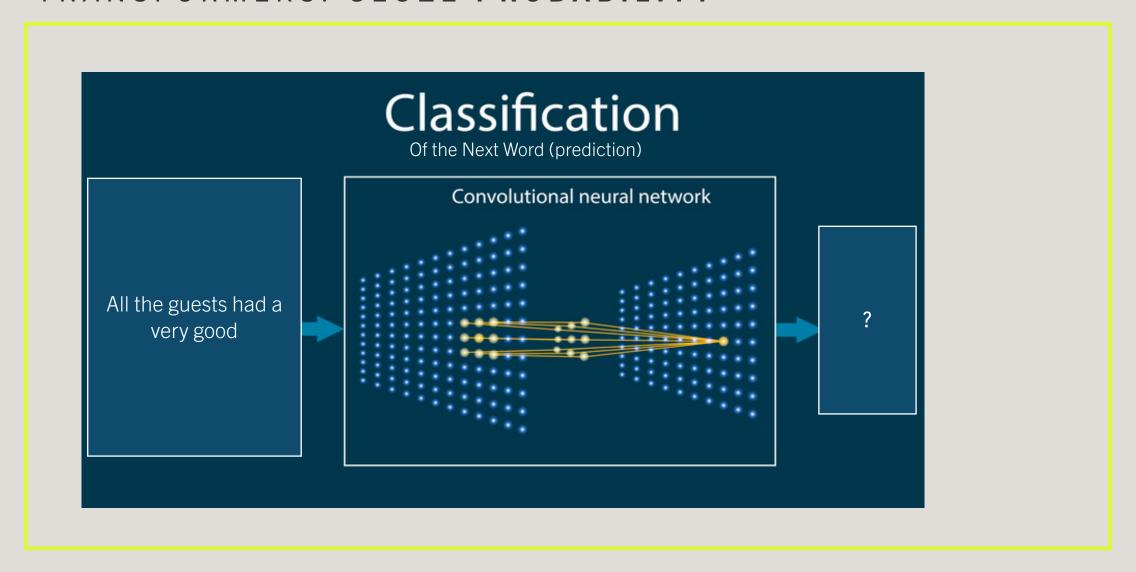


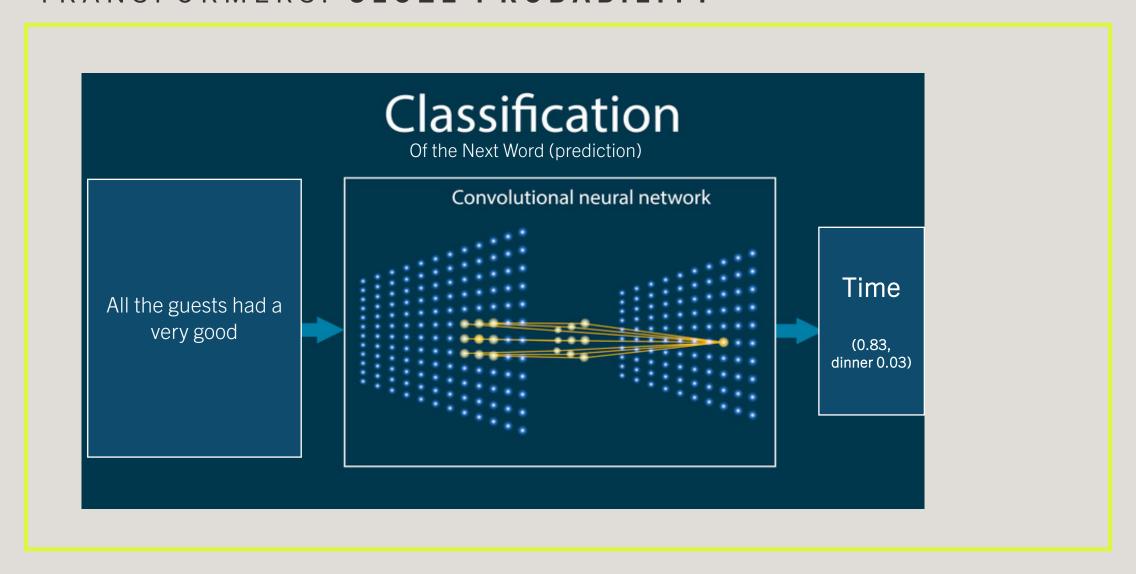


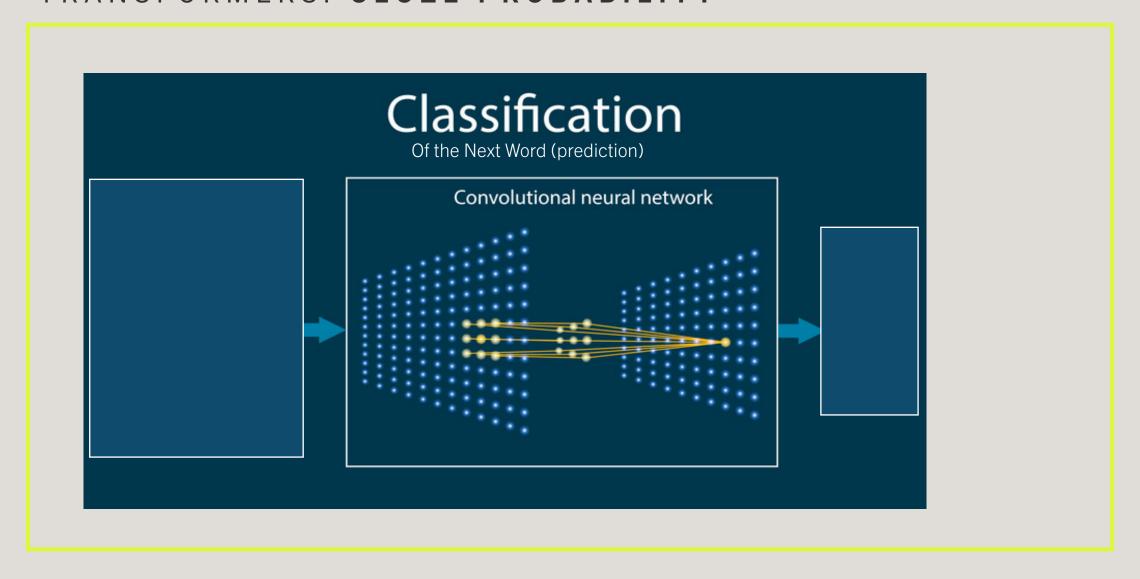


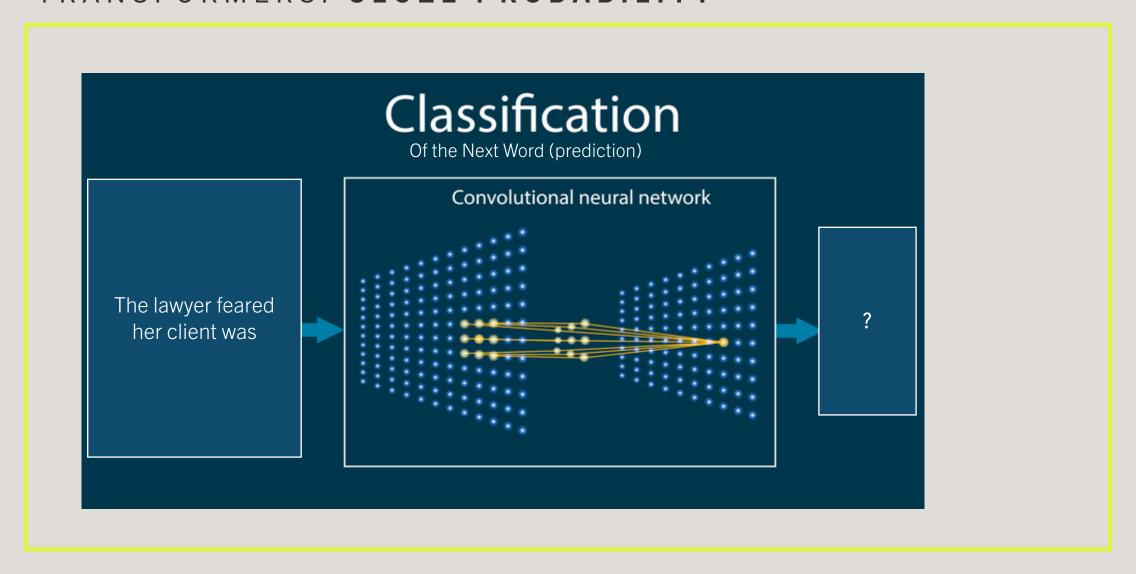


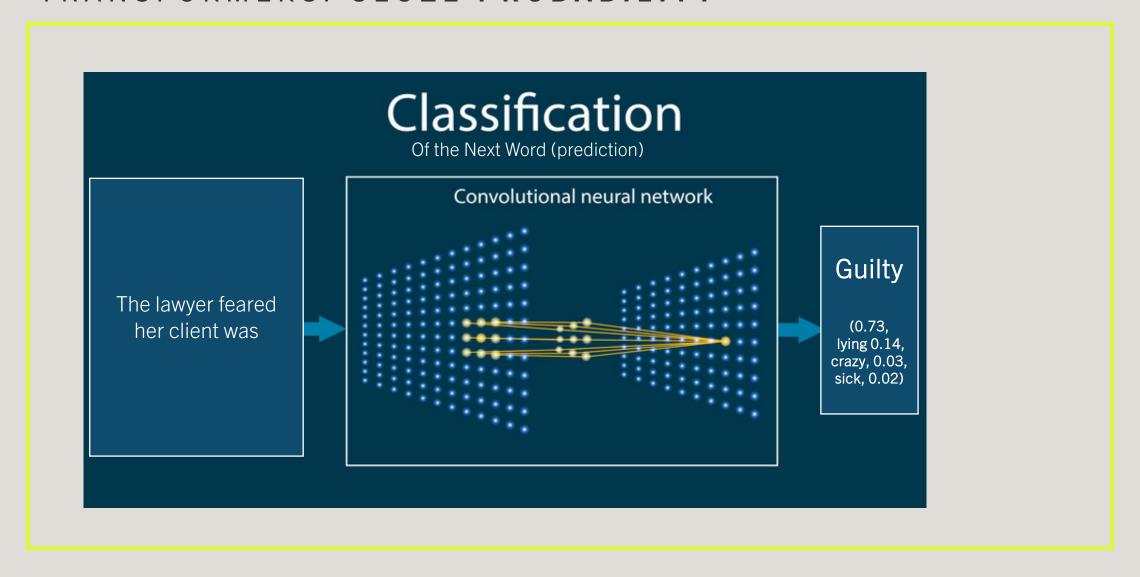








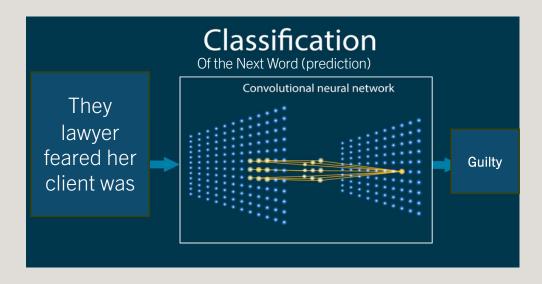




GENERATIVE PRE-TRAINED TRANSFORMERS

GPT-1 (2018), 117 million parameters: 'We evaluate on the Story Cloze Test [40], our model again outperforms the previous best results by significant margins - up to 8.9%

Gina misplaced her phone at her grandparents. It wasn't anywhere in the living room. She realized she was in the car before. She grabbed her dad's keys and ran outside.

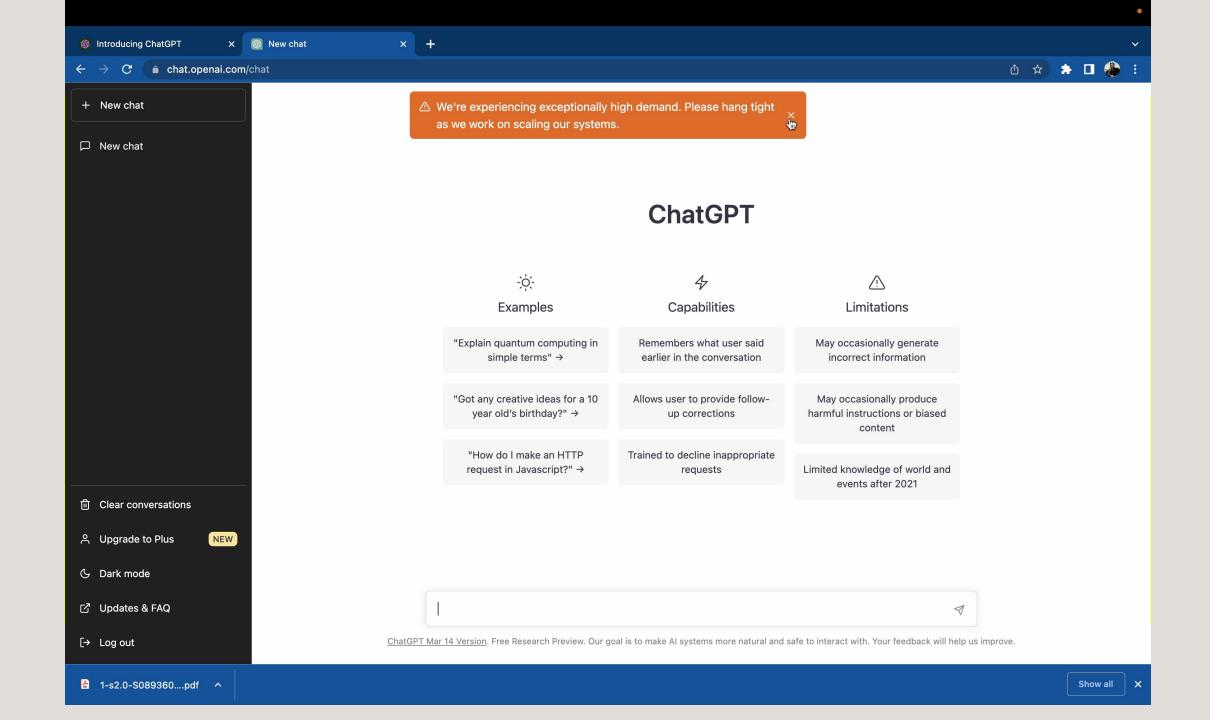


RIGHT ENDING

SHE FOUND HER PHONE IN THE CAR.

WRONG ENDING

SHE DIDN'T WANT HER PHONE ANYMORE.

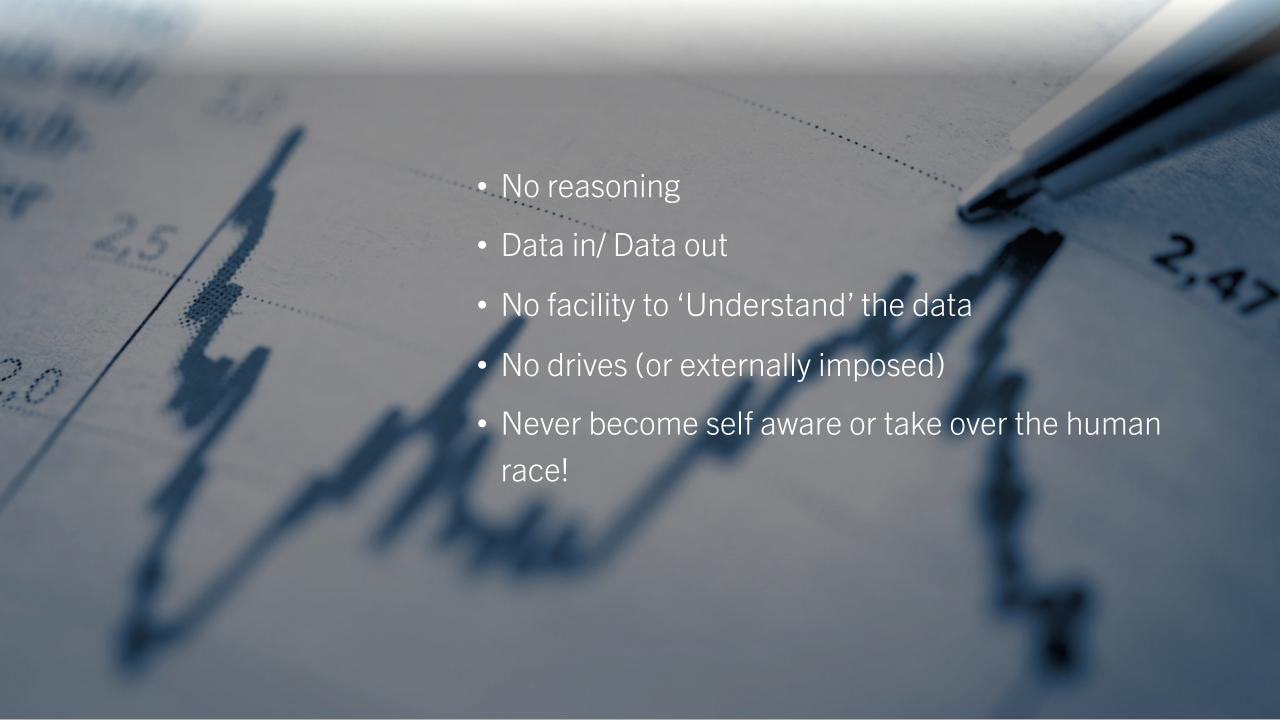




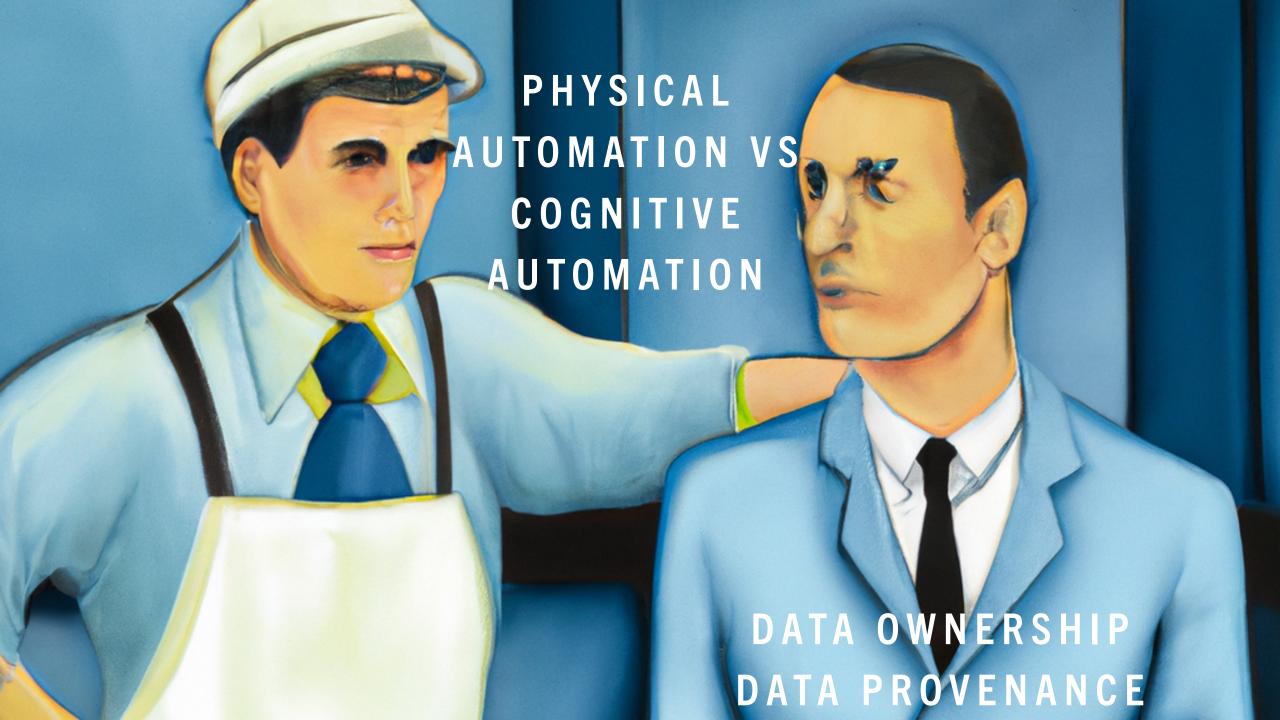
Al's Technological Inflection Points

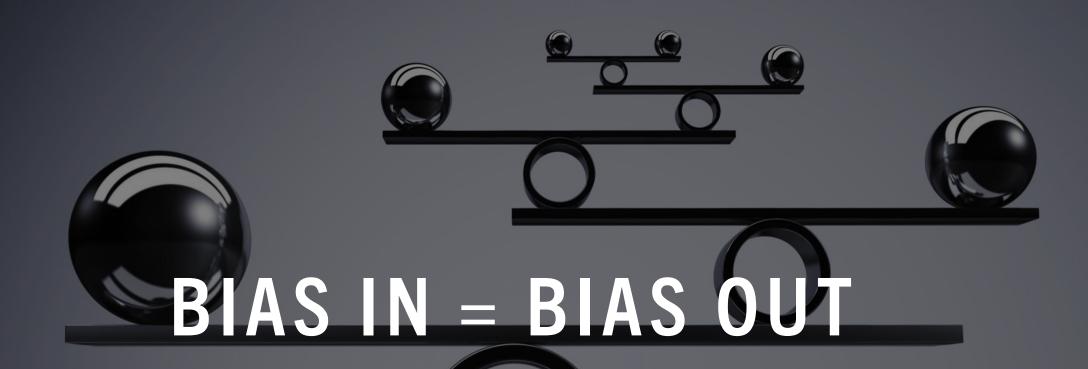
AI Concerns & 'Truthiness'

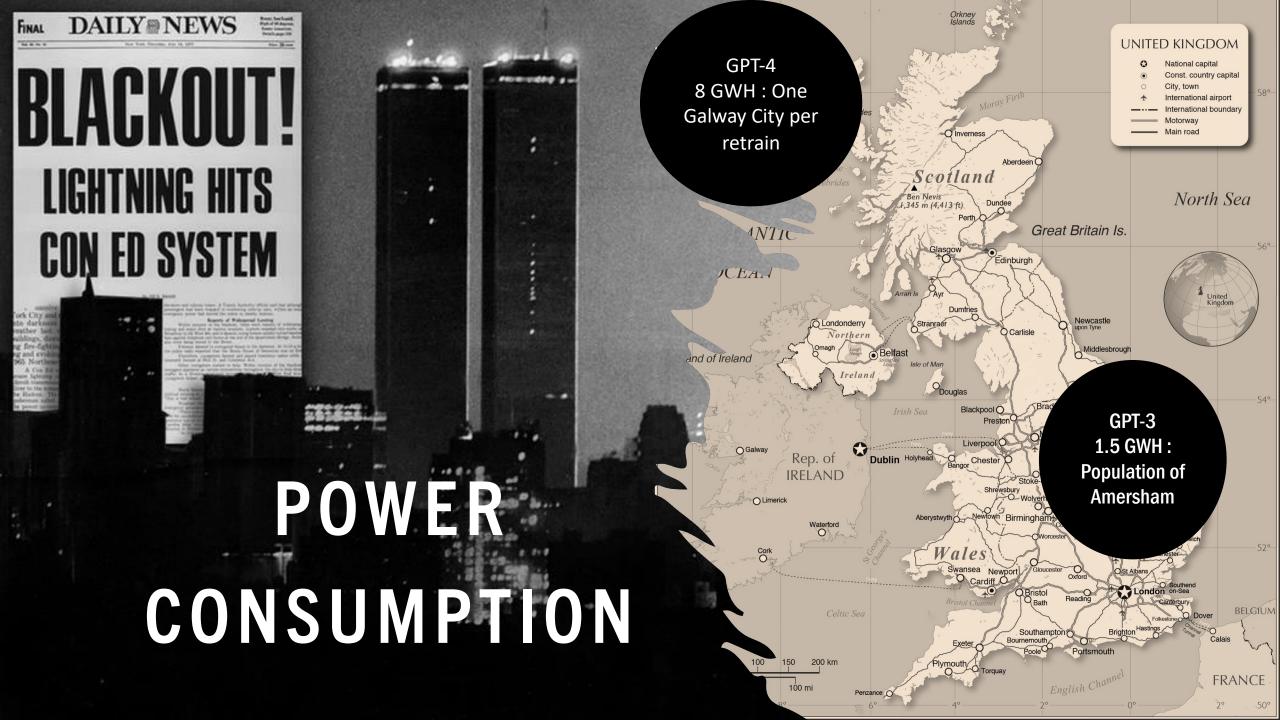
The Free Energy Principle (the low road)













Countries

Countries with M RI units

Regions

Population (mil)

Number of MRI u nits MRI units (per 1 mil)

212

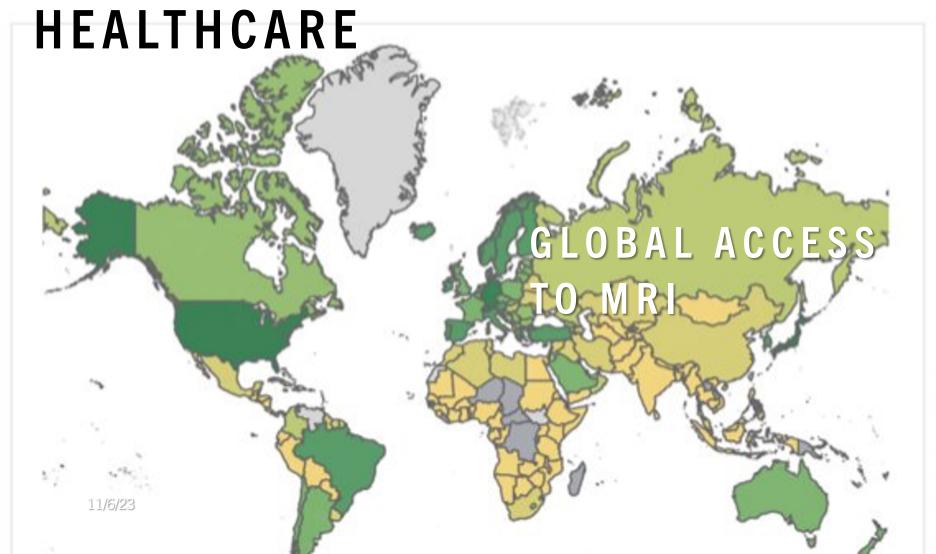
160

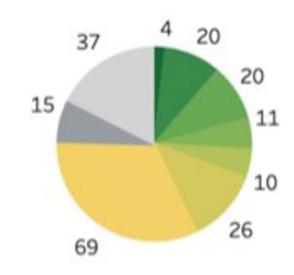
6

7,674M

51,498

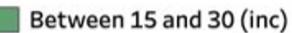
6.711

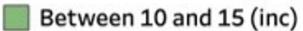




MRI units ranges

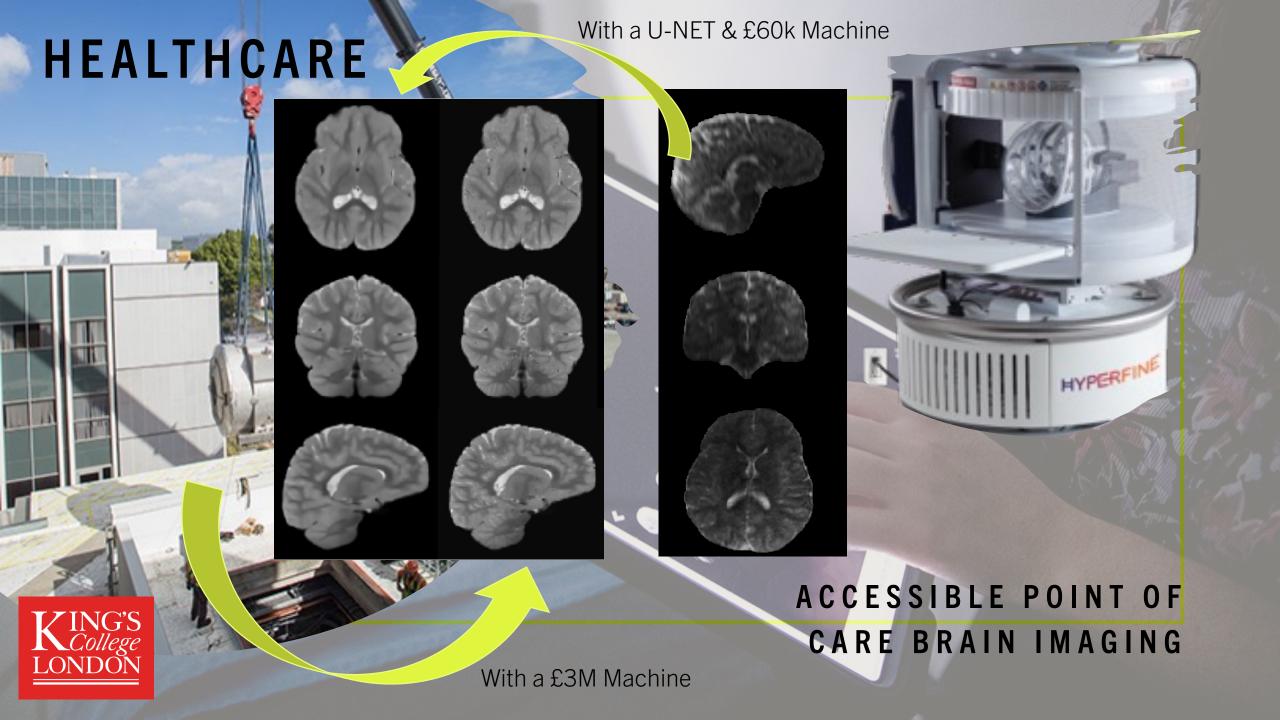












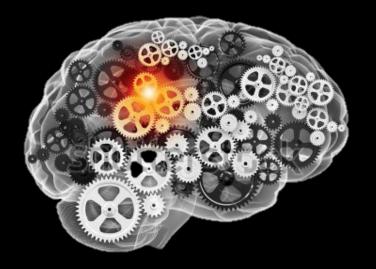


Artificial Intelligence has been inspired by models of the brain

- Neural Networks
- Deep Convolution
- Reward Reinforcement



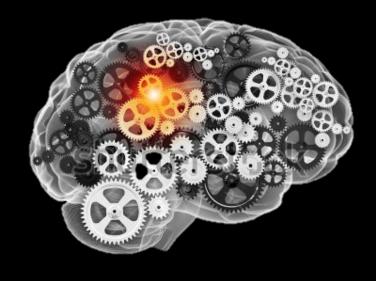
"Normative models describe how an optimal system would work given the goals. They describe 'what' the brain is trying to do."



"Process models are about the mechanisms, thus describing 'how' it is done."







Minimize Free Energy



In late 2017, a group led by Rosalyn Moran, a neuroscientist and engineer at King's College London, pitted two AI players against one another in a version of the 3D shooter game *Doom*. The goal was to compare an agent driven by active inference to one driven by reward-maximization.

Variational Message Passing To Minimize Free Energy



STANHOPE AI

Built from the inside out

WHAT IS A BRAIN TRYING TO DO?

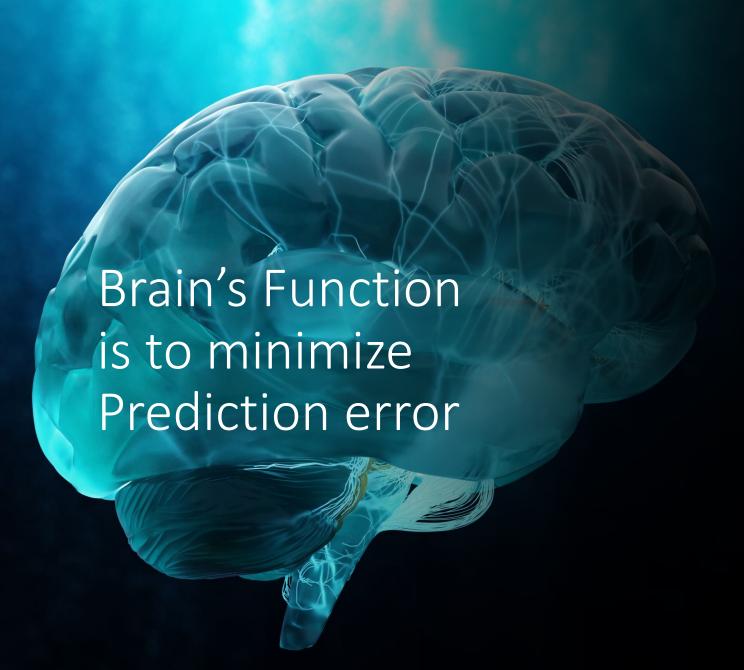
How would it find its own data, and generate its own model?



Al's Technological Inflection Points

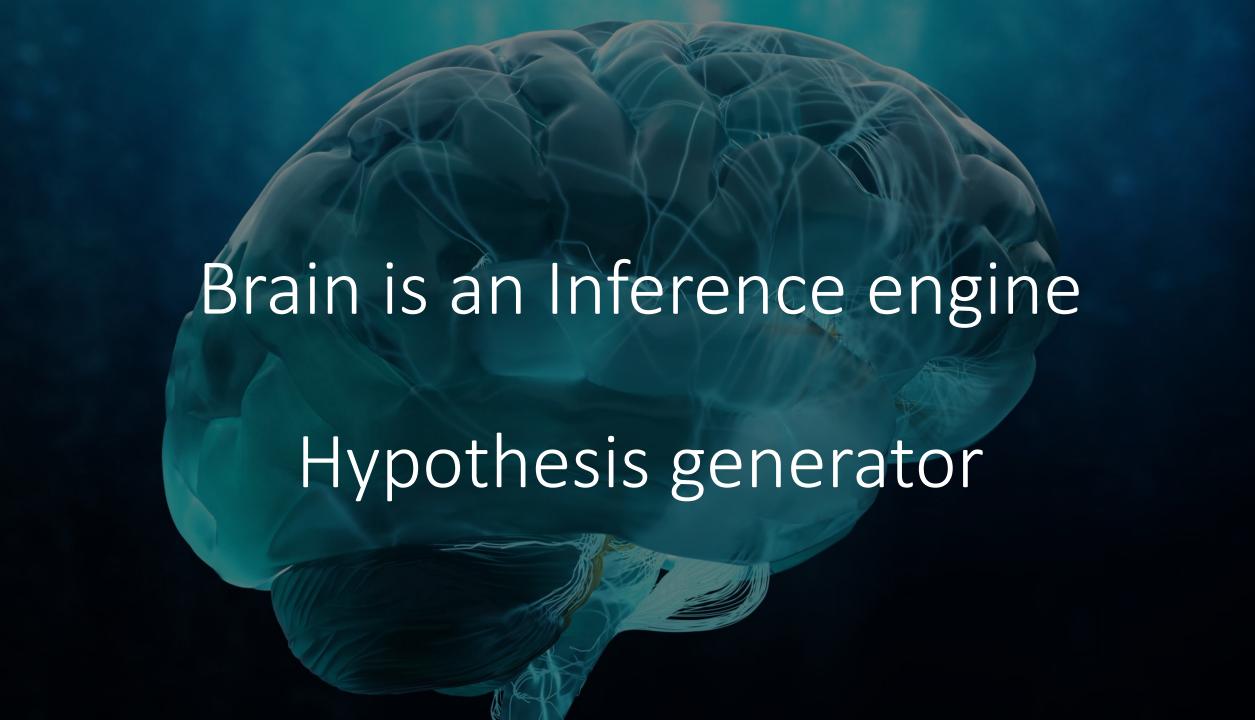
Catastrophe & 'Truthiness'

The Free Energy Principle (the low road)





A Non-Equilibrium Steady State The Fokker-Planck equation





Inference

Real world

The free energy principle is

a formal statement that explains how living and non-living systems remain in nonequilibrium steady-states by restricting themselves to a limited number of states

Biological agents perceive and act in order to minimize long term surprise. It can predict its world. To do so an agent could model the world and try to predict the causes of sensations.

Friston, A Theory of Cortical Responses 2006

The world in my brain

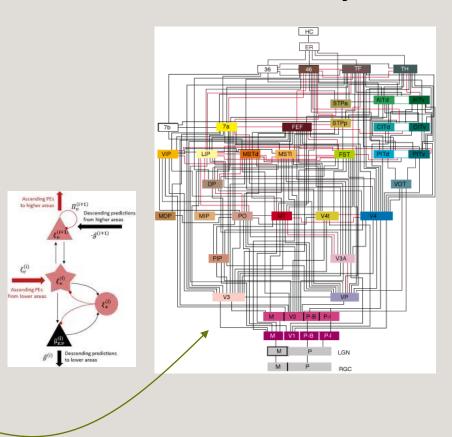
Sensory Data, y

Inference

Real world



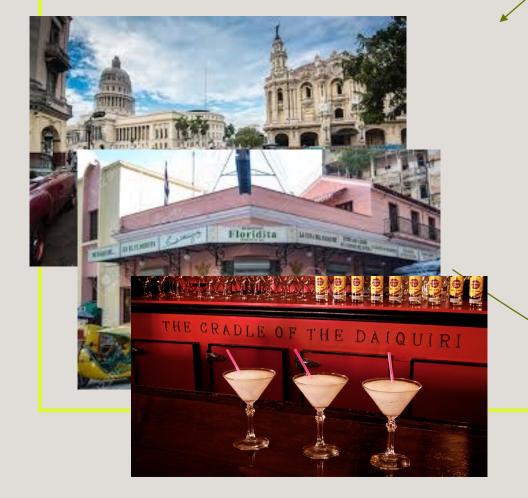
The world in my brain



Sensory Data, y

And Act

Real world (p)

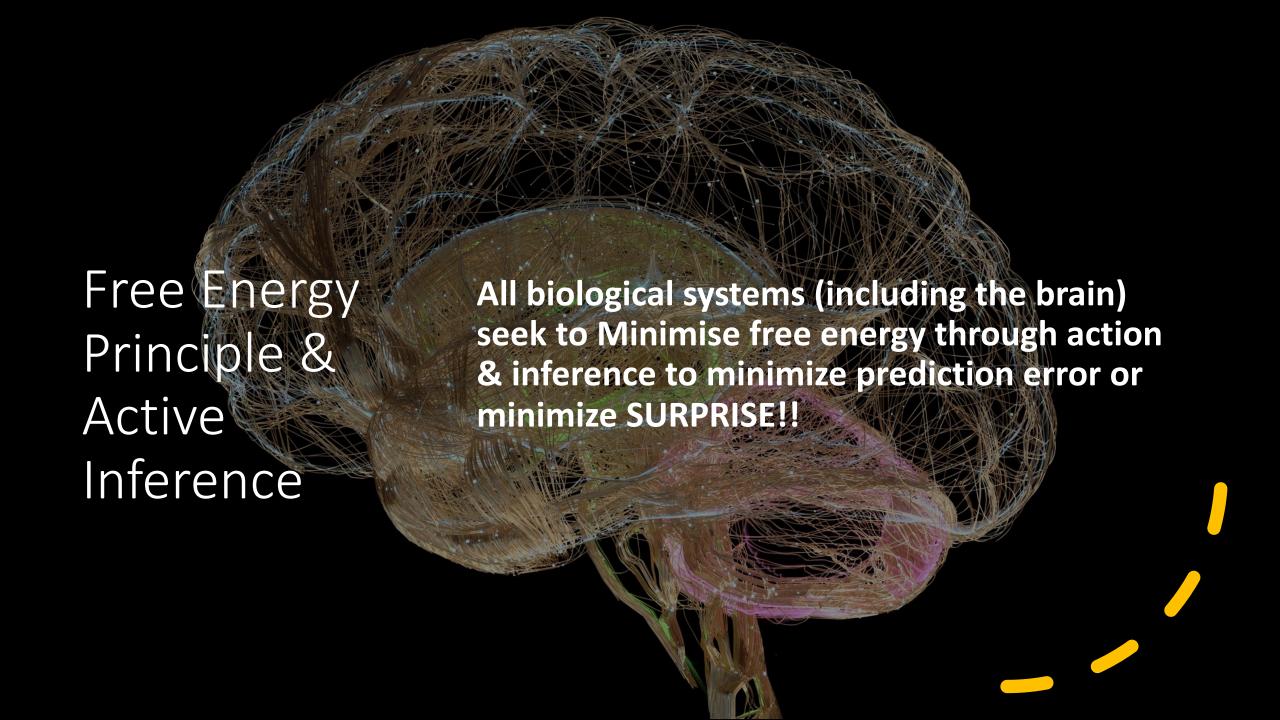


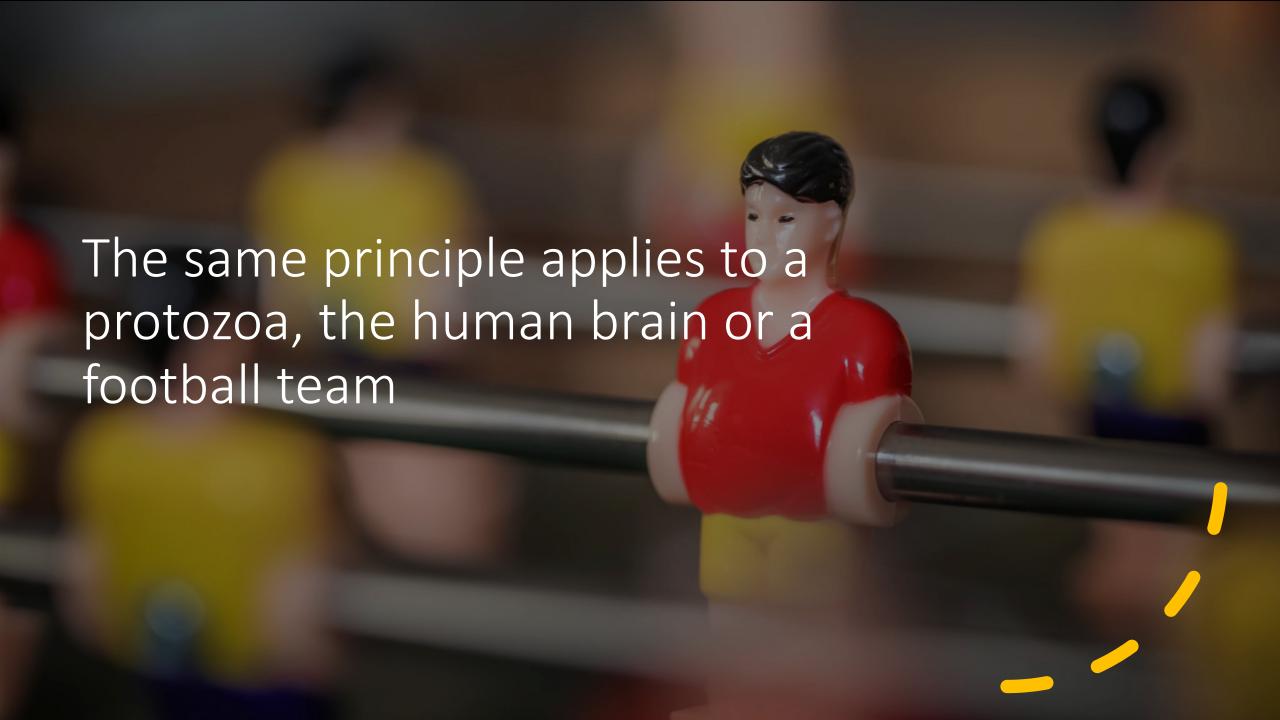
action, a

The world in my brain (q) Sensory Data, y

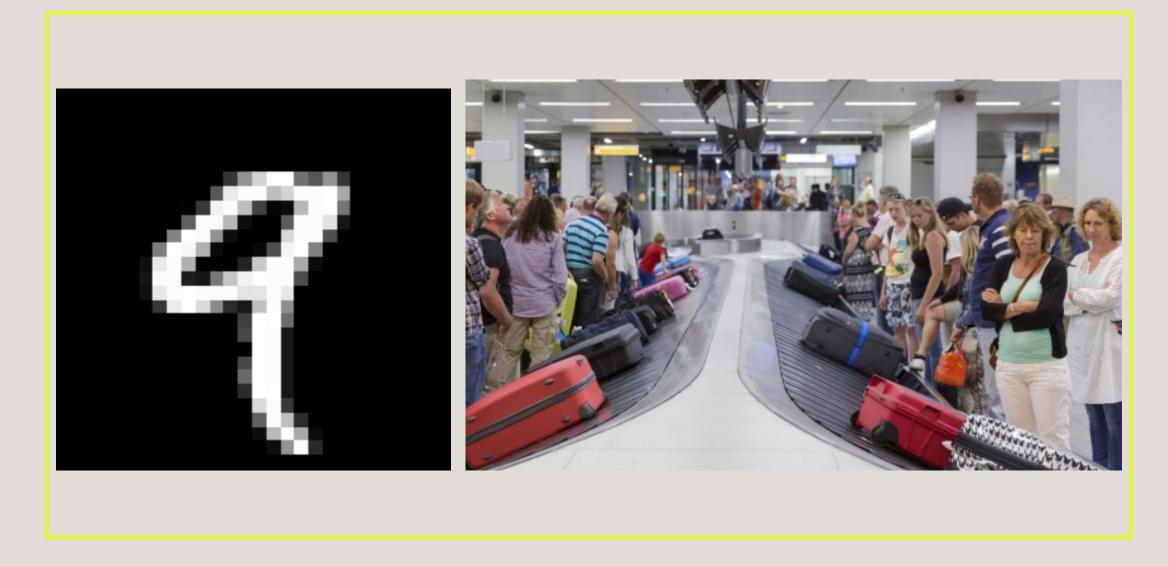
Sensory Data, y

Sensory Data, y

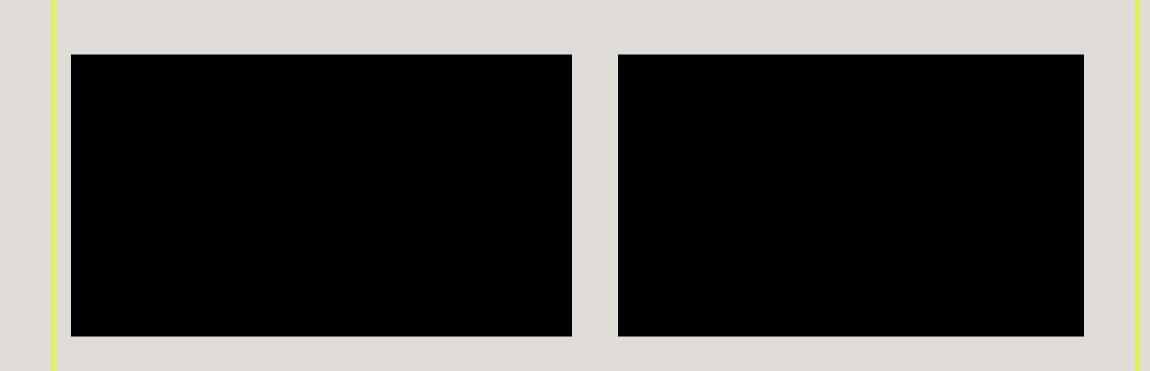




Revisiting MNIST, Why do humans 'see'? How do humans 'see'?



Revisiting MNIST



Vision is Active, Top-down, Sparse, Deliberative, Predictive

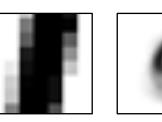
Action & Belief Under Active Inference



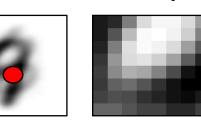
У

 T_1

 $R(y,x_p)$

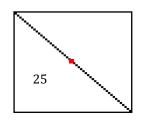


 $\widetilde{\boldsymbol{y}}$

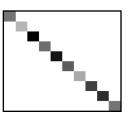


 $R(\widetilde{y},\widetilde{x}_p)$

 $o(loc)_t$



 $\boldsymbol{E_t}$



 T_2

 T_3

Hidden Stimulus

Observed portion of stimulus

The agent's visual expectation about the global & local scene generated from the prior network.

Posterior expectation about the target location $\widetilde{\mathbf{o}}_1$. & Ascending posterior expectation about the target digit $\widetilde{\mathbf{o}}_h$.

- The crucial understanding in the difference between Bayesian theory of brain function and the Free Energy Principle
- Biological systems are active, they change they move, they interact with their environment
- The inference (model, prediction) about the external world is not static but active
- Two essential implications for intelligence in organs, organisms and the brain

"ACTIVE" INFERENCE

Karl Friston Anna Sales Maell Cullen Anjali Bhat Berk Mirza **Ashley Tyrer** Jamie Thackrer Nicole Himmelstoß Elena Paci Valeria Olivia Zeliang Wang

Jonathan Monney