A BRIEF HISTORY OF

# ARTIFICIAL INTELLIGENCE

@tlberglund
linkedin.com/in/tlberglund

# CAN MACHINES THINK?

> *"LMAO who cares."*
>
> — Alan Turing, "Computing Machinery and Intelligence" (1950)



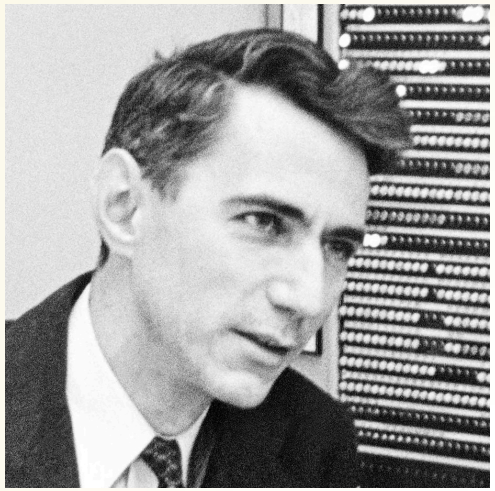
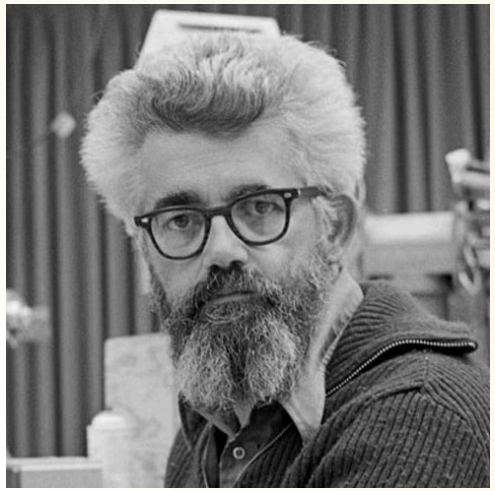
# WE JUST NEED TWO MONTHS

> *"We propose that a 2 month, 10 man study of artificial intelligence be carried out.... We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."*
>
> — The Dartmouth Workshop Proposal

**Jun 18, 1956**

1960s

# WE JUST NEED TWO MONTHS



Jun 18, 1956

1960s

# JUST TWO MONTHS?

> *The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.*
>
> — The Dartmouth Workshop Proposal

- Natural language processing
- Neural networks
- Theory of computation
- Abstraction
- Creativity

**Jun 18, 1956**

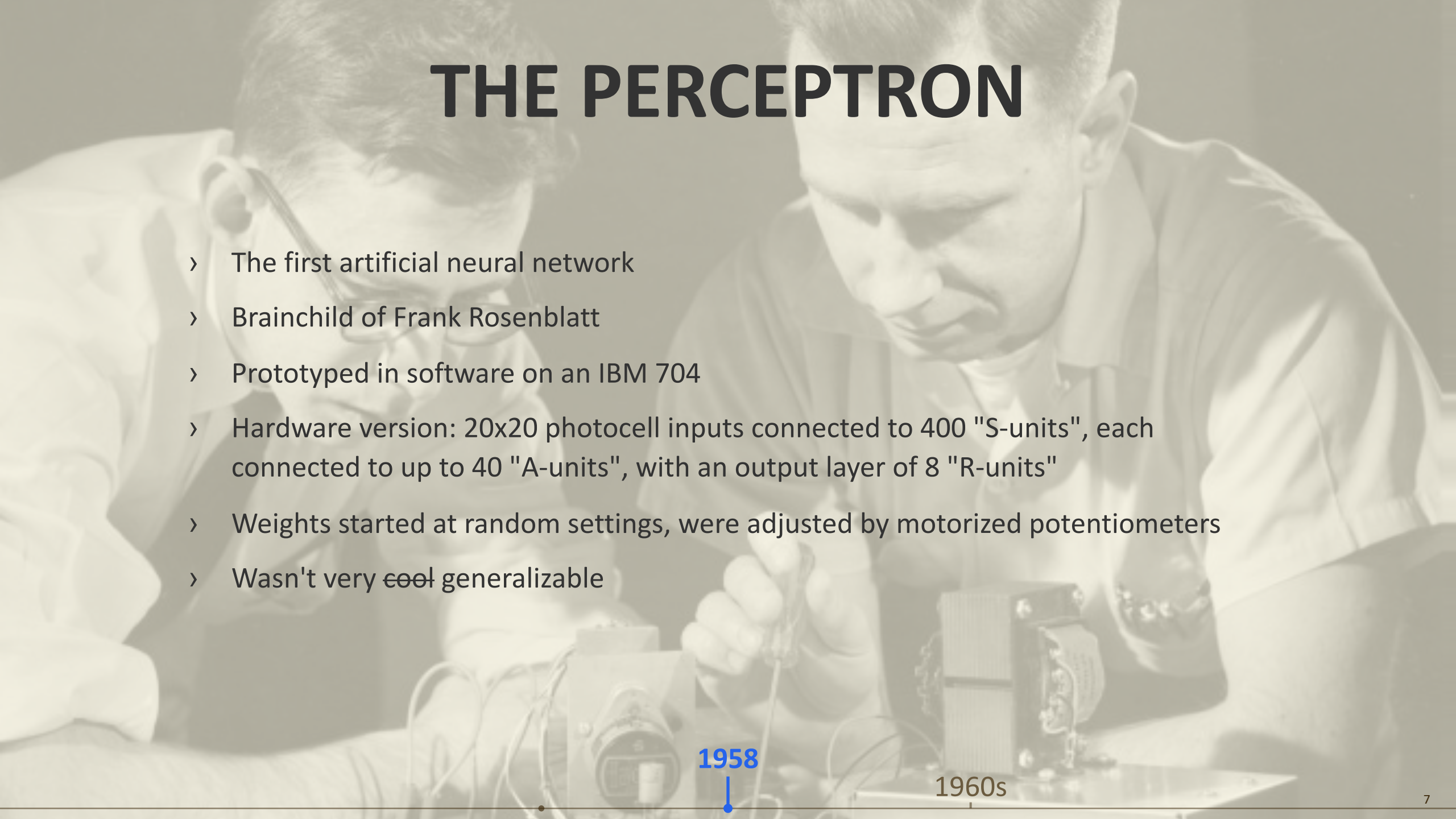
1960s

# LOL

Jun 18, 1956

1960s

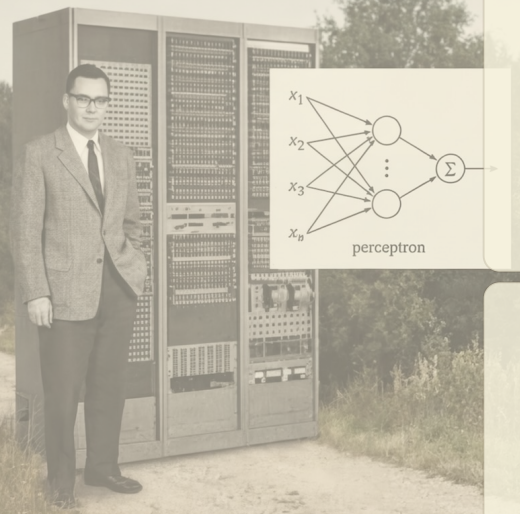
# THE PERCEPTRON



- The first artificial neural network
- Brainchild of Frank Rosenblatt
- Prototyped in software on an IBM 704
- Hardware version: 20x20 photocell inputs connected to 400 "S-units", each connected to up to 40 "A-units", with an output layer of 8 "R-units"
- Weights started at random settings, were adjusted by motorized potentiometers
- Wasn't very ~~cool~~ generalizable

**1958** 1960s

# TWO ROADS DIVERGE



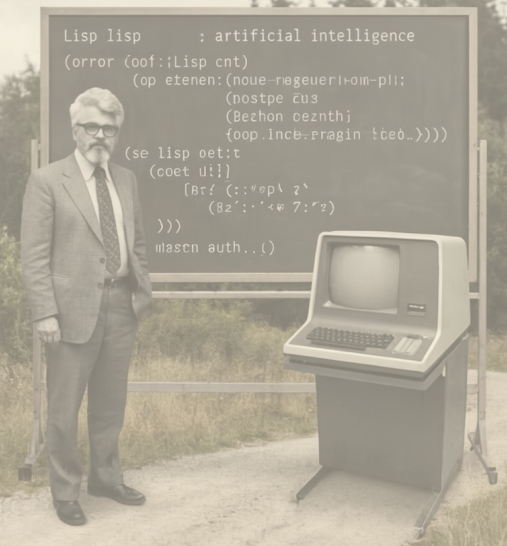



## CONNECTIONIST

Rosenblatt's Perceptron learns from data. Intelligence *emerges* from experience and weight adjustment. The brain as metaphor.

## SYMBOLIC

McCarthy's LISP encodes logic explicitly. Intelligence is *programmed*, not learned. Reason as computation.




1940s 1950s **1958** 1960s 1970s 1980s

# ELIZA




- Joseph Weizenbaum at MIT creates a chatbot
- Plays a Rogerian psychotherapist
- Staff formed emotional bonds with it
- Number of things new under the sun: 0

1966

1970s

# ELIZA



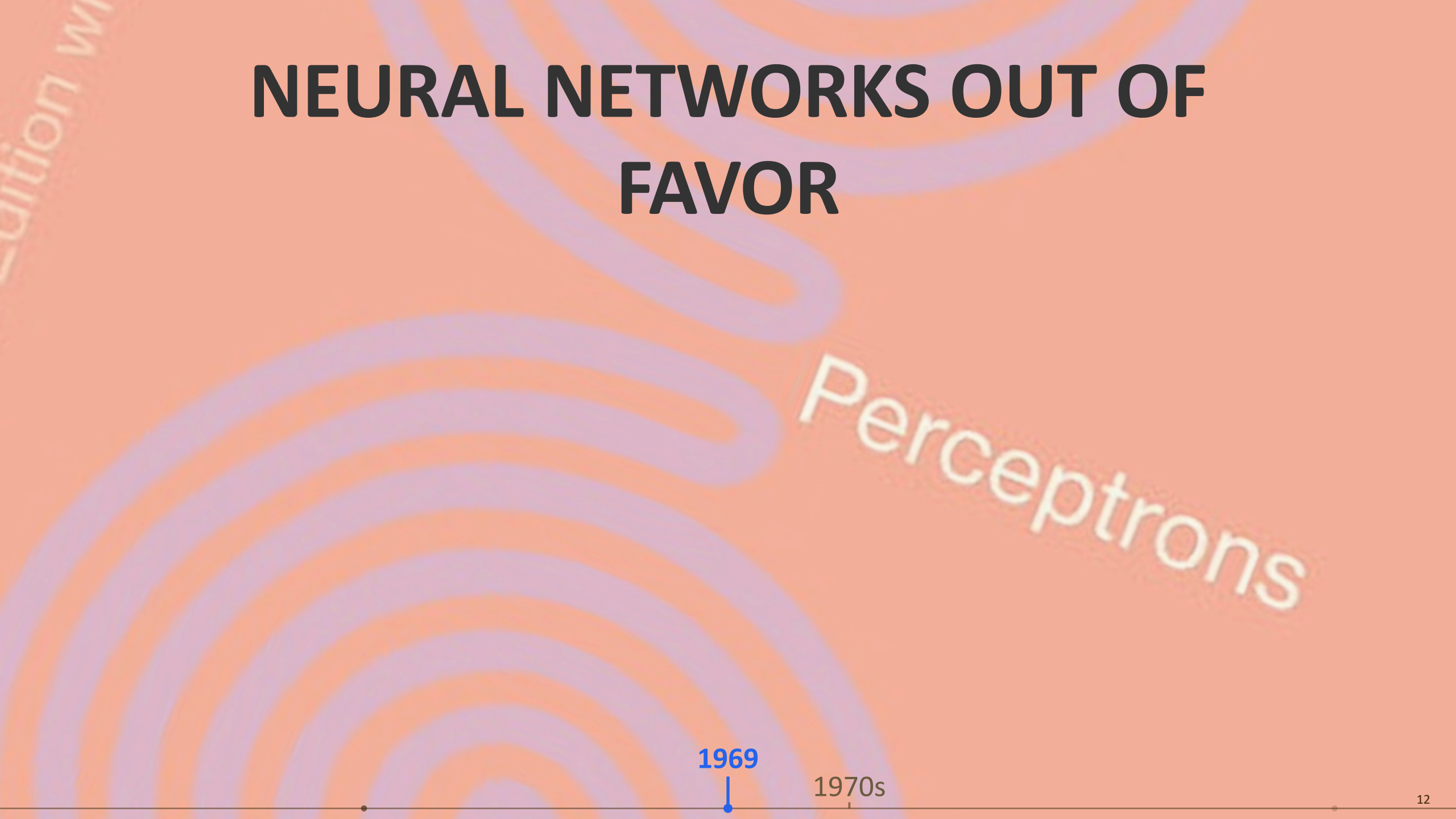

1980s **1981** 1990s

# ELIZA




1960s — **1966** — 1970s

# NEURAL NETWORKS OUT OF FAVOR



# THE FIRST AI WINTER



# AI WAS LAME

1970s — 1974 — 1980s




1980s

1985

1990s

# THE TECHNOLOGY OF EXPERT SYSTEMS

BY ROBERT H. MICHAELSEN, DONALD MICHIE, AND ALBERT BOULANGER

*Transplanting expert knowledge to machines*

THE PURPOSE OF this article is to introduce expert systems. Initially, we'll define these systems. Next, we'll discuss methods for building them, including the advantages and disadvantages of each method. Finally, we'll review the computer resources needed to build and run expert

1980s

analysis is successful, the computer program based on this analysis will be able to solve the narrowly defined problems as well as an expert. (For a discussion of successful expert systems, see reference 2.)

1985

Experts typically solve problems that are unstructured and ill-defined,

ferentiate, pursue and test hypothesis, explore and refine, ask general questions, and make a decision (see reference 11).

Experts are capable of

- Applying their expertise to the solution of problems in an efficient man-

1990s

# WHAT WE NEED IS EXPERTS

- Rule-based systems/nested conditionals until you die
- MYCIN: diagnoses bacterial infections as well as internists 💰
- XCON: configures DEC minicomputers, reportedly saves $40M annually 💰
- There must have been more than two...

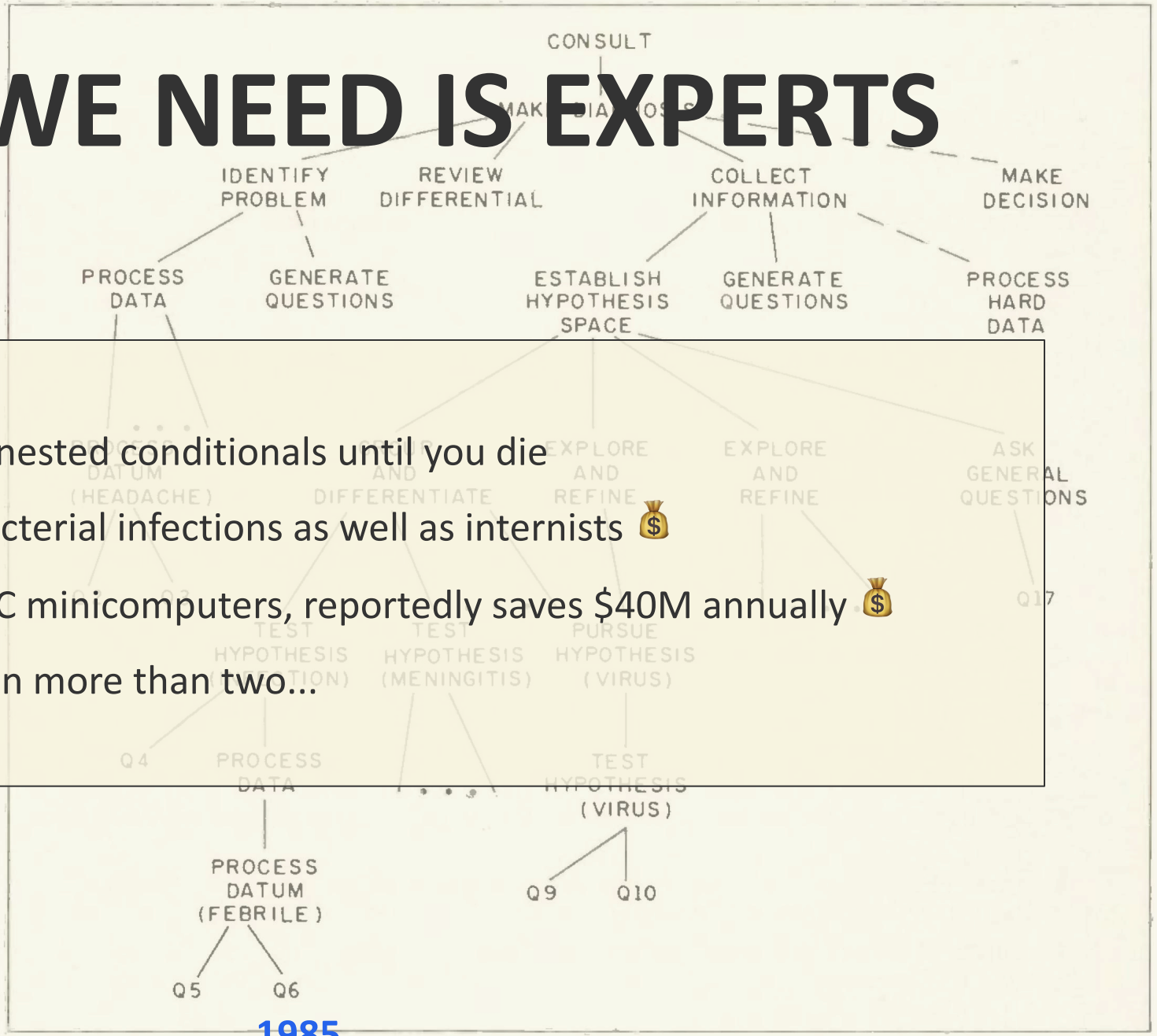
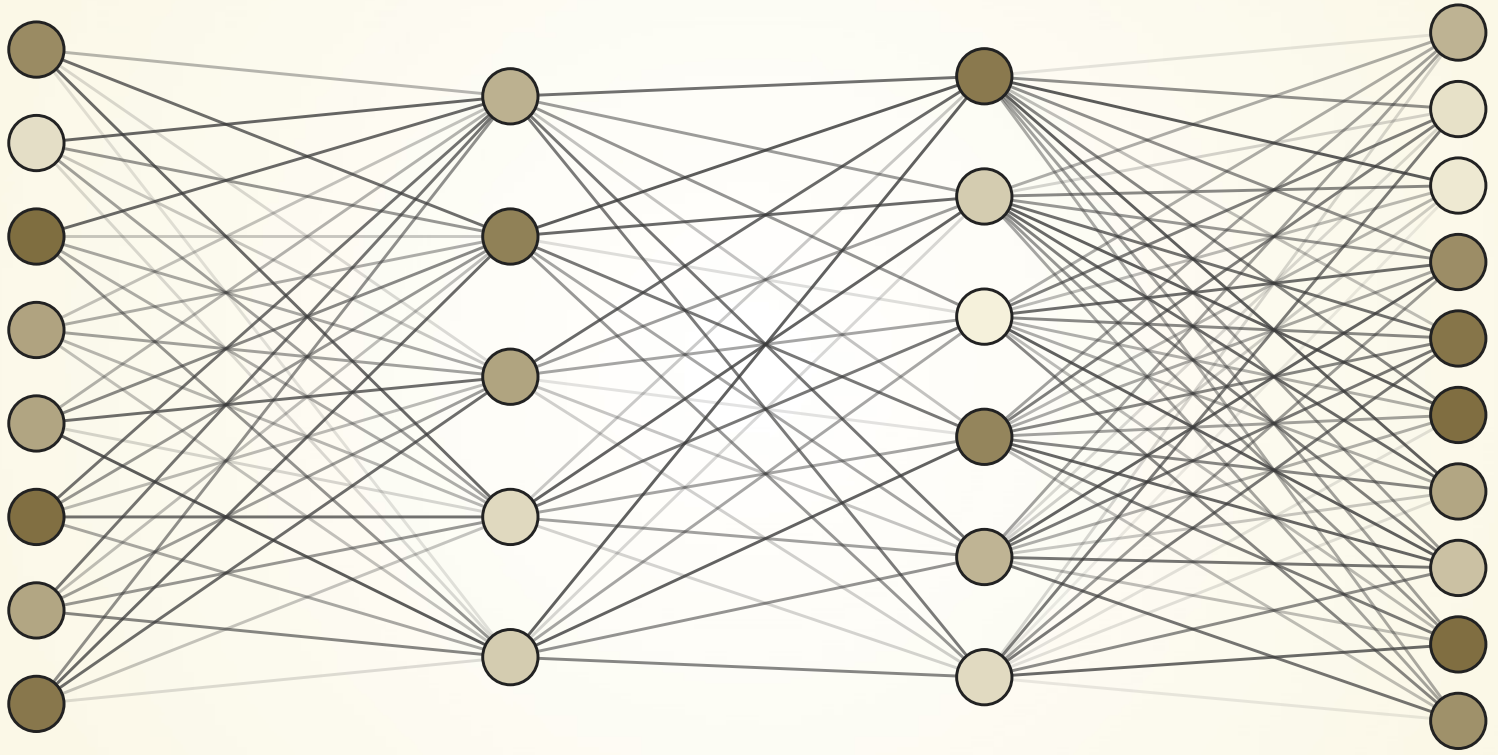



Figure 2: The MYCIN *problem-solving hierarchy. Question numbers (not shown) correspond to questions asked in the consultation. Solid lines show tasks actually done.*

1980s — **1985** — 1990s

# BACKPROPAGATION



1980s — **1986** — 1990s

# THE SECOND AI WINTER



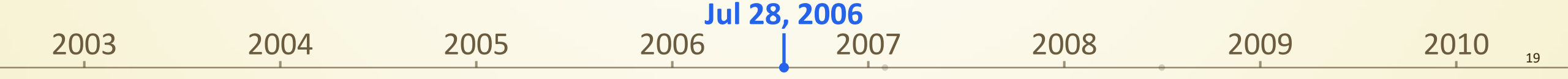
- Nobody wanted Lisp machines
- Expert systems kind of sucked
- "AI" became a marketing liability
- Computer vision? You mean *image processing*
- Geoffrey Hinton continues laboring in obscurity

1987

1990s

# DEEP LEARNING

- *Reducing the Dimensionality of Data with Neural Networks* (Hinton, 2006)
- Showed many-layered networks can be trained layer by layer
- Still regarded as somewhat crazy
- Just add layers, training data, and compute...but from where?



# NVIDIA CUDA

```
__global__ void vectorAdd(float *a, float *b, float *c, int n) {
    int i = blockIdx.x * blockDim.x + threadIdx.x;
    if (i < n)
        c[i] = a[i] + b[i];
}

// Launch with 256 threads per block
vectorAdd<<<(n + 255) / 256, 256>>>(a, b, c, n);
```



# IMAGENET



Jul 2008

2006 2007 2008 2009 2010 2011

# ALEXNET

- ImageNet Large Scale Visual Recognition Challenge (ILSVRC) started in 2010
- It was already kind of lame
- AlexNet (Krizhevsky, Sutskever, Hinton) steps up
- LMAO deep learning, awkward silence
- Trounced the nearest competitor: **15.3%** to **26.2%**
- How? Two NVIDIA GTX 580s and CUDA. In Alex's Mom's basement.
- The Valley caught on fire. The game was afoot.

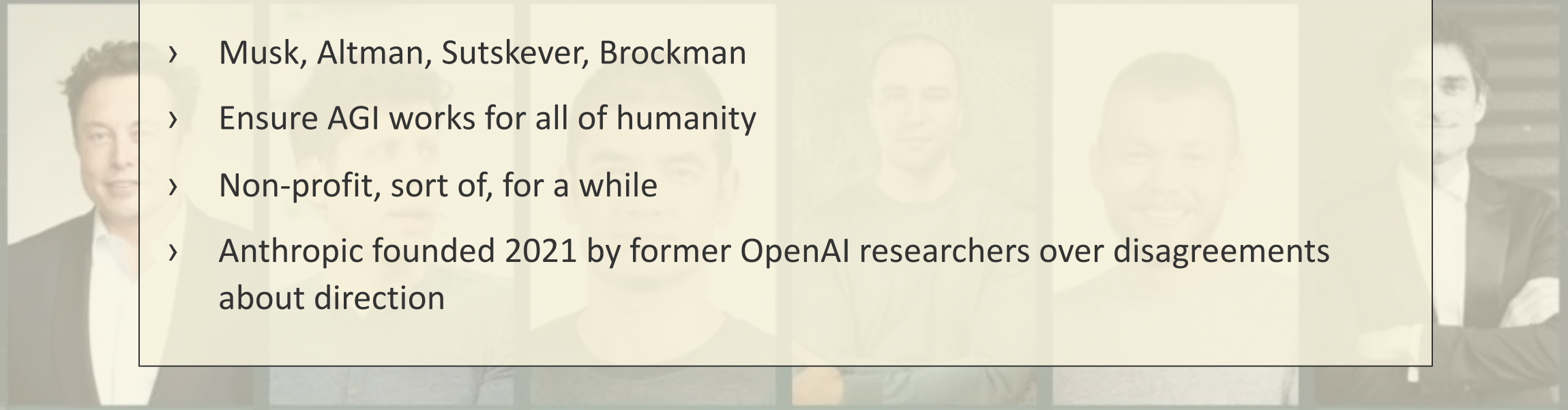


# OPENAI

## Cofounders

The organization was founded in San Francisco in 2015 by Sam Altman, Reid Hoffman, Jessica Livingston, Elon Musk, Ilya Sutskever, Peter Thiel and others, who collectively pledged US$1 billion. Musk resigned from the board in 2018 but remained a donor and eventually committed US$100 million.

- Musk, Altman, Sutskever, Brockman
- Ensure AGI works for all of humanity
- Non-profit, sort of, for a while
- Anthropic founded 2021 by former OpenAI researchers over disagreements about direction



Elon Musk | Sam Altman | Ilya Sutskever | Greg Brockman | Wojciech Zaremba | John Schulman

2014 | **2015** | 2016

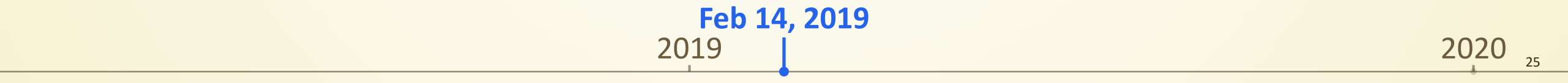
# ATTENTION IS ALL YOU NEED



- Vaswani et al. (Google Brain) introduce the Transformer architecture
- Every token can attend to every other token simultaneously
- Scales linearly with hardware
- Can train on the text of the Internet

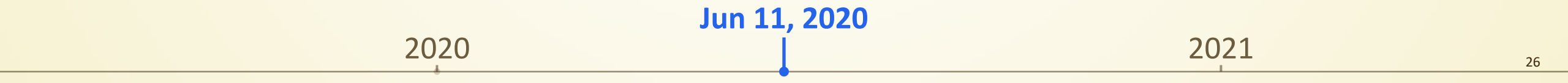
# GPT-2

- 1.5 billion parameters — 10× GPT-1
- "Too Dangerous to Release"
- Coherent, convincing long-form text on almost any topic
- OpenAI stages the release over months, citing misinformation risk
- Bigger is different, not just better



# GPT-3

- 175 billion parameters — 100× GPT-2
- Writes essays, translates languages, answers questions, generates code
- Hallucinates like your uncle who grows mescaline
- Language models are now officially Large



# CHAT GPT

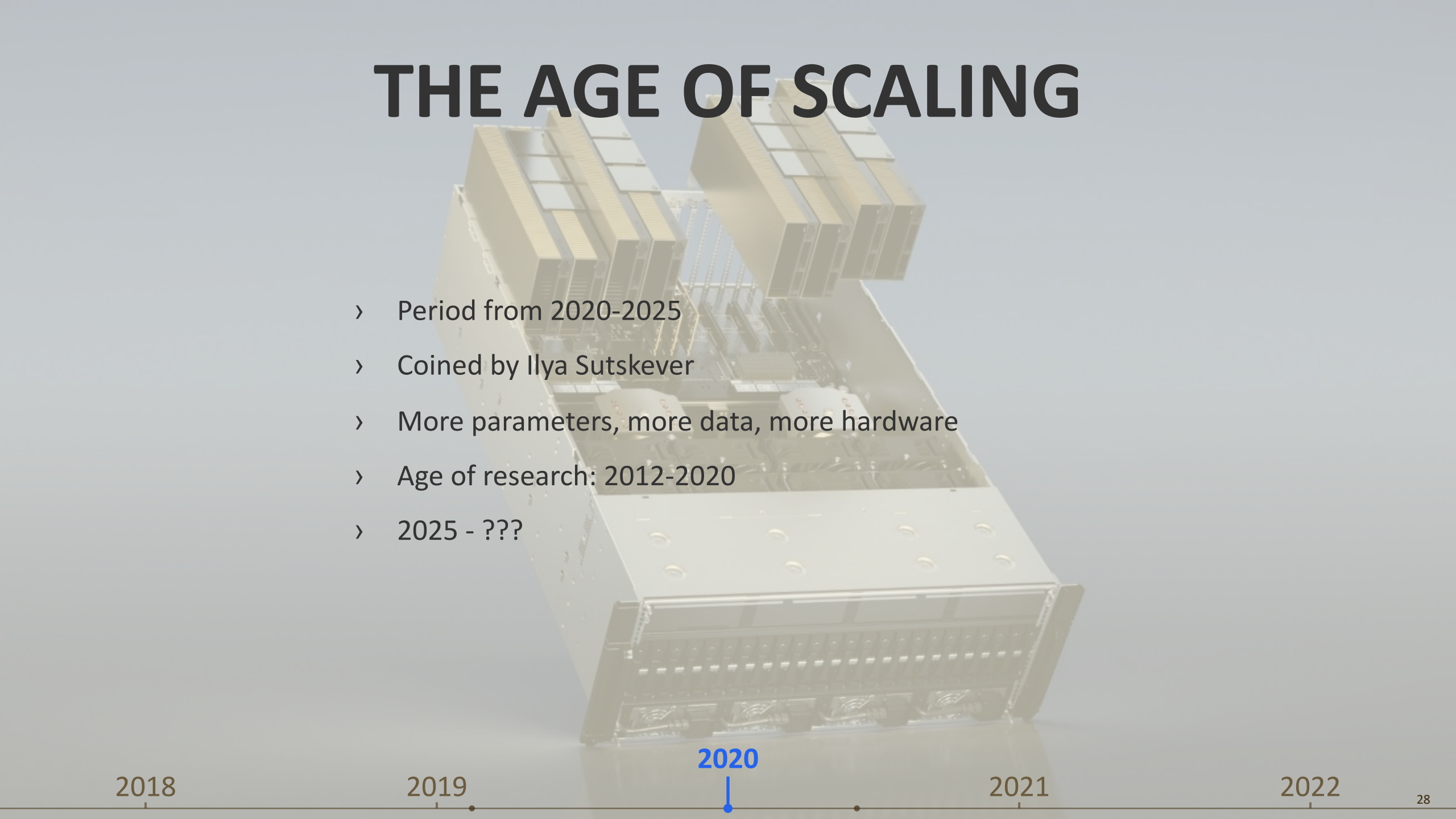


- GPT-3.5 + RLHF
- 100 million users in two months
- The world was never going to be the same
- Tim still wasn't fully buying it

Nov 30, 2022

2022 2023

# THE AGE OF SCALING



- Period from 2020-2025
- Coined by Ilya Sutskever
- More parameters, more data, more hardware
- Age of research: 2012-2020
- 2025 - ???

2018 2019 **2020** 2021 2022

# THE AGENTIC ERA

- Cursor, GitHub Copilot: AI as coding collaborator, not an oracle
- Function calling, tool use, multi-step planning over extended tasks
- Models can browse the web, write and run code, call external APIs

2025

# CLAUDE CODE

- Initial release
- Who cares? Cursor rules the world.
- Terminal-based UI
- Tim is still pretty unimpressed

Feb 2025

2025 — 2026

# OPUS 4.5 + CLAUDE CODE

## WE ALL FELT IT




Nov 24, 2025

2025 2026

# WHAT CAN WE LEARN?

Oct 2, 2026

2026

2027



# YOU'RE KNOWN BY THE COMPANY YOU KEEP

Oct 2, 2026

2026

2027

# STICK WITH YOUR PARADIGM



2026

Oct 2, 2026

2027



# DON'T STICK WITH YOUR PARADIGM

Oct 2, 2026

2026

2027

Model Context Protocol

Claude





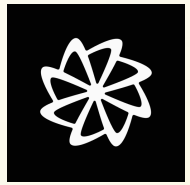


# YOU



opencode

BOSO



Oct 2, 2026

2026

2027

# THANK YOU

@tlberglund

linkedin.com/in/tlberglund

