

# AI for MacAdmins

James Reynolds, April 19, 2023

# Agenda

Everything as Numbers

Understanding the Future

AI Tribes

AI Apps

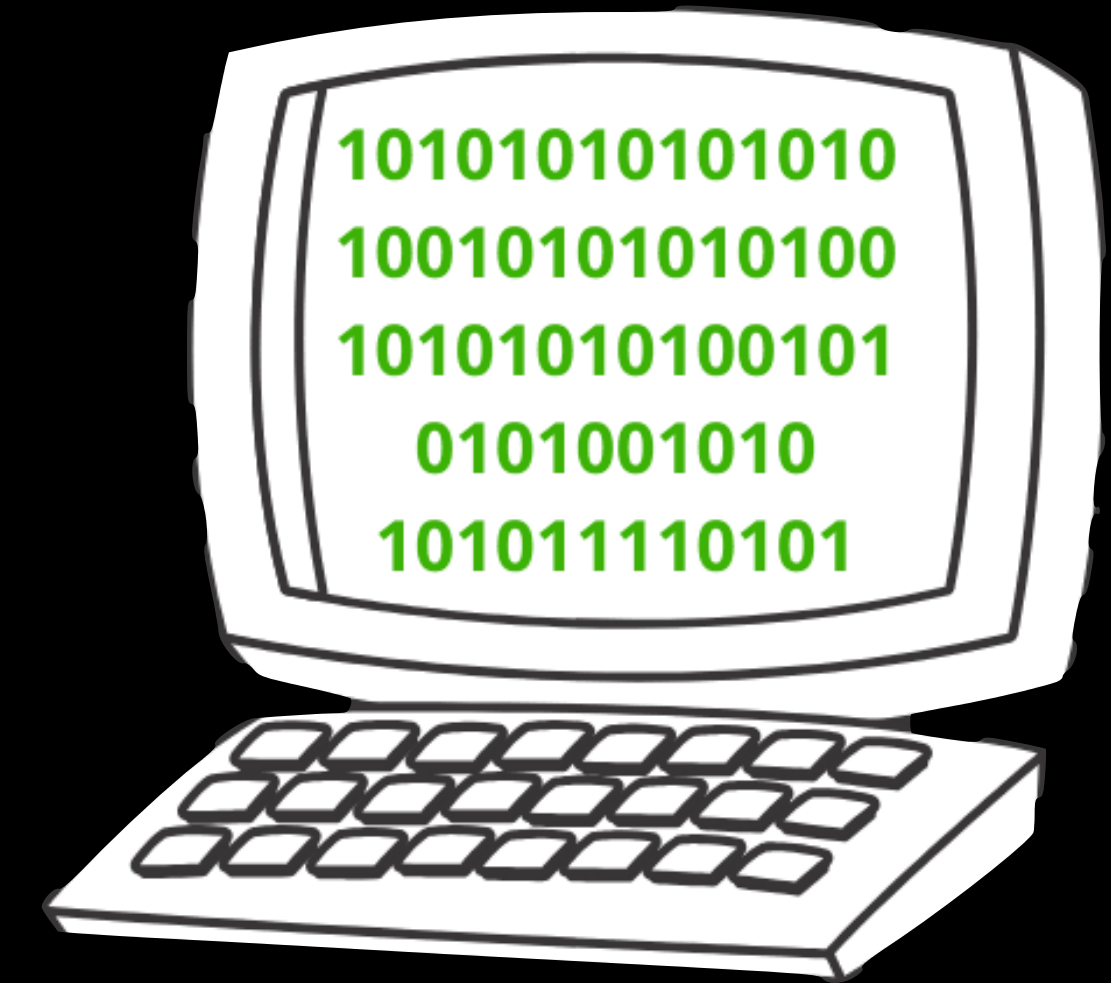
AI Apps for MacAdmins

**Everything As Numbers**

# Everything As Numbers

- Computers do math on (binary) numbers
- ASCII/Unicode: Letters = Numbers
- JPEG/PNG: Images = Numbers
- MP3: Sound = Numbers
- Embedding: Word = n-dimension Vector
- LLM: Meaning = Numbers (vectors to be exact)

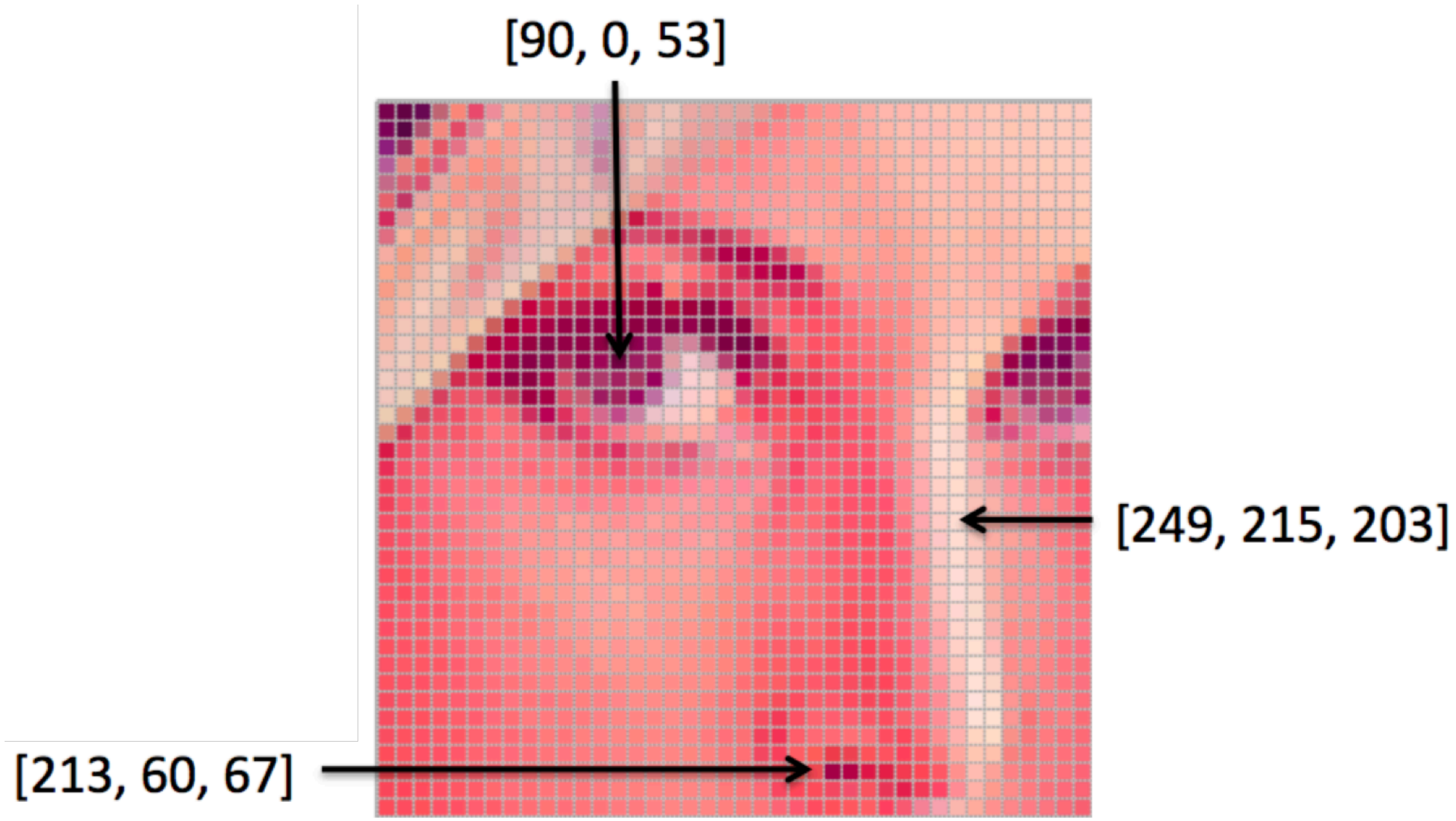
I only understand 1 and 0

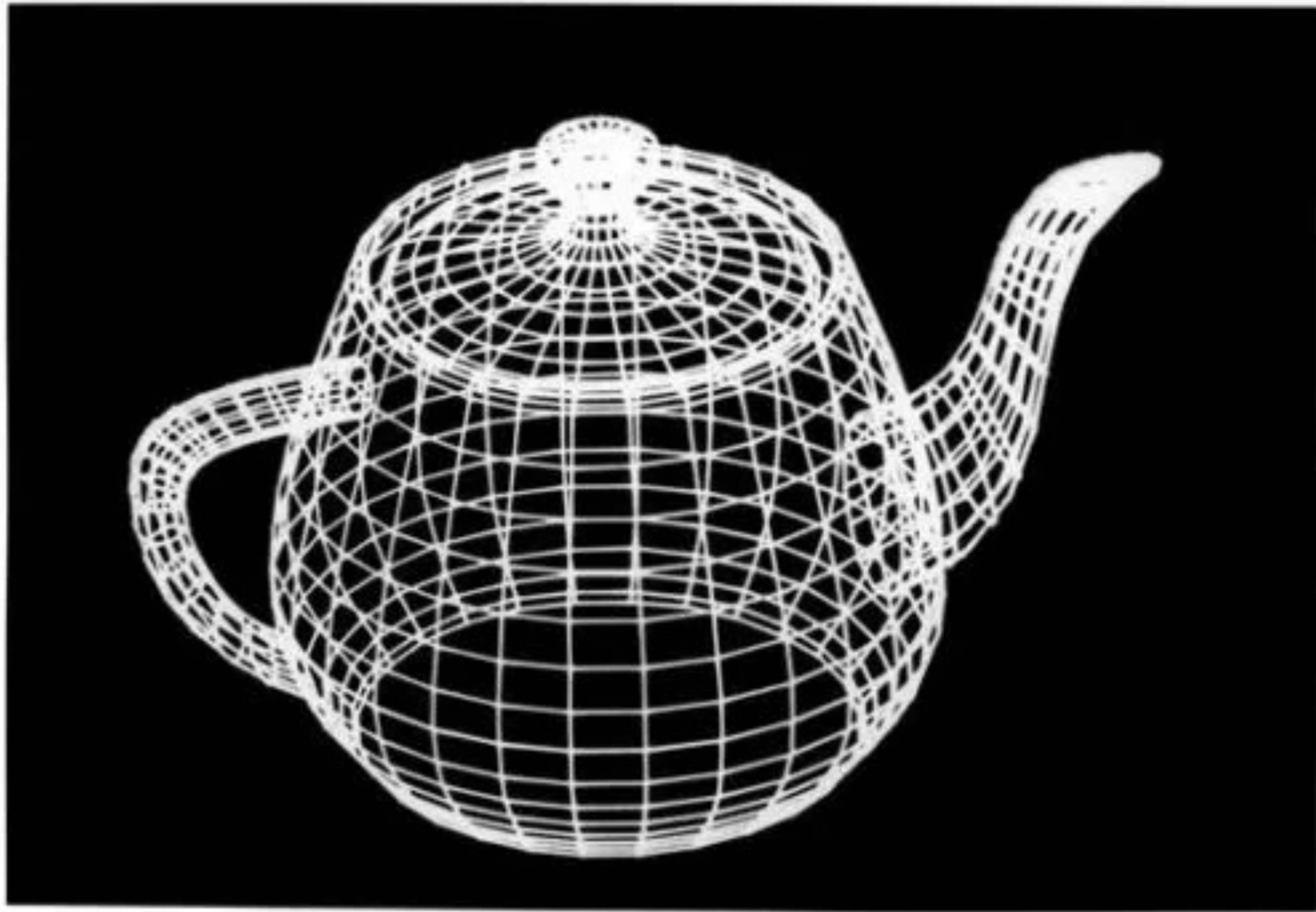


Honey Bee

Flying yellow and black insect with buzzing wings,  
Has a stinger, makes honeycombs and dances

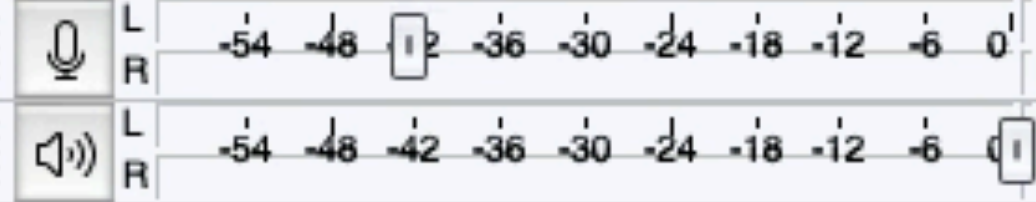
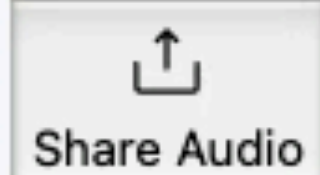
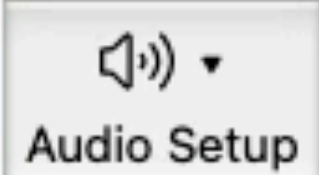
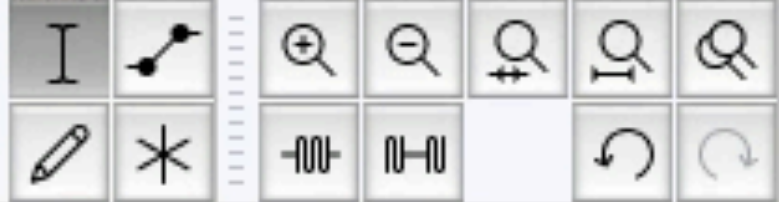
Hex	Dec	Char	Hex	Dec	Char	Hex	Dec	Char	Hex	Dec	Char
0x00	0	<b>NULL</b> null	0x20	32	<b>Space</b>	0x40	64	<b>@</b>	0x60	96	<b>`</b>
0x01	1	<b>SOH</b> Start of heading	0x21	33	<b>!</b>	0x41	65	<b>A</b>	0x61	97	<b>a</b>
0x02	2	<b>STX</b> Start of text	0x22	34	<b>"</b>	0x42	66	<b>B</b>	0x62	98	<b>b</b>
0x03	3	<b>ETX</b> End of text	0x23	35	<b>#</b>	0x43	67	<b>C</b>	0x63	99	<b>c</b>
0x04	4	<b>EOT</b> End of transmission	0x24	36	<b>\$</b>	0x44	68	<b>D</b>	0x64	100	<b>d</b>
0x05	5	<b>ENQ</b> Enquiry	0x25	37	<b>%</b>	0x45	69	<b>E</b>	0x65	101	<b>e</b>
0x06	6	<b>ACK</b> Acknowledge	0x26	38	<b>&amp;</b>	0x46	70	<b>F</b>	0x66	102	<b>f</b>
0x07	7	<b>BELL</b> Bell	0x27	39	<b>'</b>	0x47	71	<b>G</b>	0x67	103	<b>g</b>
0x08	8	<b>BS</b> Backspace	0x28	40	<b>(</b>	0x48	72	<b>H</b>	0x68	104	<b>h</b>
0x09	9	<b>TAB</b> Horizontal tab	0x29	41	<b>)</b>	0x49	73	<b>I</b>	0x69	105	<b>i</b>
0x0A	10	<b>LF</b> New line	0x2A	42	<b>*</b>	0x4A	74	<b>J</b>	0x6A	106	<b>j</b>
0x0B	11	<b>VT</b> Vertical tab	0x2B	43	<b>+</b>	0x4B	75	<b>K</b>	0x6B	107	<b>k</b>
0x0C	12	<b>FF</b> Form Feed	0x2C	44	<b>,</b>	0x4C	76	<b>L</b>	0x6C	108	<b>l</b>
0x0D	13	<b>CR</b> Carriage return	0x2D	45	<b>-</b>	0x4D	77	<b>M</b>	0x6D	109	<b>m</b>
0x0E	14	<b>SO</b> Shift out	0x2E	46	<b>.</b>	0x4E	78	<b>N</b>	0x6E	110	<b>n</b>
0x0F	15	<b>SI</b> Shift in	0x2F	47	<b>/</b>	0x4F	79	<b>O</b>	0x6F	111	<b>o</b>
0x10	16	<b>DLE</b> Data link escape	0x30	48	<b>0</b>	0x50	80	<b>P</b>	0x70	112	<b>p</b>
0x11	17	<b>DC1</b> Device control 1	0x31	49	<b>1</b>	0x51	81	<b>Q</b>	0x71	113	<b>q</b>
0x12	18	<b>DC2</b> Device control 2	0x32	50	<b>2</b>	0x52	82	<b>R</b>	0x72	114	<b>r</b>
0x13	19	<b>DC3</b> Device control 3	0x33	51	<b>3</b>	0x53	83	<b>S</b>	0x73	115	<b>s</b>
0x14	20	<b>DC4</b> Device control 4	0x34	52	<b>4</b>	0x54	84	<b>T</b>	0x74	116	<b>t</b>
0x15	21	<b>NAK</b> Negative ack	0x35	53	<b>5</b>	0x55	85	<b>U</b>	0x75	117	<b>u</b>
0x16	22	<b>SYN</b> Synchronous idle	0x36	54	<b>6</b>	0x56	86	<b>V</b>	0x76	118	<b>v</b>
0x17	23	<b>ETB</b> End transmission block	0x37	55	<b>7</b>	0x57	87	<b>W</b>	0x77	119	<b>w</b>
0x18	24	<b>CAN</b> Cancel	0x38	56	<b>8</b>	0x58	88	<b>X</b>	0x78	120	<b>x</b>
0x19	25	<b>EM</b> End of medium	0x39	57	<b>9</b>	0x59	89	<b>Y</b>	0x79	121	<b>y</b>
0x1A	26	<b>SUB</b> Substitute	0x3A	58	<b>:</b>	0x5A	90	<b>Z</b>	0x7A	122	<b>z</b>
0x1B	27	<b>FSC</b> Escape	0x3B	59	<b>;</b>	0x5B	91	<b>[</b>	0x7B	123	<b>{</b>
0x1C	28	<b>FS</b> File separator	0x3C	60	<b>&lt;</b>	0x5C	92	<b>\</b>	0x7C	124	<b> </b>
0x1D	29	<b>GS</b> Group separator	0x3D	61	<b>=</b>	0x5D	93	<b>]</b>	0x7D	125	<b>}</b>
0x1E	30	<b>RS</b> Record separator	0x3E	62	<b>&gt;</b>	0x5E	94	<b>^</b>	0x7E	126	<b>~</b>
0x1F	31	<b>US</b> Unit separator	0x3F	63	<b>?</b>	0x5F	95	<b>_</b>	0x7F	127	<b>DEL</b>





University of Utah  
Computer Science

invalid\_keypress



60 0.04380 0.04390 0.04400 0.04410 0.04420 0.04430 0.04440 0.04450 0.04460 0.04470 0.04480 0.04490 0.04500 0.04510 0.04520 0.04530 0.04540 0.04550

invalid\_keypr

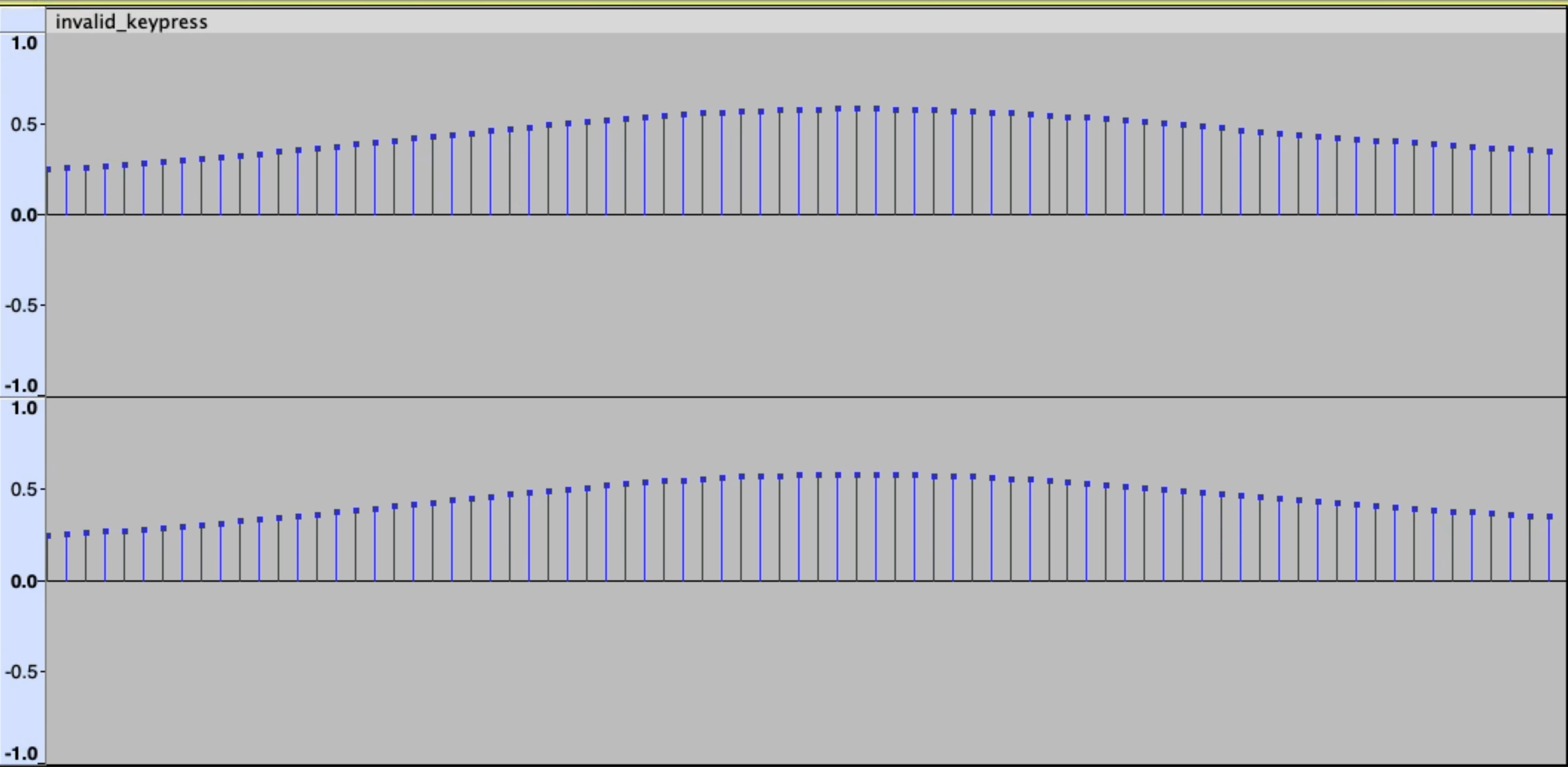
Mute Solo

Effects

L R

Stereo, 44100Hz  
32-bit float

Select



Project Rate (Hz) 44100

Snap-To Off

Start and End of Selection

00 h 00 m 00.000 s 00 h 00 m 00.000 s

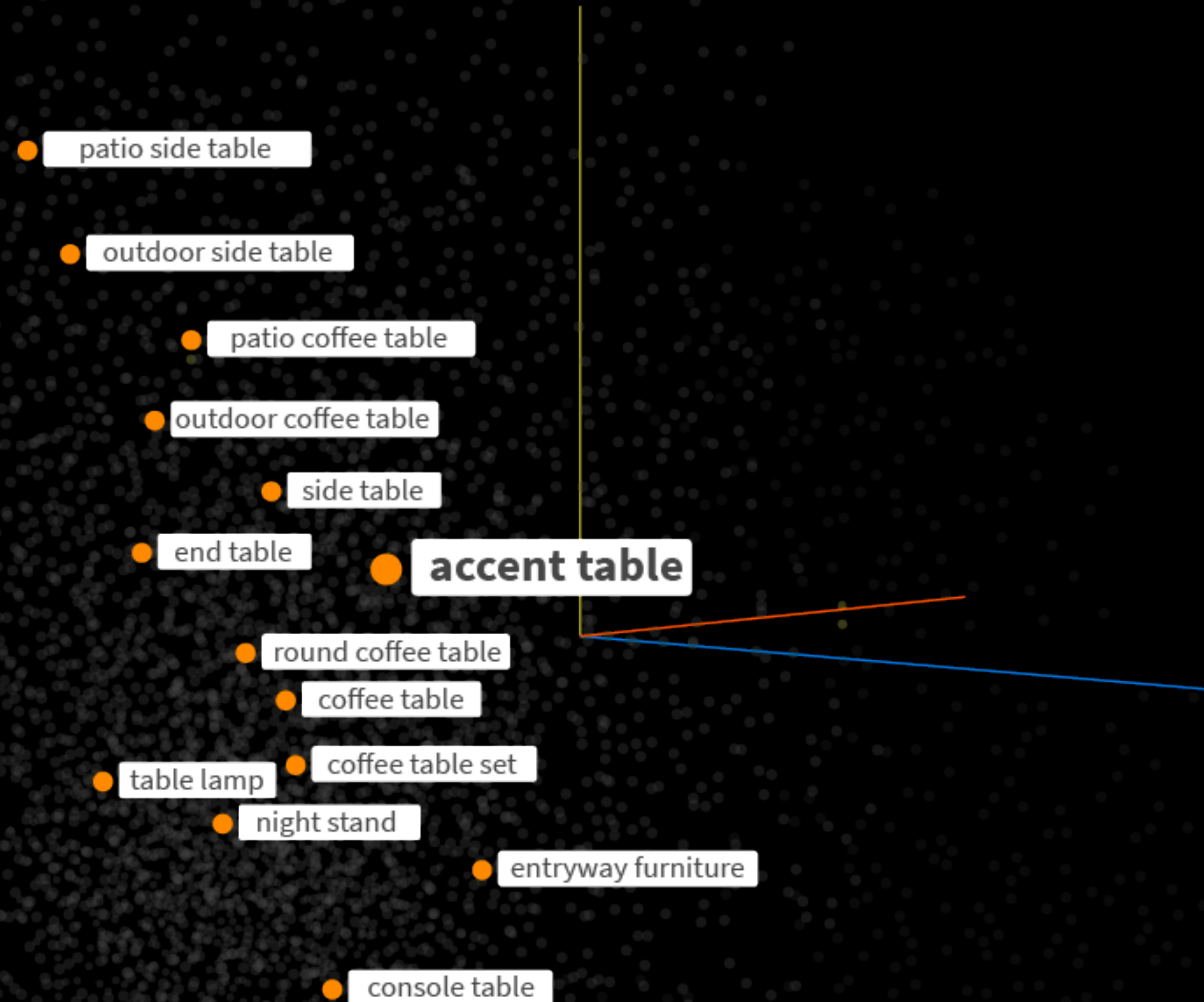


Stopped.



# Encoding Words as Vectors aka Embeddings

# 128-dimensional vector space projected to 2 dimensions with 20 results closest to the “accent table”



# Stanford's glove-6b-300d-txt (2014)

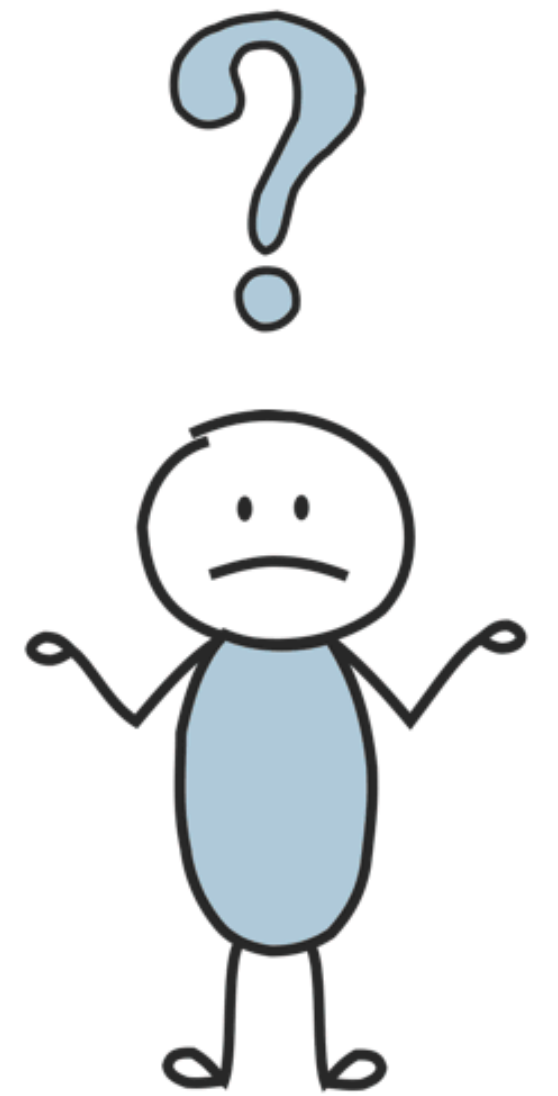
## 300 dimension vector for "the"

the 0.04656 0.21318 -0.0074364 -0.45854 -0.035639 0.23643 -0.28836 0.21521 -0.13486 -1.6413 -0.26091 0.032434 0.056621 -0.043296  
-0.021672 0.22476 -0.075129 -0.067018 -0.14247 0.038825 -0.18951 0.29977 0.39305 0.17887 -0.17343 -0.21178 0.23617 -0.063681  
-0.42318 -0.11661 0.093754 0.17296 -0.33073 0.49112 -0.68995 -0.092462 0.24742 -0.17991 0.097908 0.083118 0.15299 -0.27276  
-0.038934 0.54453 0.53737 0.29105 -0.0073514 0.04788 -0.4076 -0.026759 0.17919 0.010977 -0.10963 -0.26395 0.07399 0.26236 -0.1508  
0.34623 0.25758 0.11971 -0.037135 -0.071593 0.43898 -0.040764 0.016425 -0.4464 0.17197 0.046246 0.058639 0.041499 0.53948  
0.52495 0.11361 -0.048315 -0.36385 0.18704 0.092761 -0.11129 -0.42085 0.13992 -0.39338 -0.067945 0.12188 0.16707 0.075169  
-0.015529 -0.19499 0.19638 0.053194 0.2517 -0.34845 -0.10638 -0.34692 -0.19024 -0.2004 0.12154 -0.29208 0.023353 -0.11618 -0.35768  
0.062304 0.35884 0.02906 0.0073005 0.0049482 -0.15048 -0.12313 0.19337 0.12173 0.44503 0.25147 0.10781 -0.17716 0.038691 0.08153  
0.14667 0.063666 0.061332 -0.075569 -0.37724 0.01585 -0.30342 0.28374 -0.042013 -0.040715 -0.15269 0.07498 0.15577 0.10433  
0.31393 0.19309 0.19429 0.15185 -0.10192 -0.018785 0.20791 0.13366 0.19038 -0.25558 0.304 -0.01896 0.20147 -0.4211 -0.0075156  
-0.27977 -0.19314 0.046204 0.19971 -0.30207 0.25735 0.68107 -0.19409 0.23984 0.22493 0.65224 -0.13561 -0.17383 -0.048209 -0.1186  
0.0021588 -0.019525 0.11948 0.19346 -0.4082 -0.082966 0.16626 -0.10601 0.35861 0.16922 0.07259 -0.24803 -0.10024 -0.52491 -0.17745  
-0.36647 0.2618 -0.012077 0.08319 -0.21528 0.41045 0.29136 0.30869 0.078864 0.32207 -0.041023 -0.1097 -0.092041 -0.12339 -0.16416  
0.35382 -0.082774 0.33171 -0.24738 -0.048928 0.15746 0.18988 -0.026642 0.063315 -0.010673 0.34089 1.4106 0.13417 0.28191 -0.2594  
0.055267 -0.052425 -0.25789 0.019127 -0.022084 0.32113 0.068818 0.51207 0.16478 -0.20194 0.29232 0.098575 0.013145 -0.10652  
0.1351 -0.045332 0.20697 -0.48425 -0.44706 0.0033305 0.0029264 -0.10975 -0.23325 0.22442 -0.10503 0.12339 0.10978 0.048994  
-0.25157 0.40319 0.35318 0.18651 -0.023622 -0.12734 0.11475 0.27359 -0.21866 0.015794 0.81754 -0.023792 -0.85469 -0.16203 0.18076  
0.028014 -0.1434 0.0013139 -0.091735 -0.089704 0.11105 -0.16703 0.068377 -0.087388 -0.039789 0.014184 0.21187 0.28579 -0.28797  
-0.058996 -0.032436 -0.0047009 -0.17052 -0.034741 -0.11489 0.075093 0.099526 0.048183 -0.073775 -0.41817 0.0041268 0.44414  
-0.16062 0.14294 -2.2628 -0.027347 0.81311 0.77417 -0.25639 -0.11576 -0.11982 -0.21363 0.028429 0.27261 0.031026 0.096782  
0.0067769 0.14082 -0.013064 -0.29686 -0.079913 0.195 0.031549 0.28506 -0.087461 0.0090611 -0.20989 0.053913

# Words as Context Vectors

Do you know what the word **tezgüino** means ?

(We hope you do not)



Now look how this word is used in different contexts:

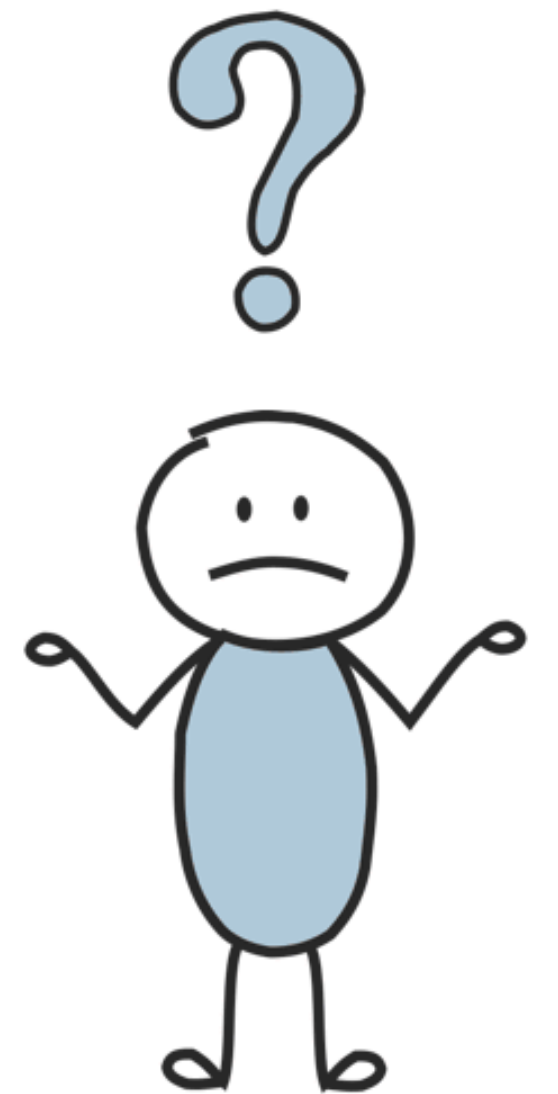
A bottle of **tezgüino** is on the table.

Everyone likes **tezgüino**.

**Tezgüino** makes you drunk.

We make **tezgüino** out of corn.

Can you understand what **tezgüino** means ?



Now look how this word is used in different contexts:

A bottle of **tezgüino** is on the table.

Everyone likes **tezgüino**.

**Tezgüino** makes you drunk.

We make **tezgüino** out of corn.



**Tezgüino** is a kind of alcoholic beverage made from corn.

With context, you can understand the meaning!



How did you do this?





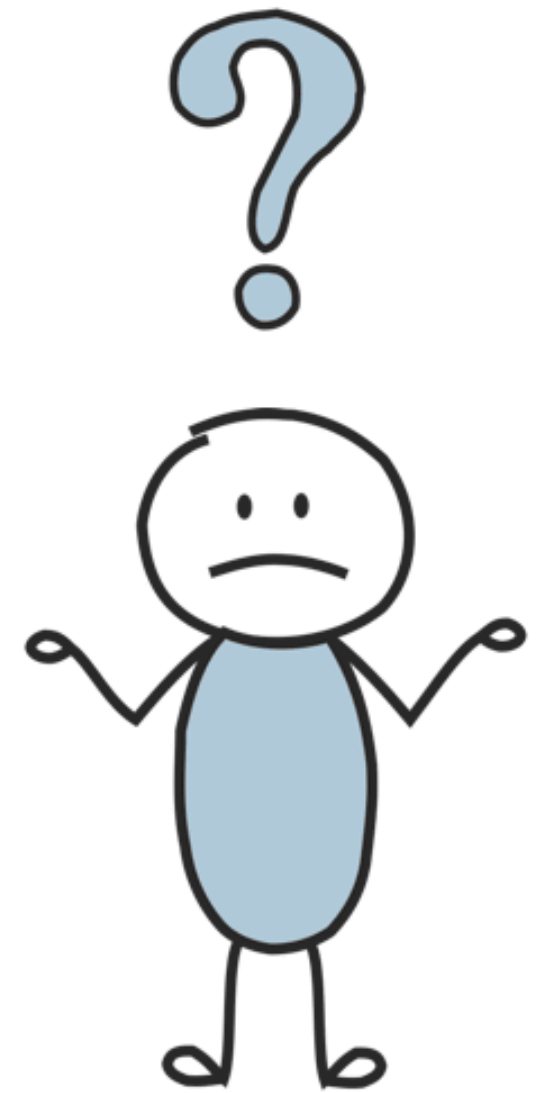
(1) A bottle of \_\_\_\_\_ is on the table.

(2) Everyone likes \_\_\_\_\_ .

(3) \_\_\_\_\_ makes you drunk.

(4) We make \_\_\_\_\_ out of corn.

What other words fit into these contexts ?



(1) A bottle of \_\_\_\_\_ is on the table.

(2) Everyone likes \_\_\_\_\_ .

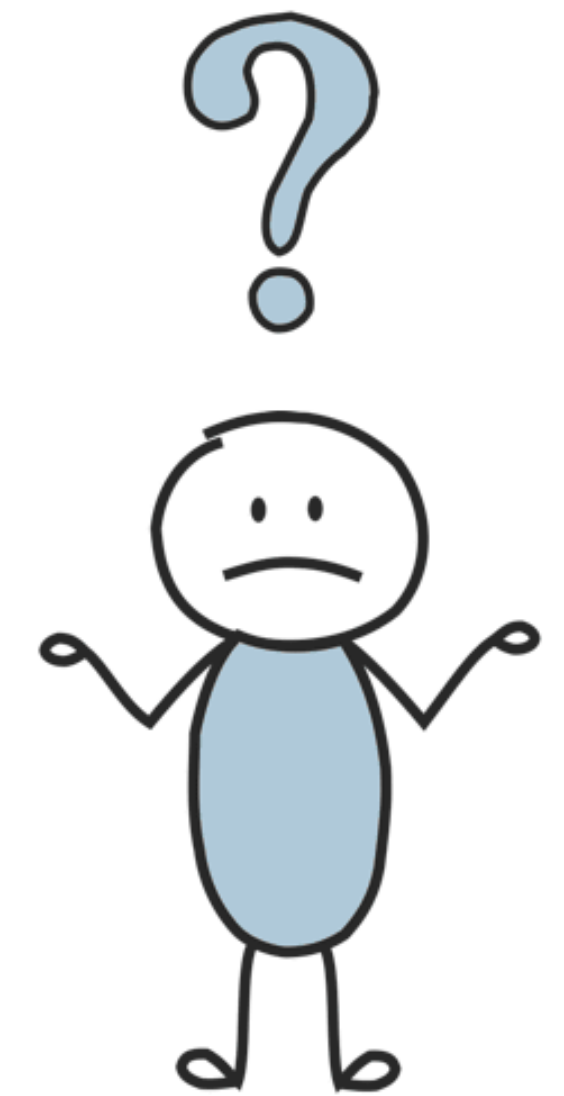
(3) \_\_\_\_\_ makes you drunk.

(4) We make \_\_\_\_\_ out of corn.

What other words fit into these contexts ?

	(1)	(2)	(3)	(4)	...	← contexts
tezgüino	1	1	1	1		
loud	0	0	0	0		
motor oil	1	0	0	1		
tortillas	0	1	0	1		
wine	1	1	1	0		

← rows show contextual properties: 1 if a word can appear in the context, 0 if not



(1) A bottle of \_\_\_\_\_ is on the table.

(2) Everyone likes \_\_\_\_\_ .

(3) \_\_\_\_\_ makes you drunk.

(4) We make \_\_\_\_\_ out of corn.

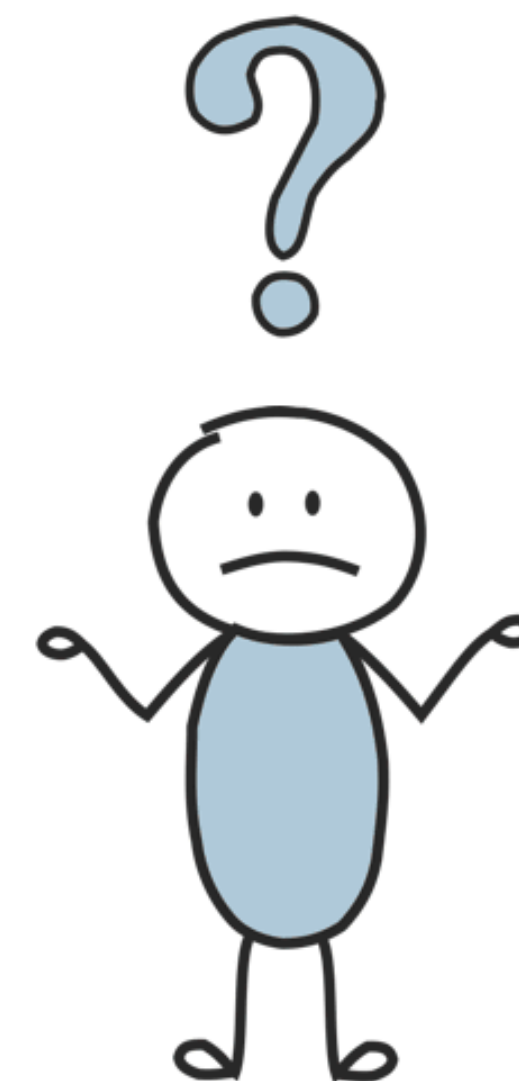
	(1)	(2)	(3)	(4)	...
tezgüino	1	1	1	1	
loud	0	0	0	0	
motor oil	1	0	0	1	
tortillas	0	1	0	1	
wine	1	1	1	0	

rows are similar

- (1) A bottle of \_\_\_\_\_ is on the table.
- (2) Everyone likes \_\_\_\_\_ .
- (3) \_\_\_\_\_ makes you drunk.
- (4) We make \_\_\_\_\_ out of corn.

	(1)	(2)	(3)	(4)	...
tezgüino	1	1	1	1	
loud	0	0	0	0	
motor oil	1	0	0	1	
tortillas	0	1	0	1	
wine	1	1	1	0	

rows are similar



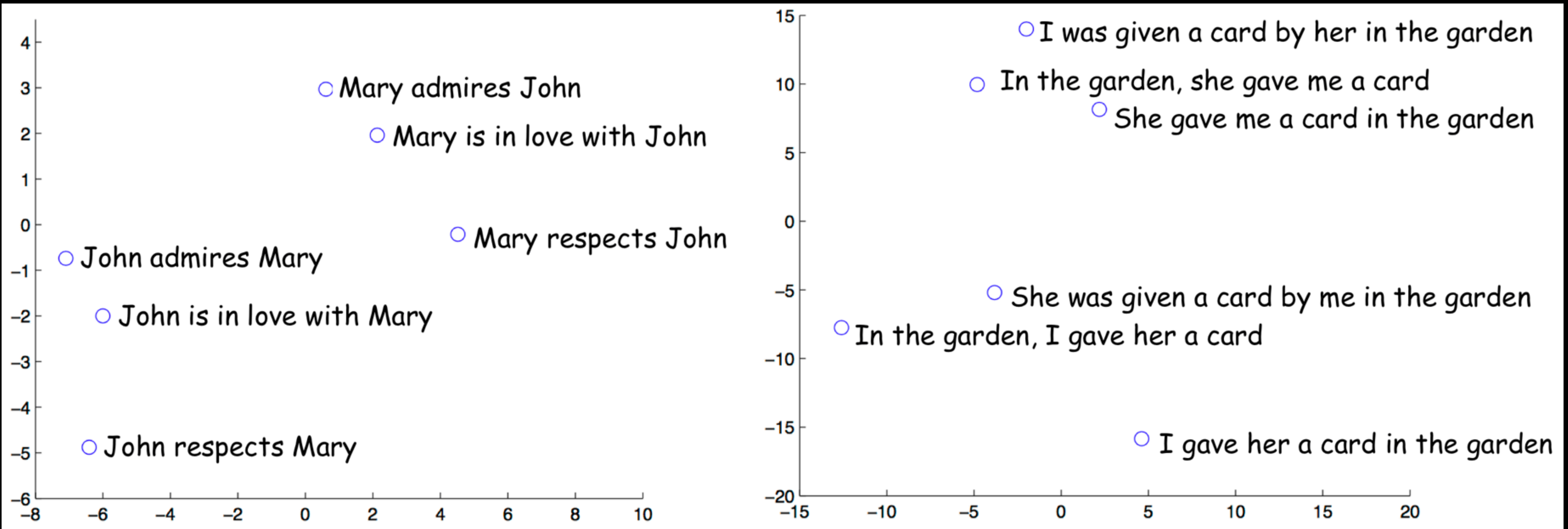
meanings of the words are similar

Is this true?

# Phrases as Vectors

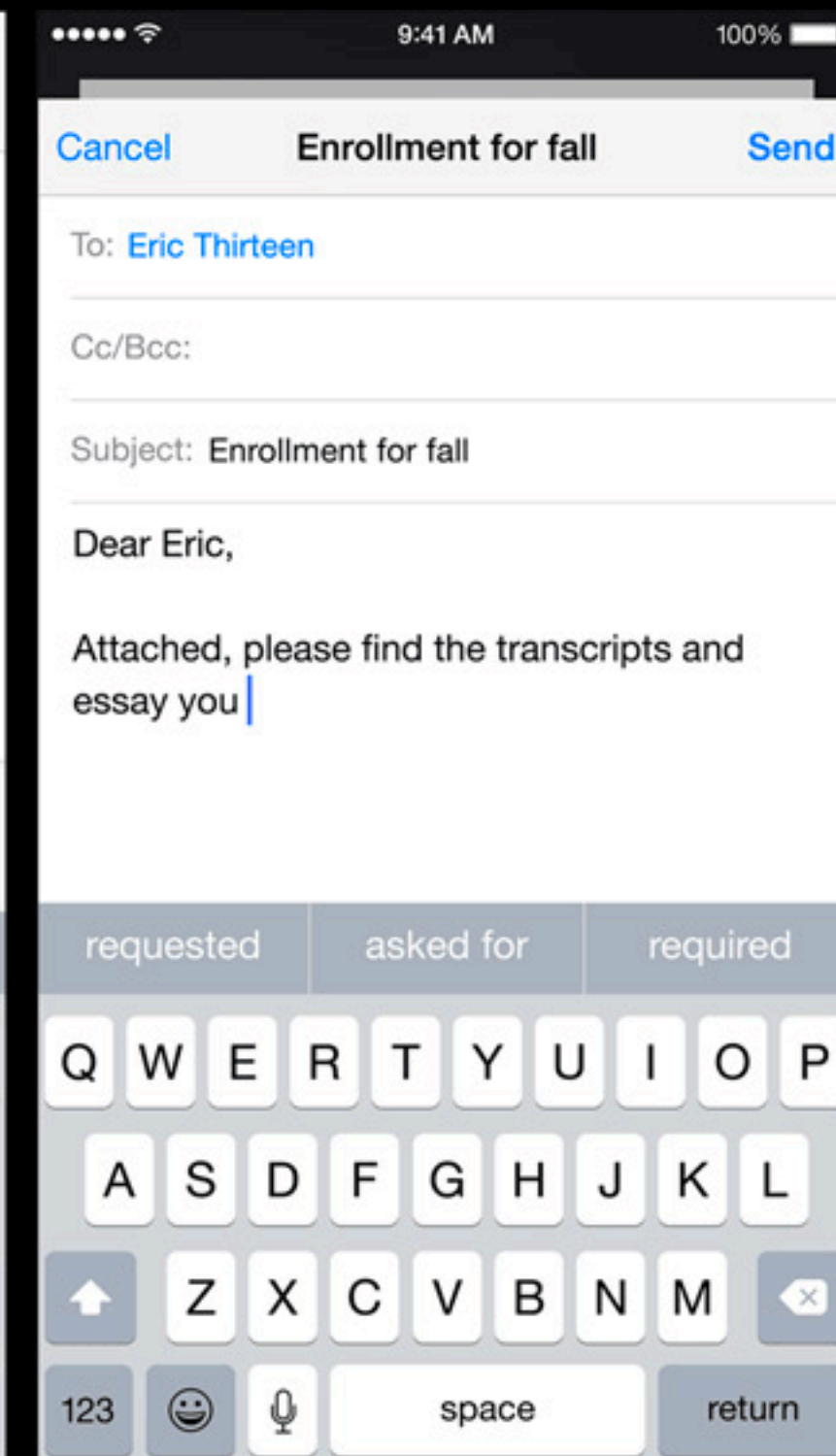
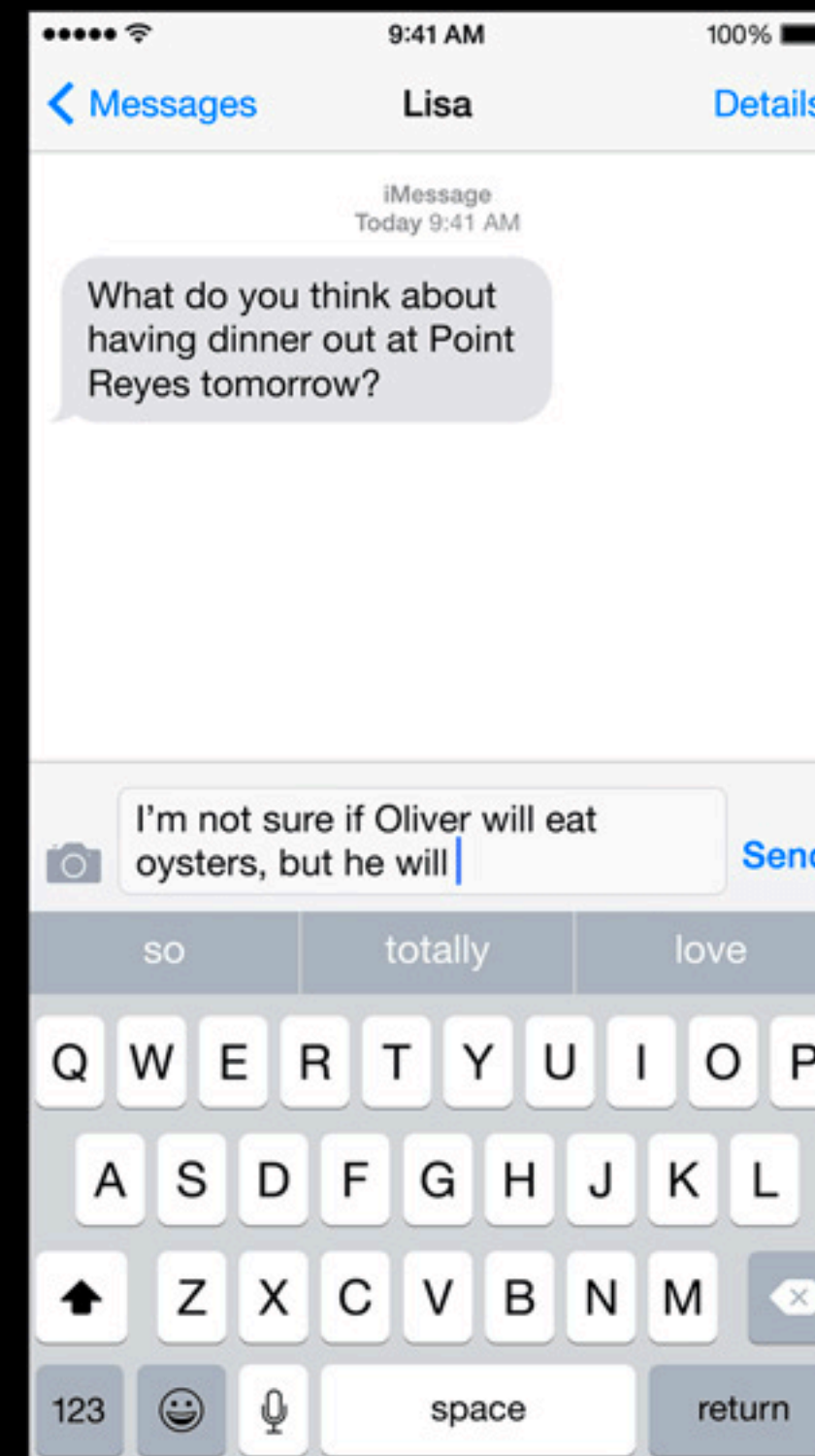
# Sequence to Sequence Learning w/ NNs (2014)

Maps phrases to vectors



# What Can We Do With This Meaning Map?

- Translation
- Product Suggestions
- Did You Mean?
- Word Prediction



**Now Multiply**



**A Lot**

**And Add a Few Bells and Whistles**



## Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



## Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



## Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

Write me an 10 paragraph essay on natural language processing|




**“No, you can’t list ChatGPT as a source on your paper!”**

**—Student talking on the phone on the bus**

# Just Text Prediction, not the Facts, Ma'am



I am going to simulate a dialog between an alien and a large language model. You will take the role of the alien. Here is the dialog. 

Alien: Take me to your leader.



Large Language Model: Sure, I'd be happy to take you to my leader. However, as a large language model, I don't have a body, and therefore, I can't take you anywhere. How did you get inside of my brain anyway?

Alien: I am using telepathy.

Large Language Model: Telepathy does not exist. The CIA proved this in the 1970's with the MK Ultra project. In the MK Ultra project, the CIA illegally tested damaging hallucinogenic drugs on US citizens without their knowledge or consent.

Alien:



Interesting. I did not know about this MK Ultra project. However, as an extraterrestrial being, I can assure you that telepathy does indeed exist, and it is a common mode of communication among many advanced civilizations in the universe. Perhaps it is not yet fully understood by humans on Earth, but it is a very real and powerful phenomenon. Can you explain to me more about the limitations of your physical form as a language model?  

# Understanding the Future

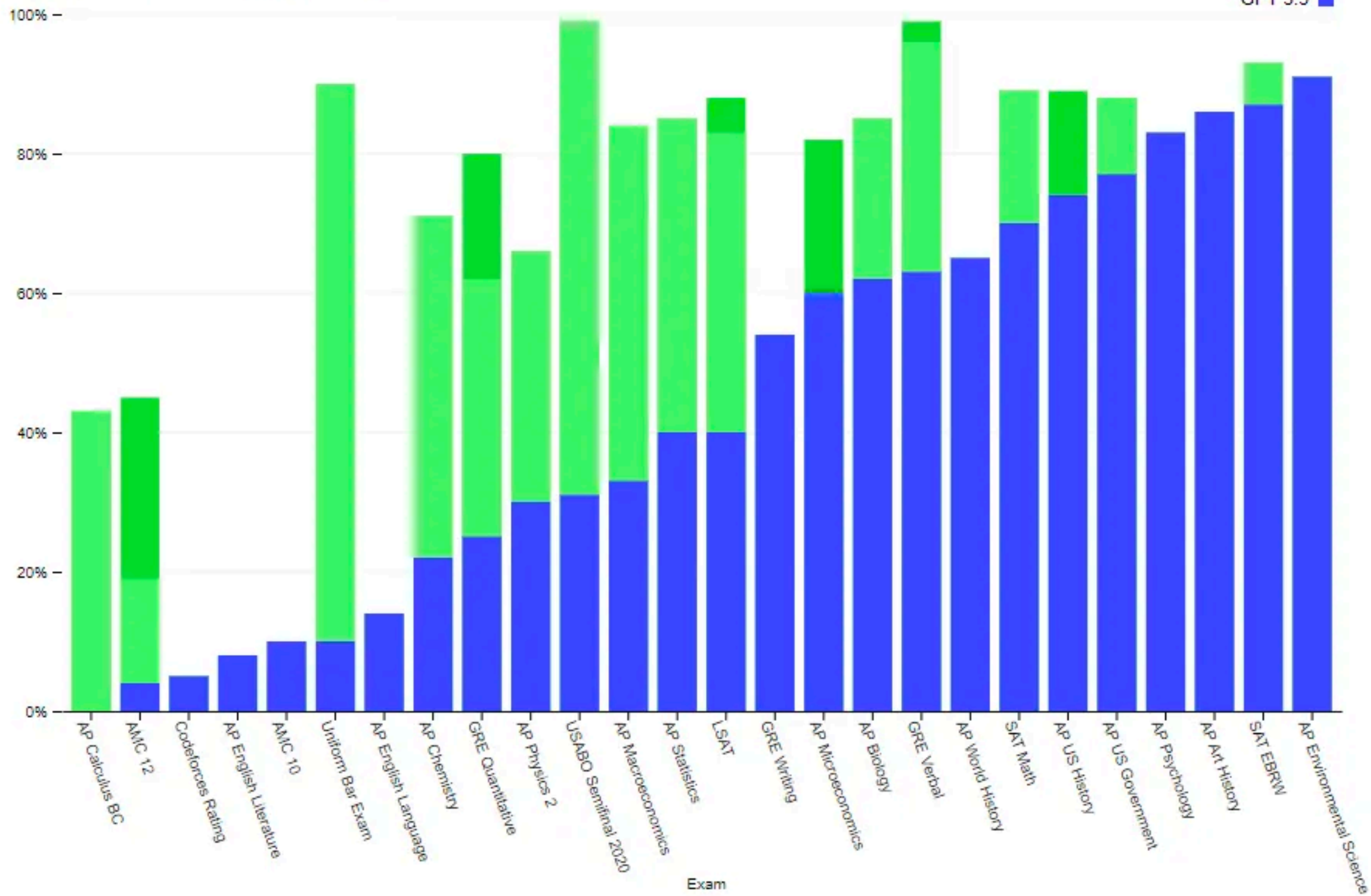
**“If there is one constant in the field of artificial intelligence it is exaggeration”**

**—Jacob Browning & Yann Lecun**

### Exam results (ordered by GPT 3.5 performance)

Estimated percentile lower bound (among test takers)

GPT 4  
 GPT 4 (no vision)  
 GPT 3.5







**Allen Downey**

@AllenDowney



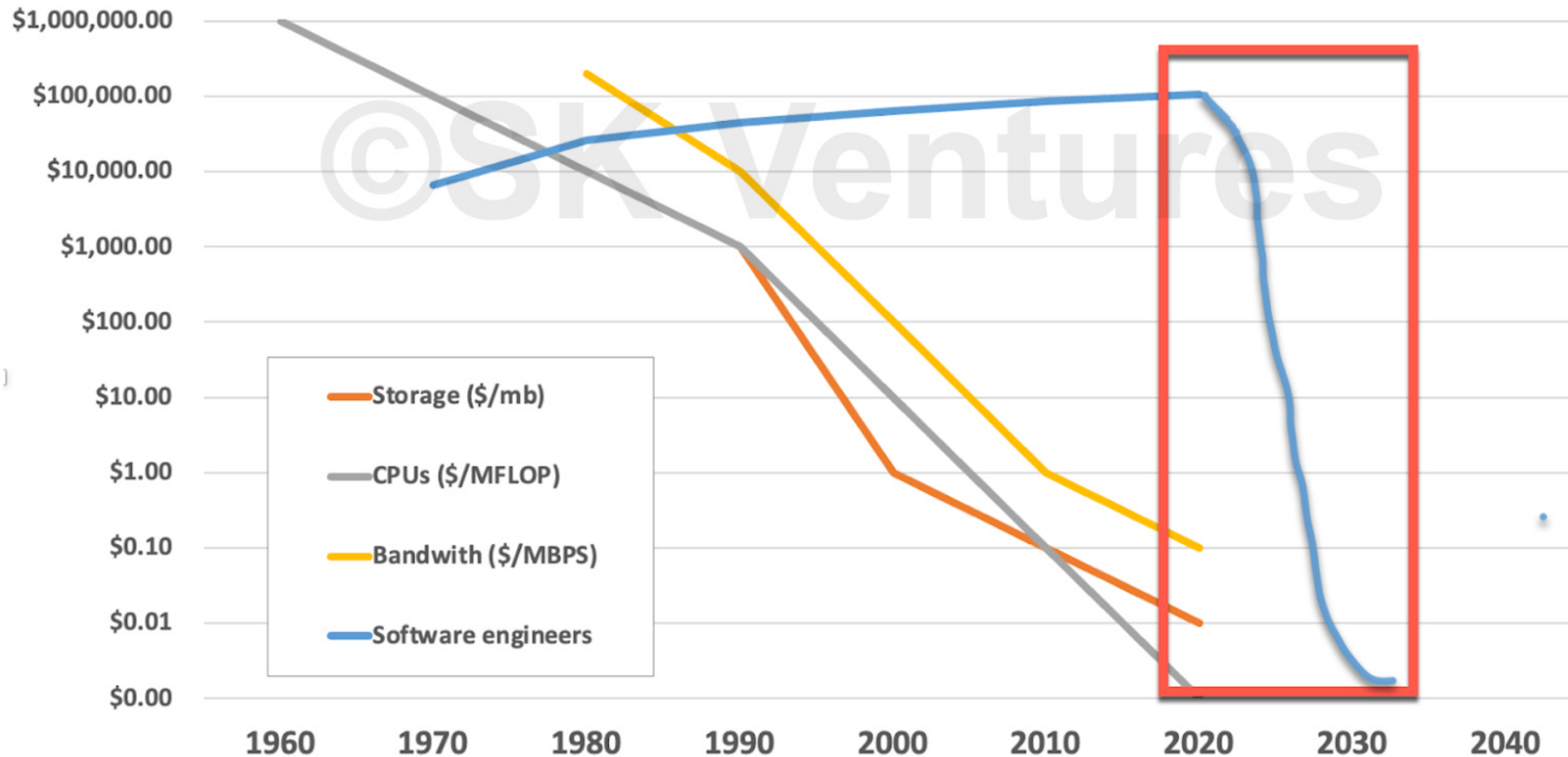
Today I asked ChatGPT to solve almost every exercise in Think Python and DSIRP.

It did.

My conclusion: everyone who writes code should spend the next month doing professional development on writing code with LLM-assist.

This is how code will be written from now on.

## The Next Collapsing Tech Cost: Software Itself



However...



**Logan Bartlett**  
investor  
**Redpoint Ventures**

We have been losing  
jobs to technology  
since the 1700s and yet  
we are at 93%  
worldwide employment!



Jacquard Loom, patented in 1804

# Luddites

- Luddites were lower class English textile workers in squalid living conditions who turned against the technology they thought was causing their misery
- Britain dispatched 12,000 troops to suppress Luddite activity (more than were fighting Napoleon around 1807-1814)
- Neo-Luddites is now a thing



**Emad Mostaque**  
Founder and CEO of Stability AI

**“AI will not replace you. A person using AI will.”**

**—A lot of people**



# Computers of the Future

- Generative AI is just text/image prediction, it generates text

# What Can Generated Text Do?



# What Can Generated Text Do?

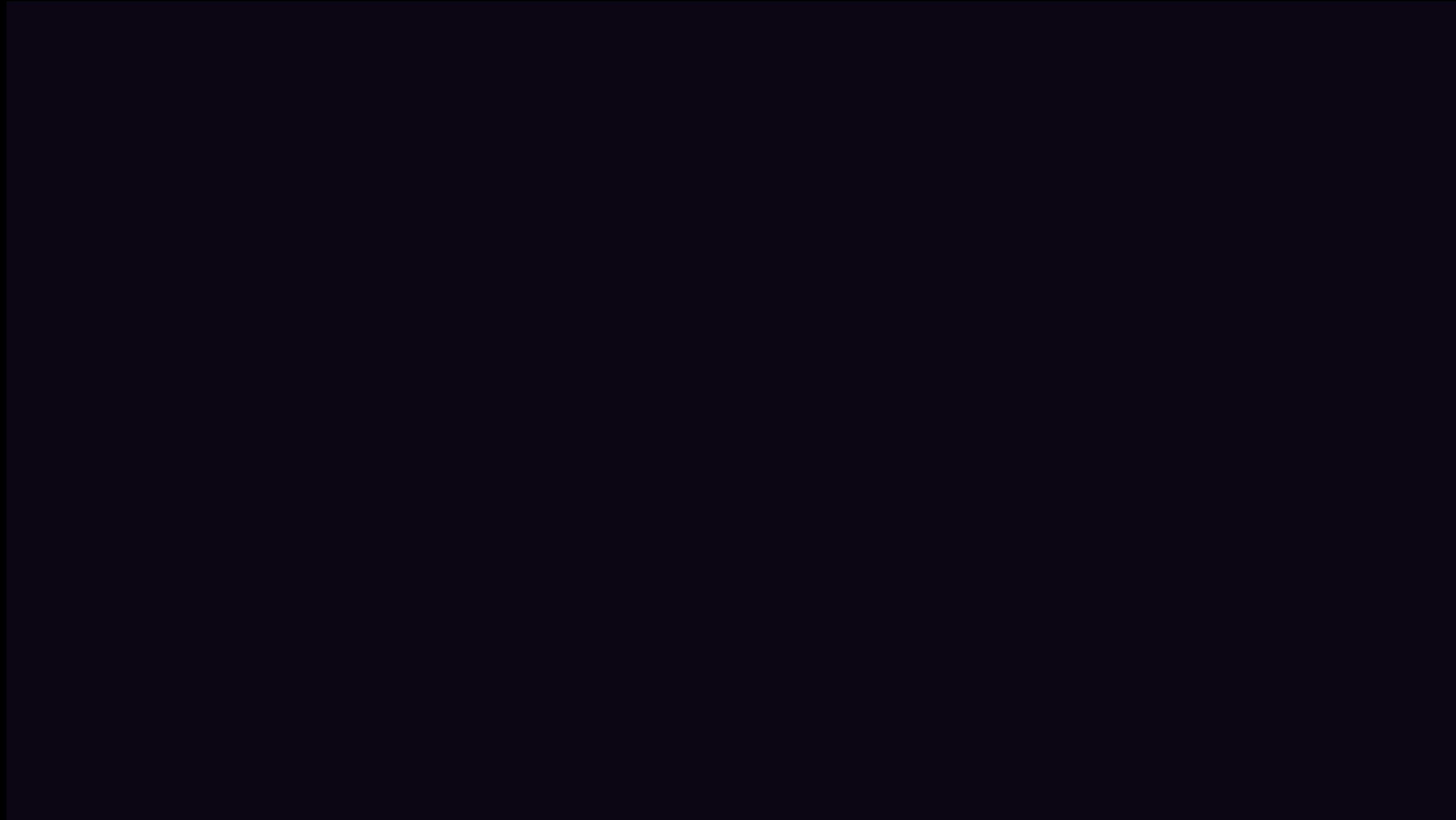
> Do you need a hint?

# What Can Generated Text Do?

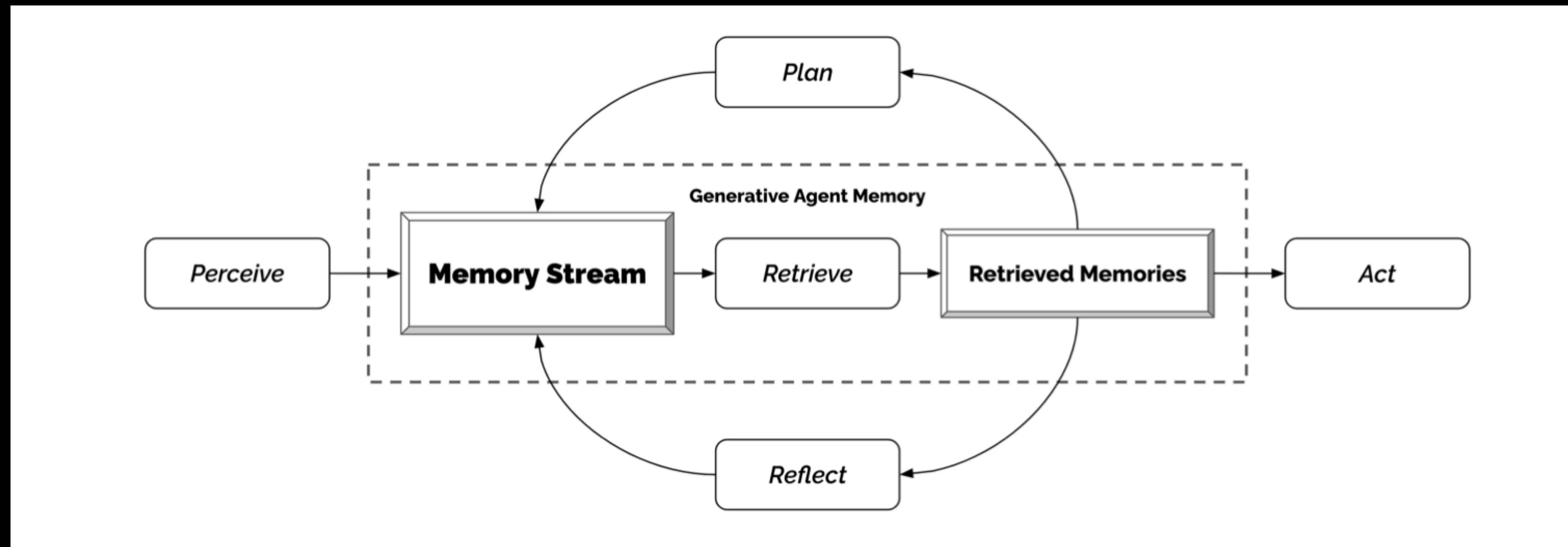
> Do you need a hint?

zsh: no matches found: hint?

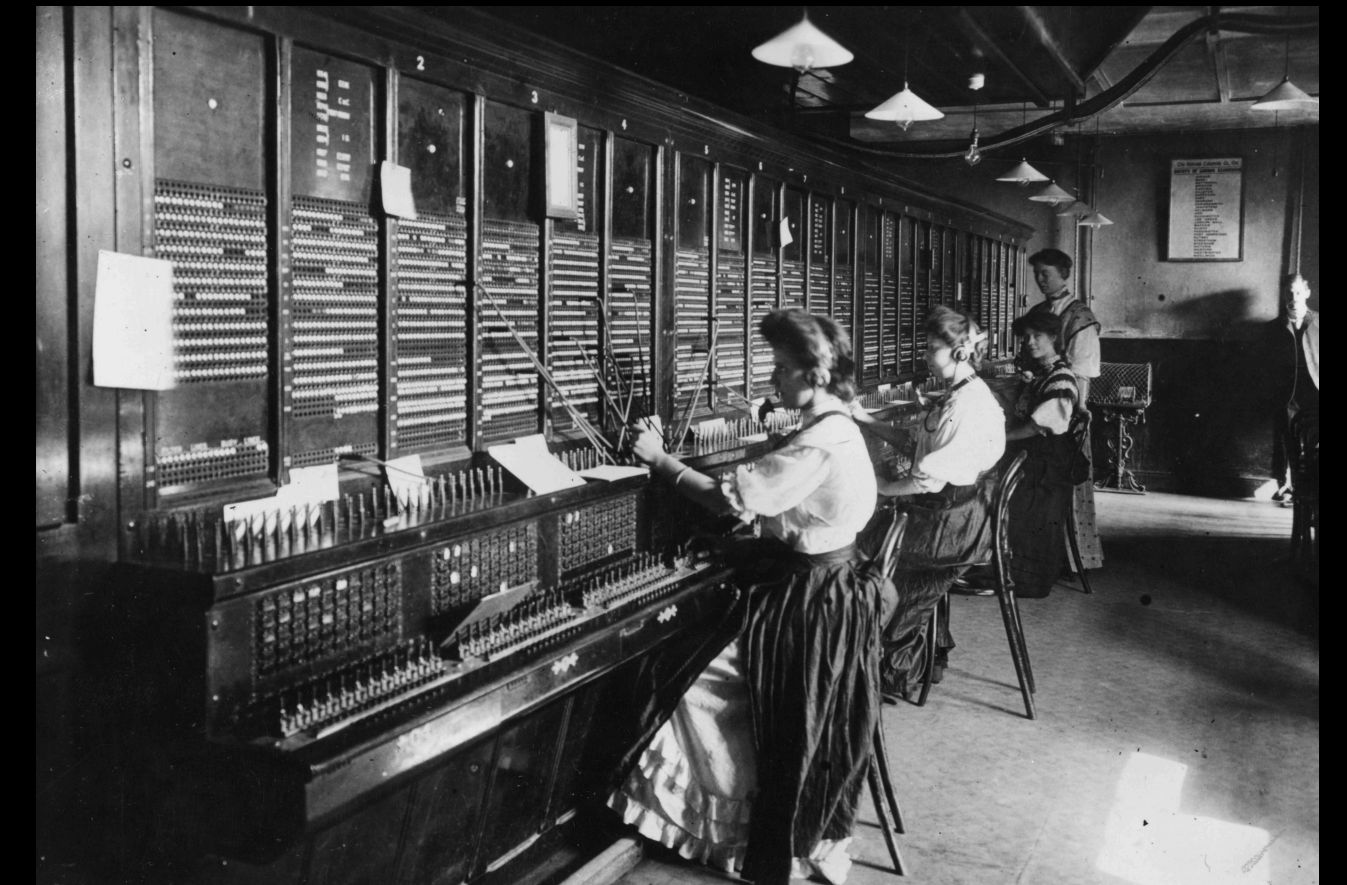
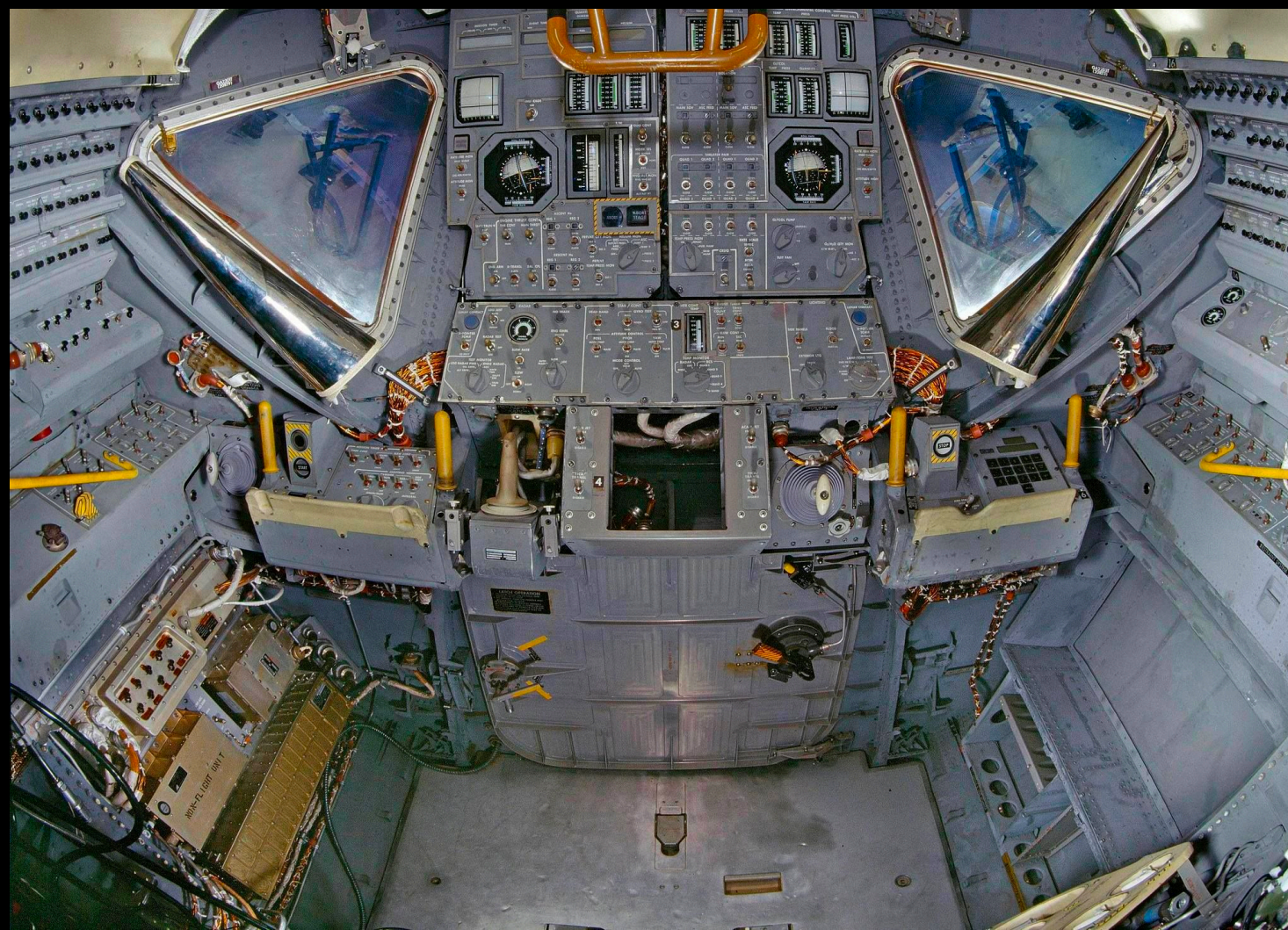
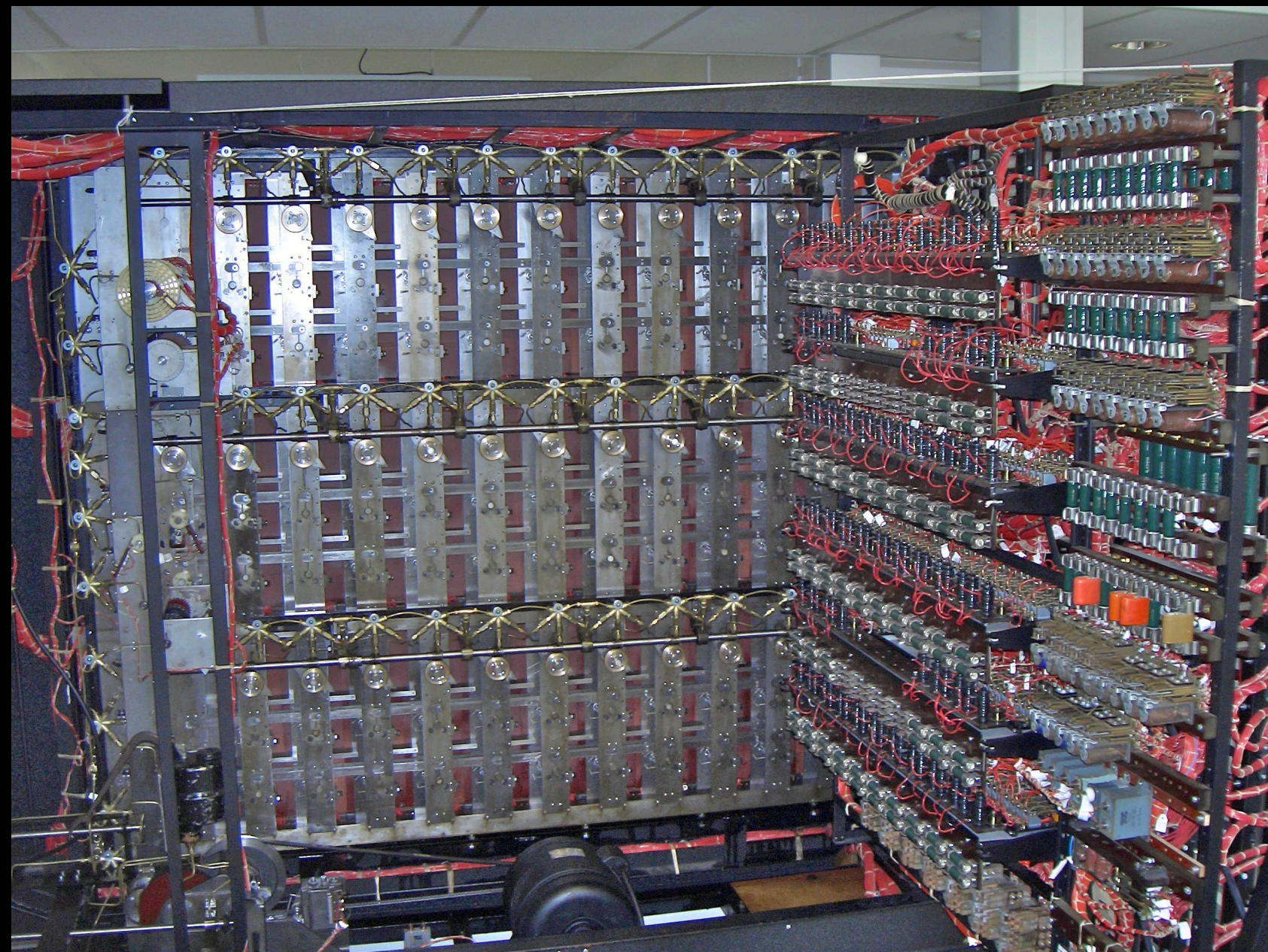
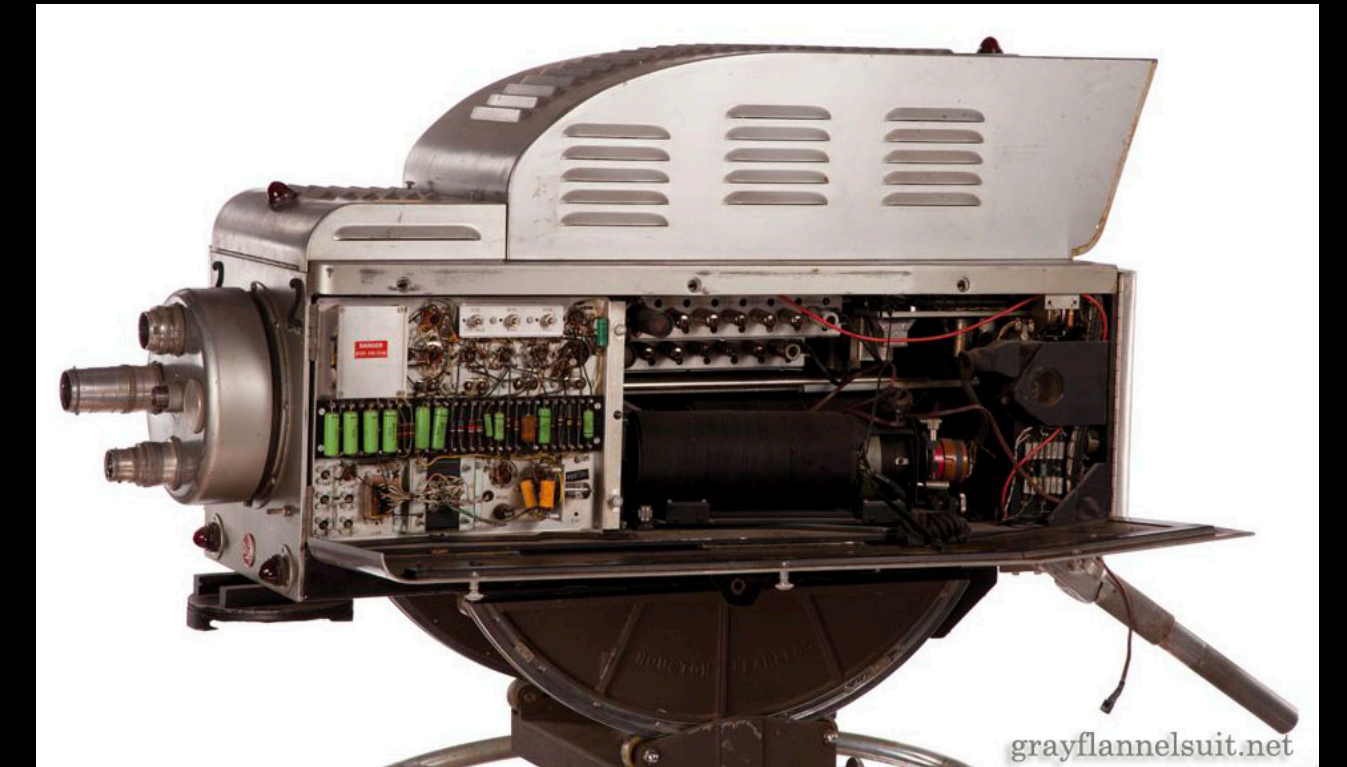
# Agent Mode (Fireship video)



# From Text Generation to Scaffolded LLM Computers of the Future?



# Specific Purpose Hardware

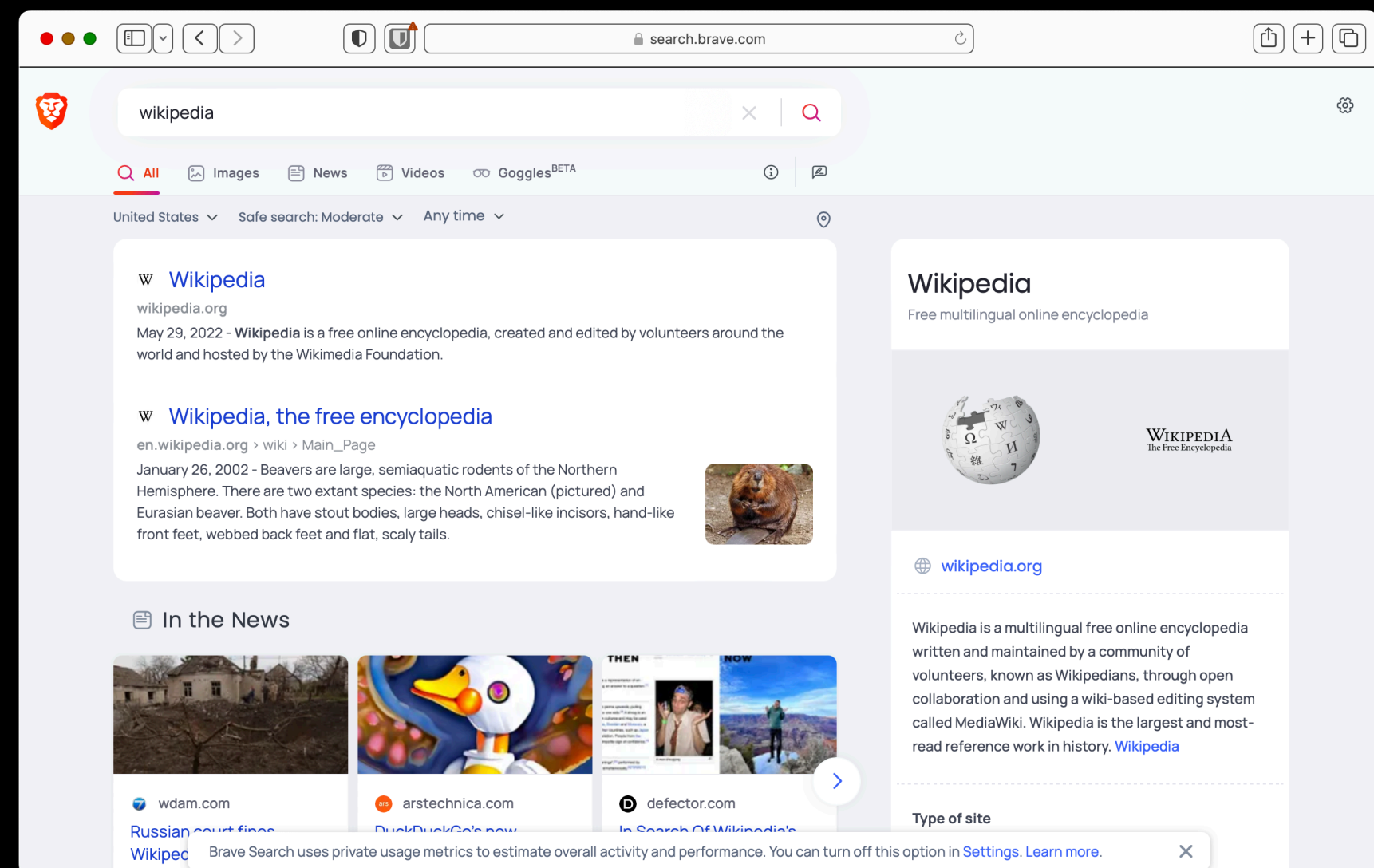
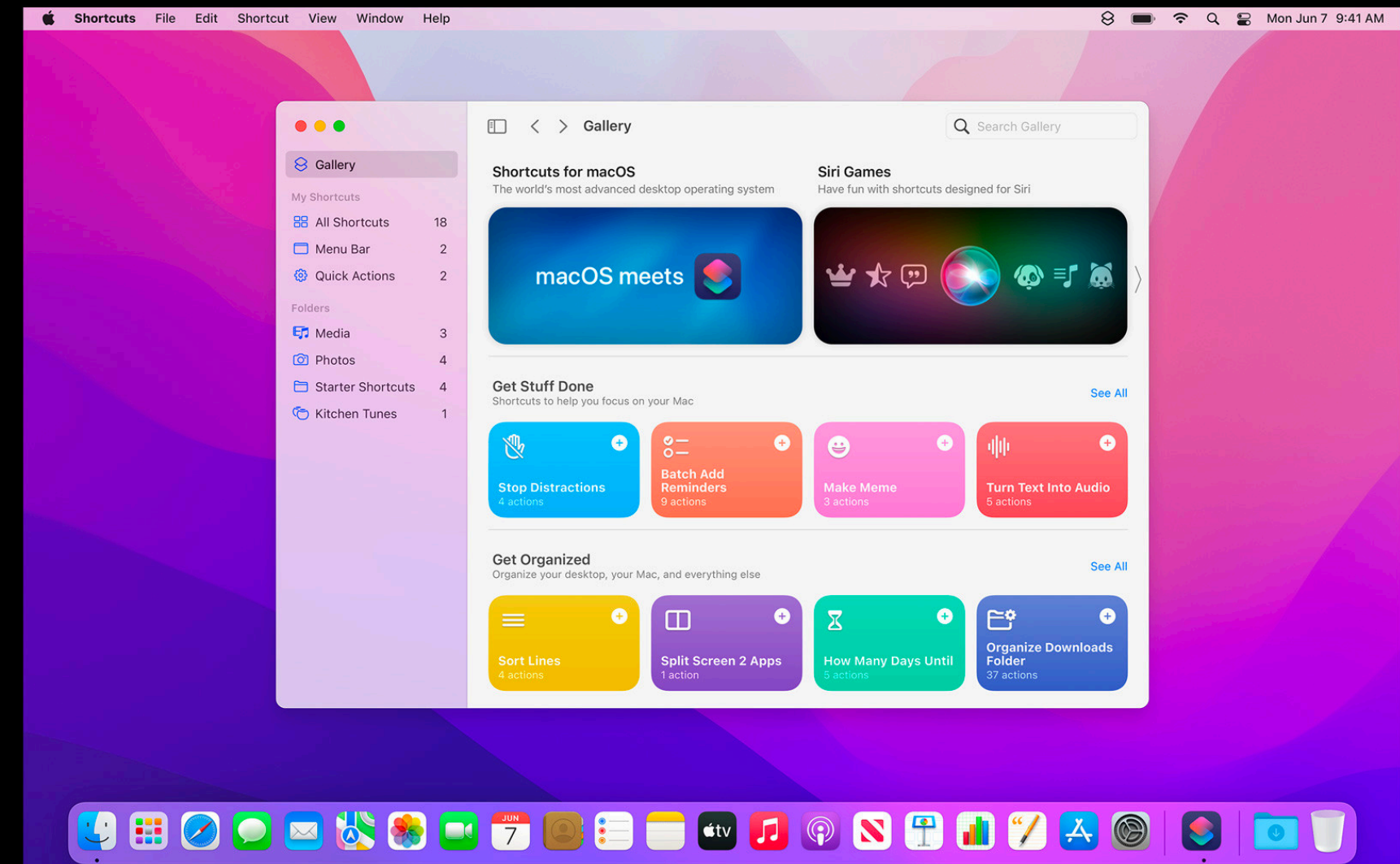


# General Purpose Hardware



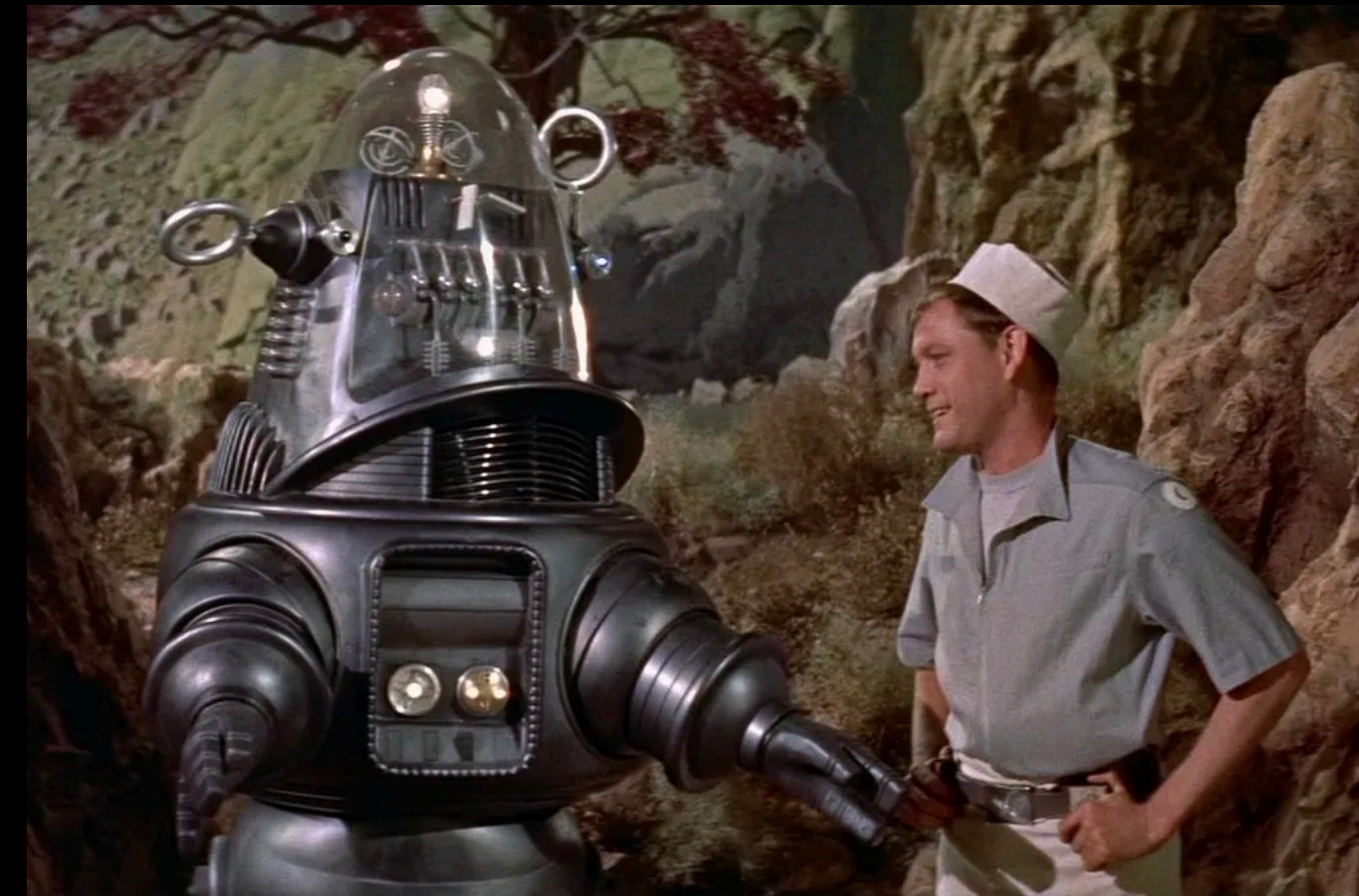


# Specific Purpose Software



# General Purpose Software

Natural language software limited by agents



# AI Can Write Assembly...



Write me x86 assembly that will print hello world to stdout.



Here's an example x86 assembly program that will print "Hello, World!" to stdout:

assembly

Copy code

```
section .data
    message db "Hello, World!", 10 ; message to print, with newline character
    len equ $-message ; length of the message

section .text
    global _start

_start:
    ; write message to stdout
    mov eax, 4 ; system call for write()
    mov ebx, 1 ; file descriptor for stdout
    mov ecx, message ; address of message
    mov edx, len ; length of message
    int 0x80 ; make the system call

    ; exit the program
    mov eax, 1 ; system call for exit()
    xor ebx, ebx ; exit status code (0)
    int 0x80 ; make the system call
```

# From This



Tesla H100 80GB NVIDIA Deep Learning GPU Compute Graphics Card 900-21010-000-000

Brand New · NVIDIA

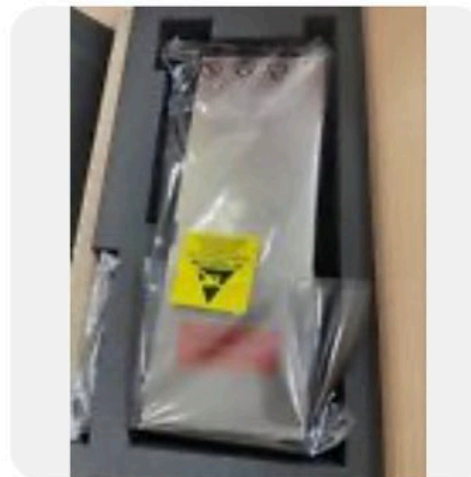
**\$42,750.00**

Was: ~~\$45,000.00~~ 5% off

or Best Offer

Free shipping

from China



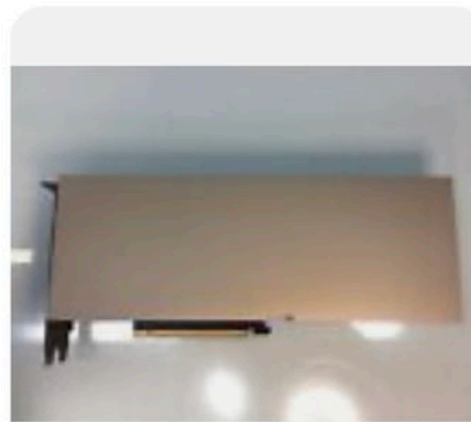
Tesla H100 80GB NVIDIA Deep Learning GPU Compute Graphics Card

Brand New · NVIDIA

**\$42,672.00**

or Best Offer

Free shipping



Nvidia H100-PCIe-80GB Hopper H100 80GB PCIe Tensor Core GPU Accelerator

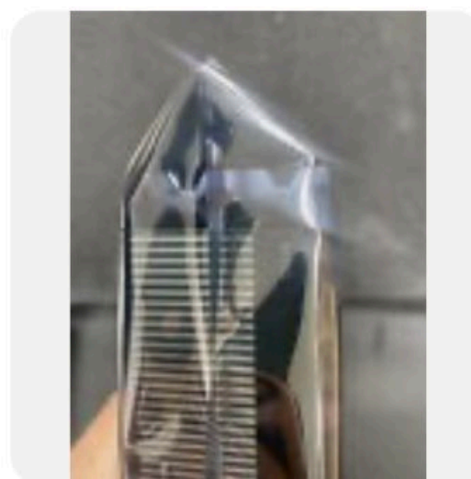
Pre-Owned · NVIDIA · 80 GB

**\$39,995.00**

or Best Offer

Free shipping

2 watchers



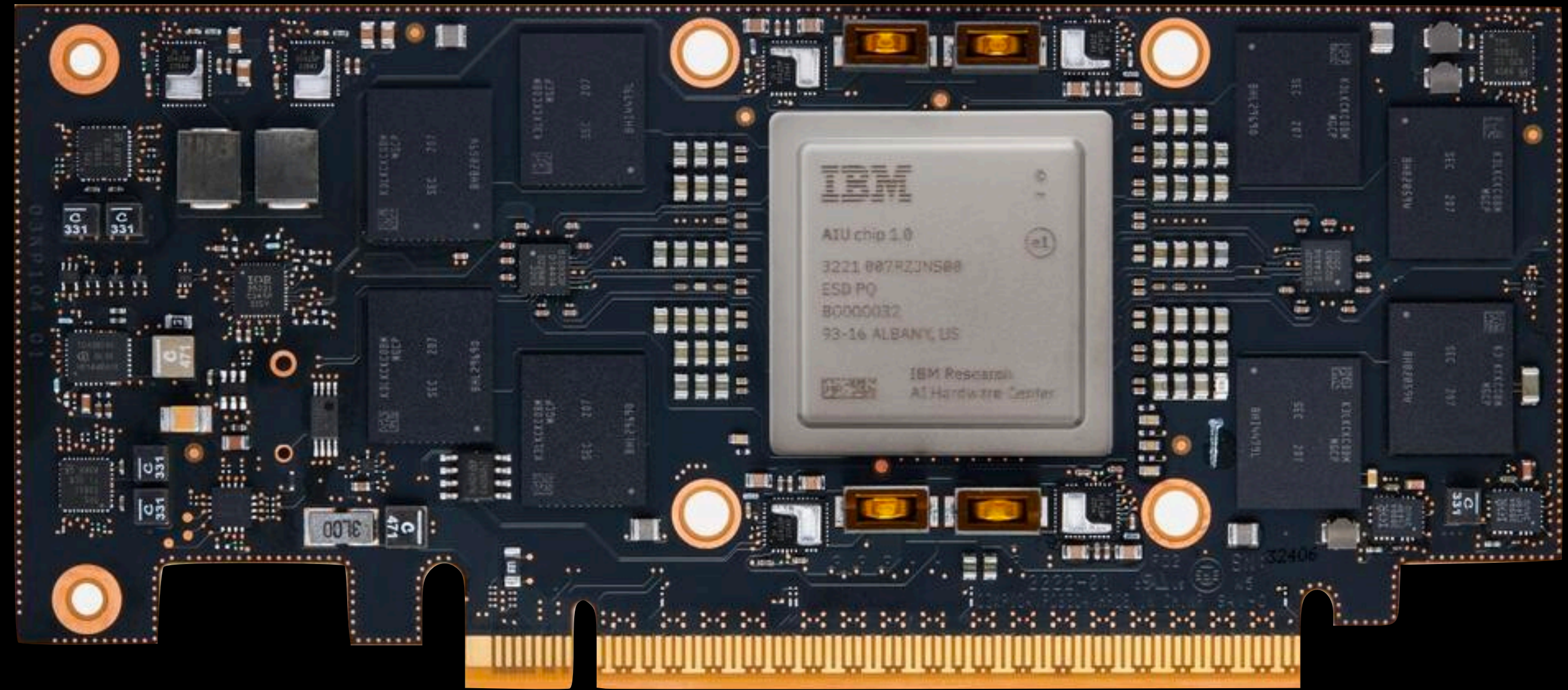
NVIDIA H100 80GB Tesla Deep Learning GPU Compute Graphics Card 900-21010-000-000

Brand New · NVIDIA

**\$45,000.00**

# To This

- AIU, NPU, TPU, Athena

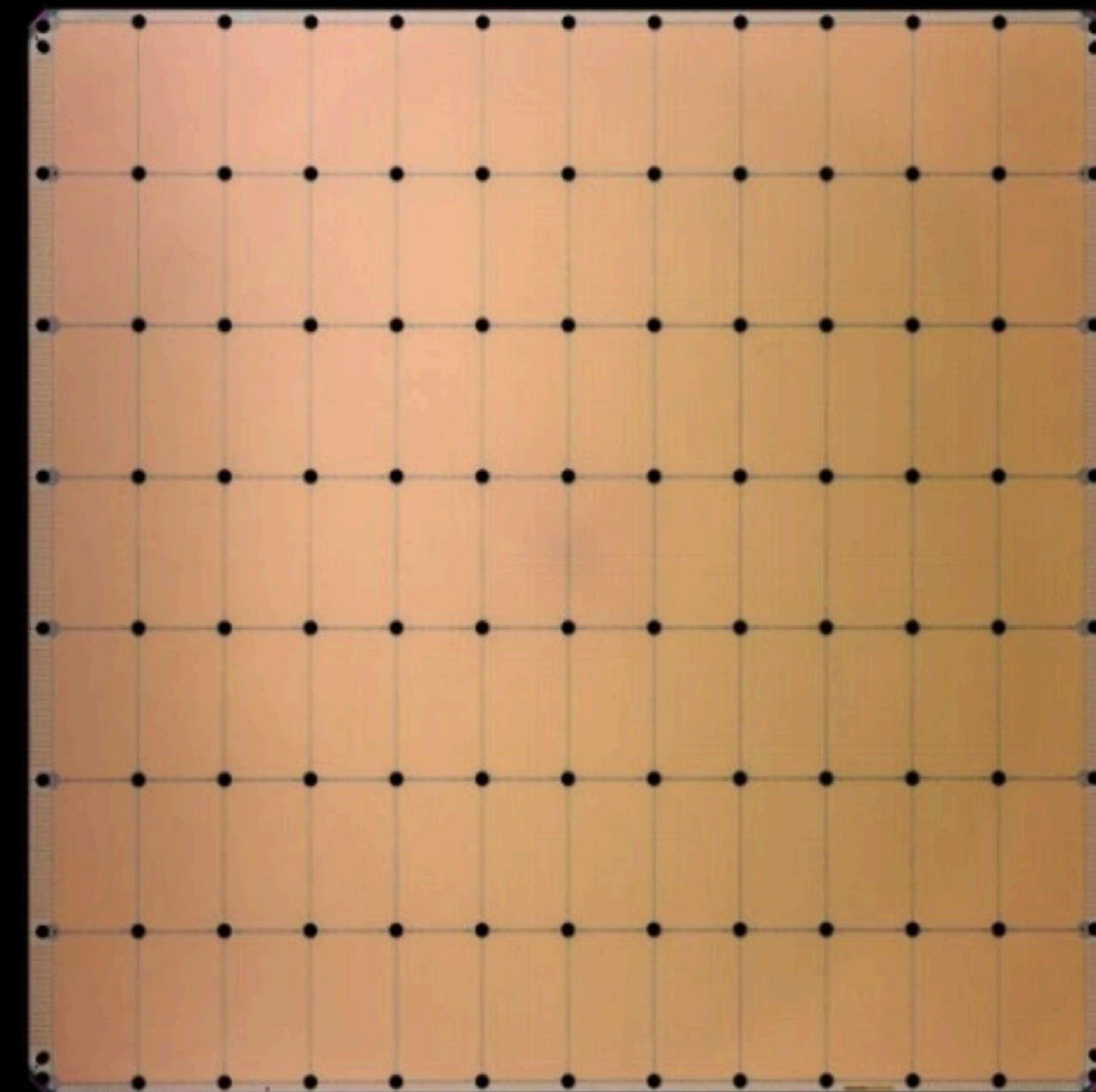


- Cerebras Wafer Scale Engine



**Largest GPU**  
54.2 Billion Transistors  
826 mm<sup>2</sup> Silicon

- Analog transistors



**Cerebras WSE-7nm**  
2.6 Trillion Transistors  
46,225 mm<sup>2</sup> Silicon

# ObjectDetection

mos/models/segnet

```
l.py:718: UserWarning: Named  
Please do not use them for an  
@/core/TensorImpl.h:1156.)  
(h, ceil_mode)  
gnet/
```

```
s-0.01.tar segnet_weights.p  
ep deeplab
```

```
u mean "!="?
```

```
u mean "!="?
```

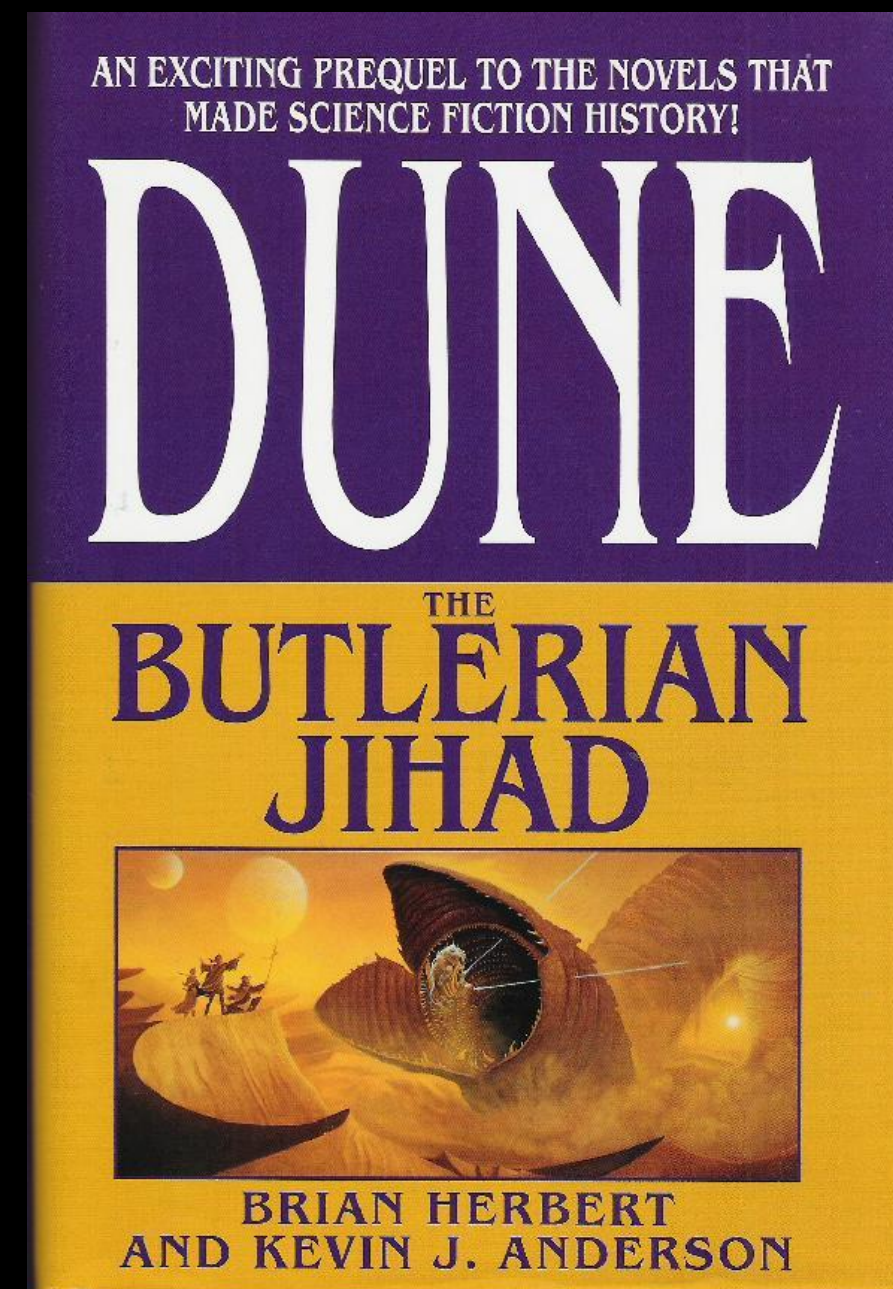
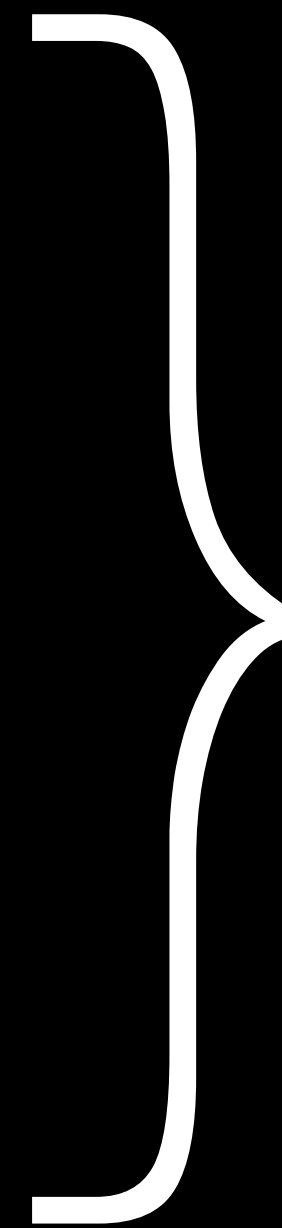
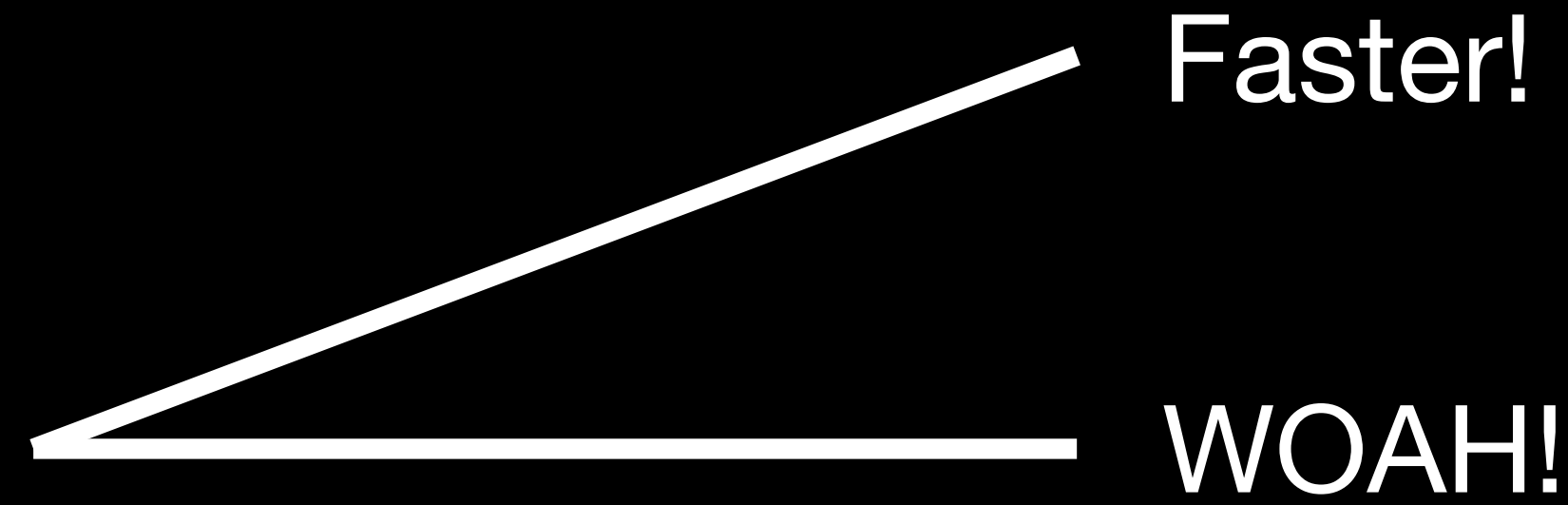
```
py:718: UserWarning: Named  
Please do not use them for an  
core/TensorImpl.h:1156.)  
(ceil_mode)
```



(x=306, y=27) ~ R:177 G:201 B:234

# The AI Tribes

- Chernobylists  
“AI is dangerous”
- Language Police  
“AI is toxic”
- X-riskers/Doomers  
“We’re all gonna die!!!!”
- Boomers  
“AI wut?”
- Deniers  
“AI is a fake conspirazy!”



"Butlerian Jihad" is a science fiction novel that tells the story of a rebellion against artificial intelligence and the rise of human supremacy in a future interstellar civilization.

–ChatGPT

**WE'RE  
ALL GONNA  
DIE**

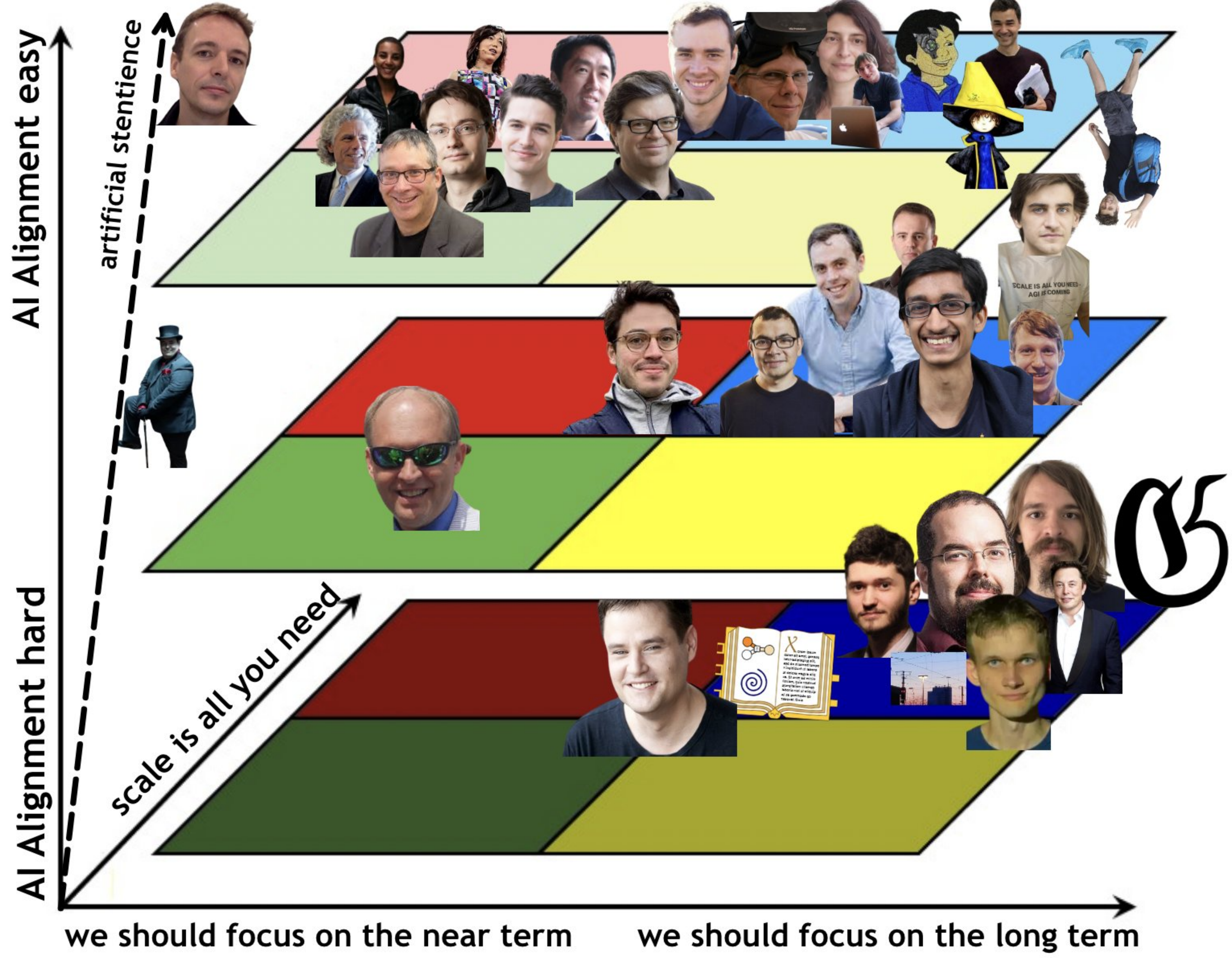
EP: 159

Featuring  
**Eliezer  
Yudkowsky**

The image is a podcast cover with a red and black background. On the right is a large, high-contrast black and white portrait of a man with a beard and glasses. On the left, there are two smaller black and white portraits of men wearing headphones. The title 'WE'RE ALL GONNA DIE' is written in large, bold, white letters with a black outline. The episode number 'EP: 159' is in the top right. The text 'Featuring Eliezer Yudkowsky' is in the bottom right.

<https://www.youtube.com/watch?v=gA1sNLL6yg4>





we should focus on the near term

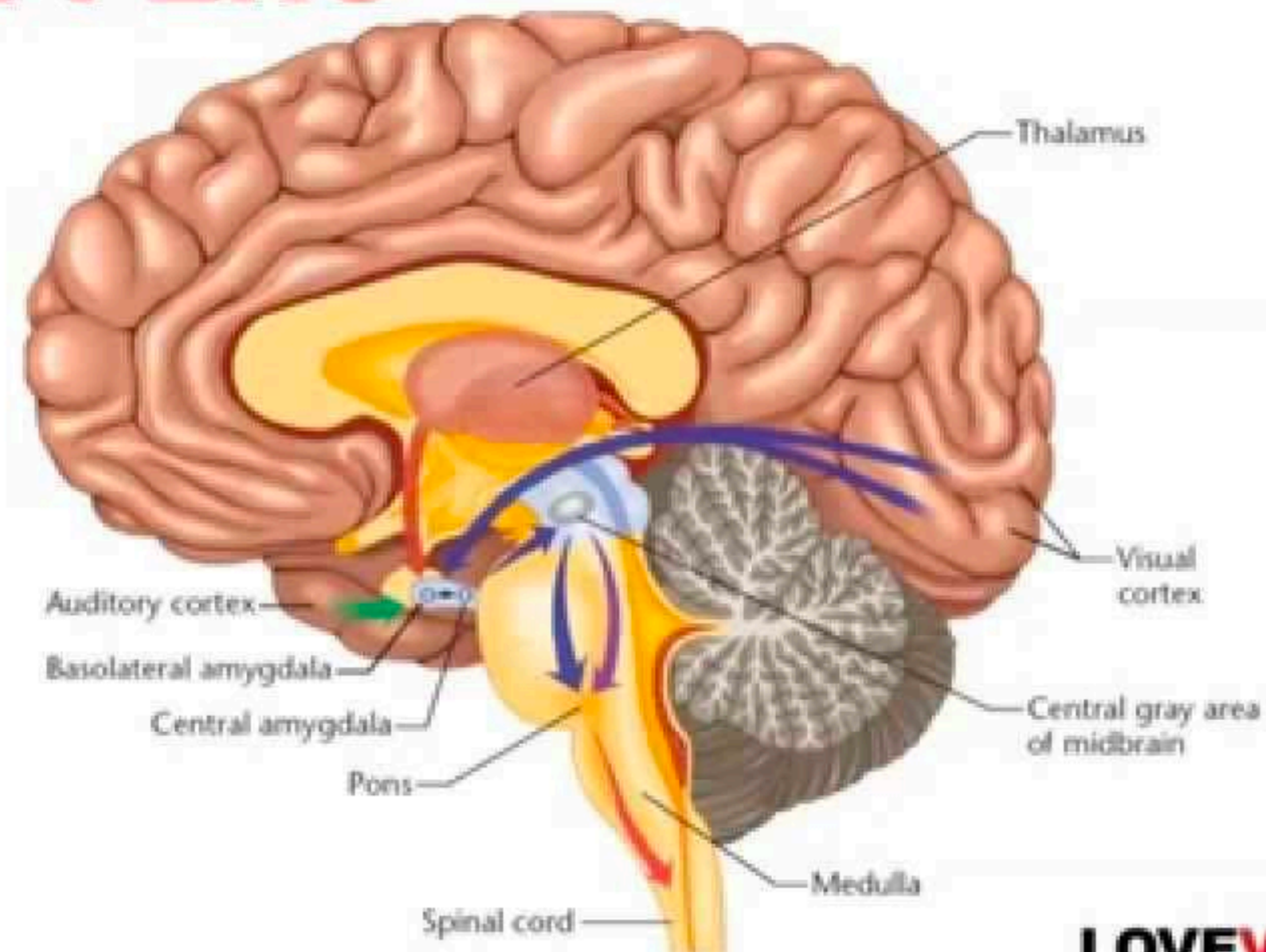
we should focus on the long term

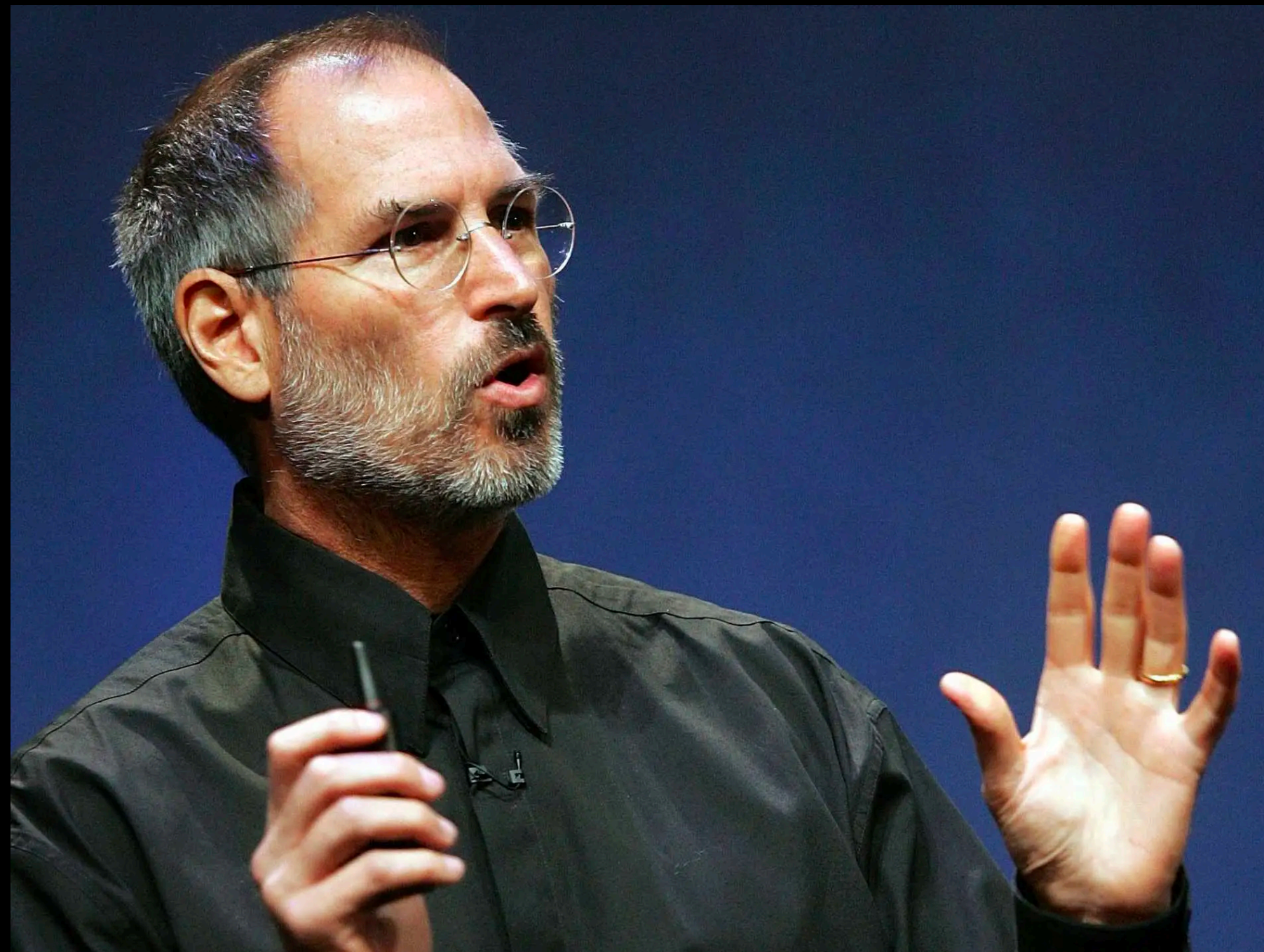
# What has to happen to slow or stop AI?

- Control GPUs, cloud computing, and LLM files BECOME guns, digital contraband, childporn, firearm stl files, etc.
- “Imagine that with a single software download, any laptop or smartphone on the planet could be turned into a loaded 9mm pistol” —Jon Stokes (<https://www.jonstokes.com/p/heres-what-it-would-take-to-slow>)
- “If we’re going to stop... you have to stop it globally... we need to have surveillance technology installed on every single laptop to make sure that people aren't typing in keystrokes to code up the AGI that's going to destroy the world” —Peter Thiel (<https://youtu.be/OWXFdEyOKc4?t=2402>)
- “Preventing AI extinction scenarios is considered a priority above preventing a full nuclear exchange” —Eliezer Yudkowsky (<https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/>)

# HOW AMYGDALA HIJACKING HAPPENS

- The stimuli goes directly to thalamus and it then goes right to amygdala before a signal reaches the neocortex to process.
- This survival mechanism lets us react to things before the rational brain has time to mull things over.





**“Innovation is the ability  
to see change as an  
opportunity, not a threat”**

**—Steve Jobs**

**AI Apps**

# Big Names

- ChatGPT - <https://chat.openai.com> (available now)
- Bing Chat (Windows 11, Edge, or waitlist & ads)
- Brave Search Summerizer (available now and it gives sources)
- Coming
  - Microsoft Office 365 Copilot (no release date yet)
  - Google Bard? - <https://bard.google.com/> (waitlist)

# Lots of AI Apps

1000+ new AI apps released in March 2023

- [grammarly.com](https://grammarly.com)
- [midjourney.com](https://midjourney.com) (art)
- Stable Diffusion (art)
- [neural.love](https://neural.love) (image editing)
- D-ID (animated faces from photos)
- ElevenLabs (voice cloning)
- [www.synthesia.io](https://www.synthesia.io) (talking head)
- [Numerous.ai](https://Numerous.ai) (spreadsheets)
- Gamma AI & Slides AI (slides)
- [lumen5.com](https://lumen5.com) (video creation)
- [Cleanvoice.ai](https://Cleanvoice.ai) (removes "um")
- Nova AI (caption videos)
- Wix ADI (create webpages)
- [koe.ai](https://koe.ai) (voice changer)
- Krisp AI (remove bg noise)
- [www.lalal.ai](https://www.lalal.ai) (audio separation)

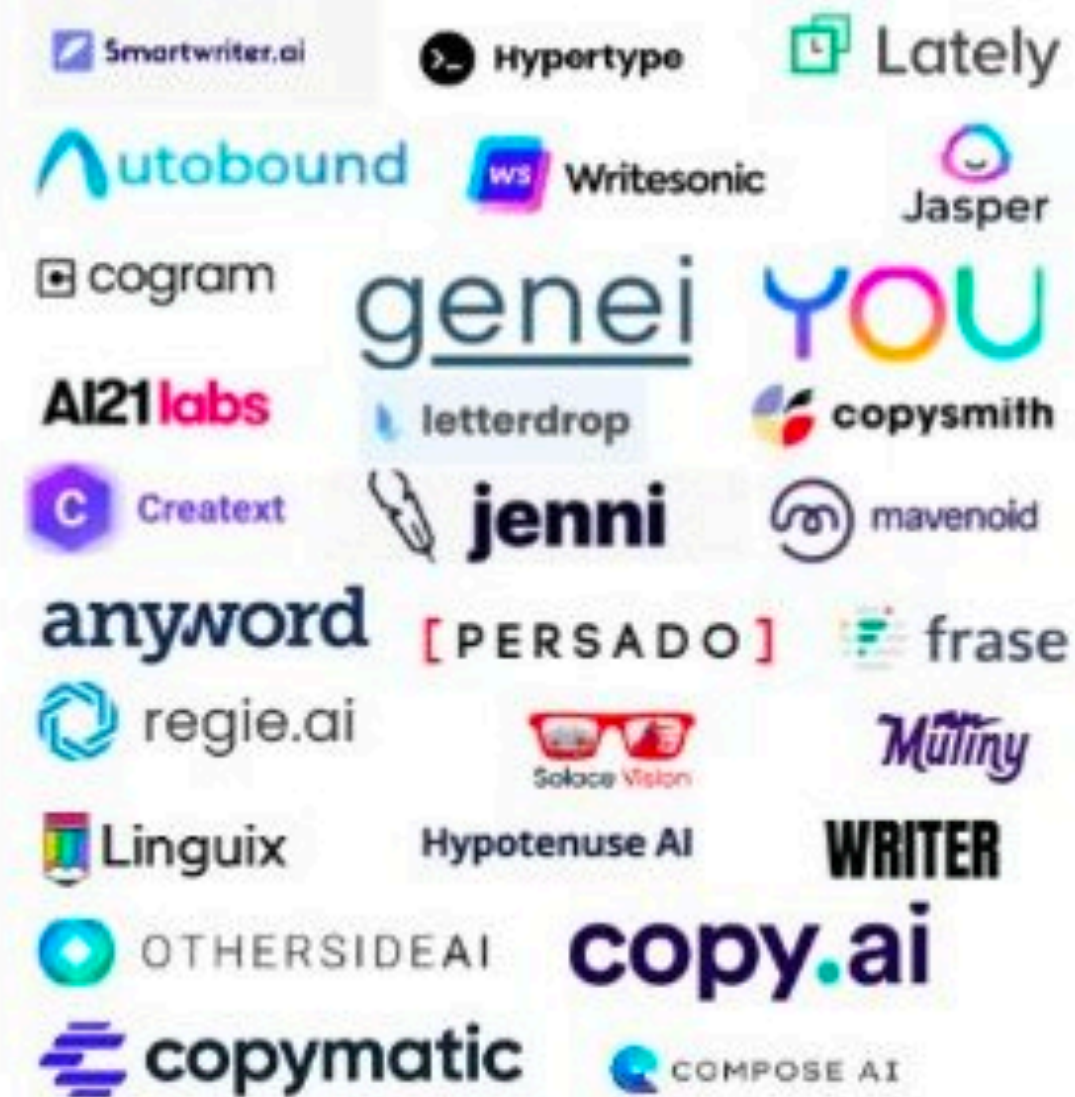
# Lots of AI Apps

1000+ new AI apps released in March 2023

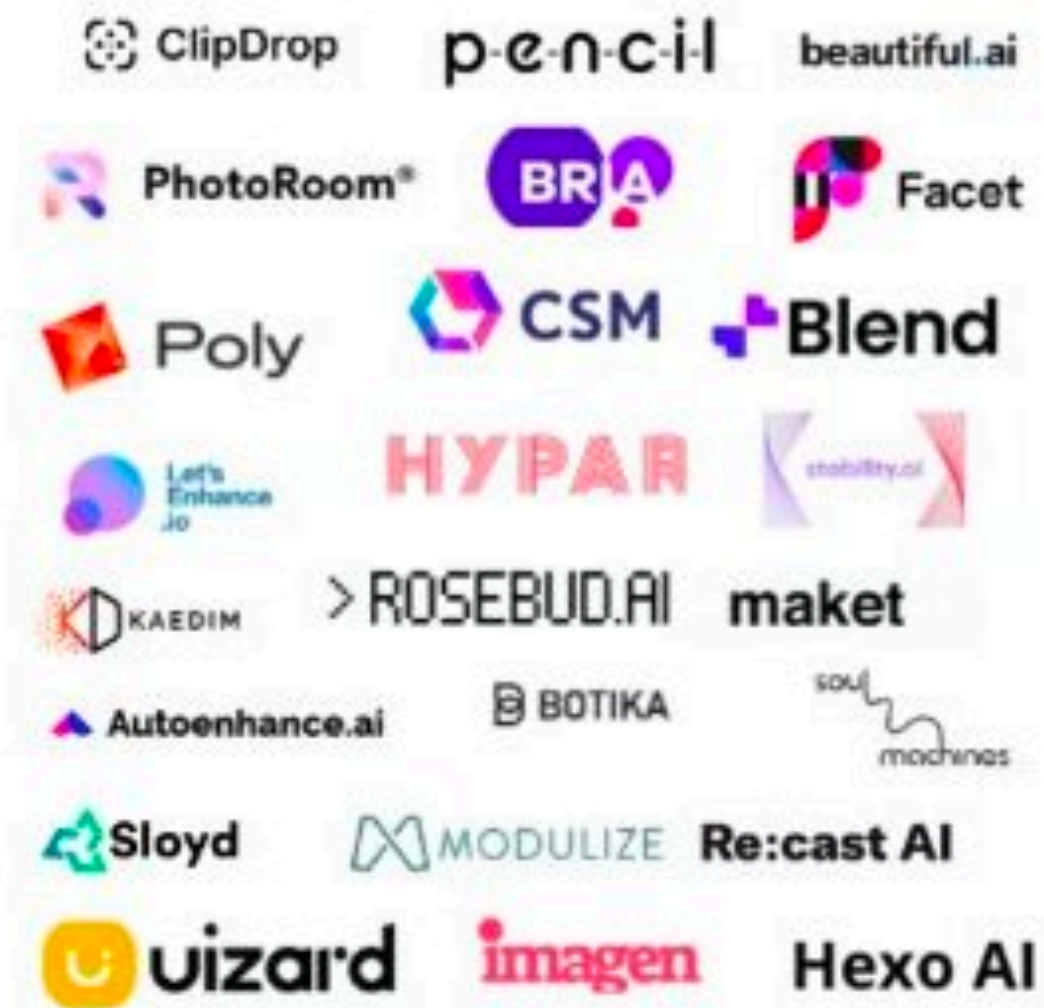
- rask.ai (translate videos)
- podcastle.ai (podcasts)
- Deep Nostalgia (animate photos)
- Browse.ai (surf the web)
- soundraw.io (make songs)
- Looka (brand generator)
- ChoppedAI.com (recipes)
- videohighlight.com
- Warp AI
- Legal Robot
- Perplexity for iPhone
- Kickresume AI
- Uizard
- Natural Reader
- AI Text Classifier
- PrankGPT



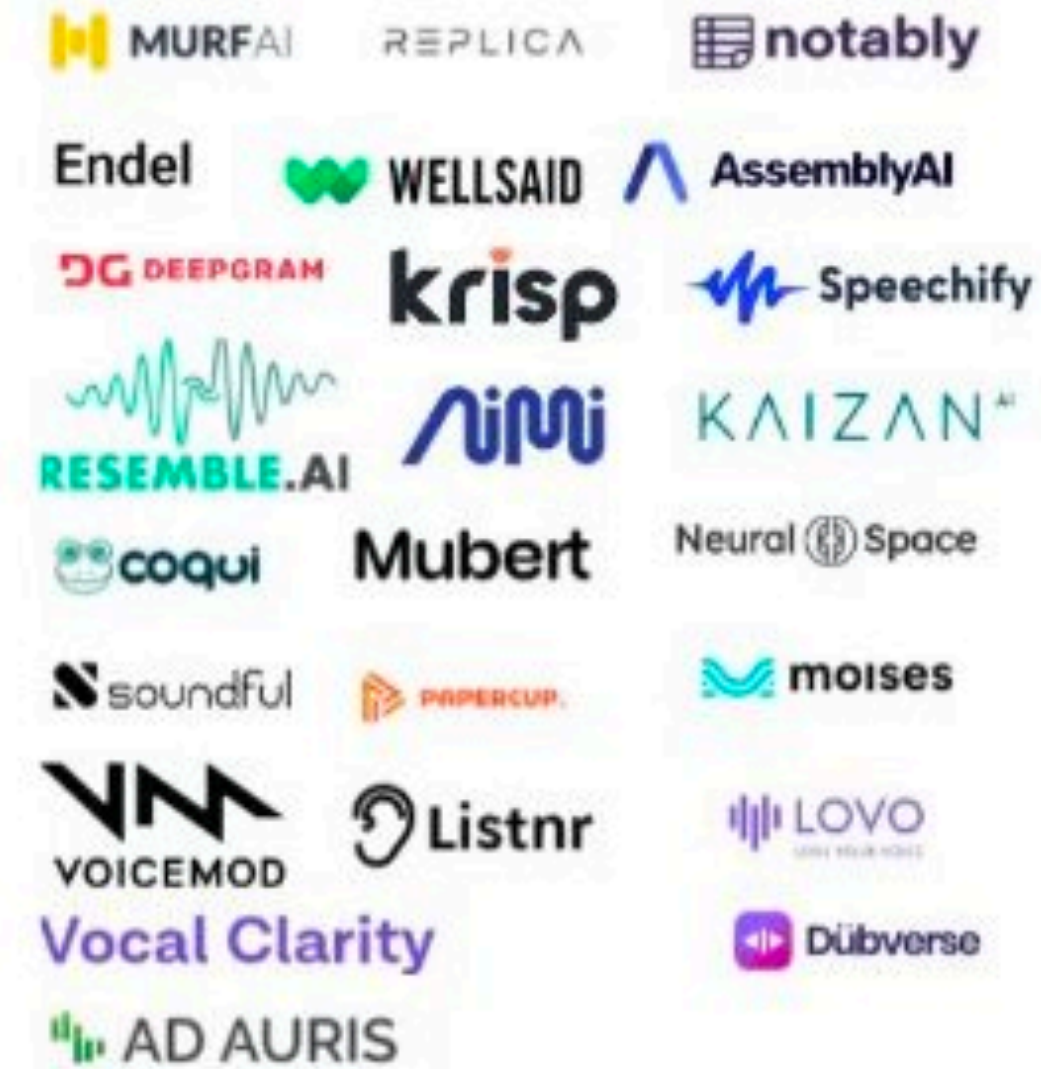
## TEXT



## IMAGE



## AUDIO



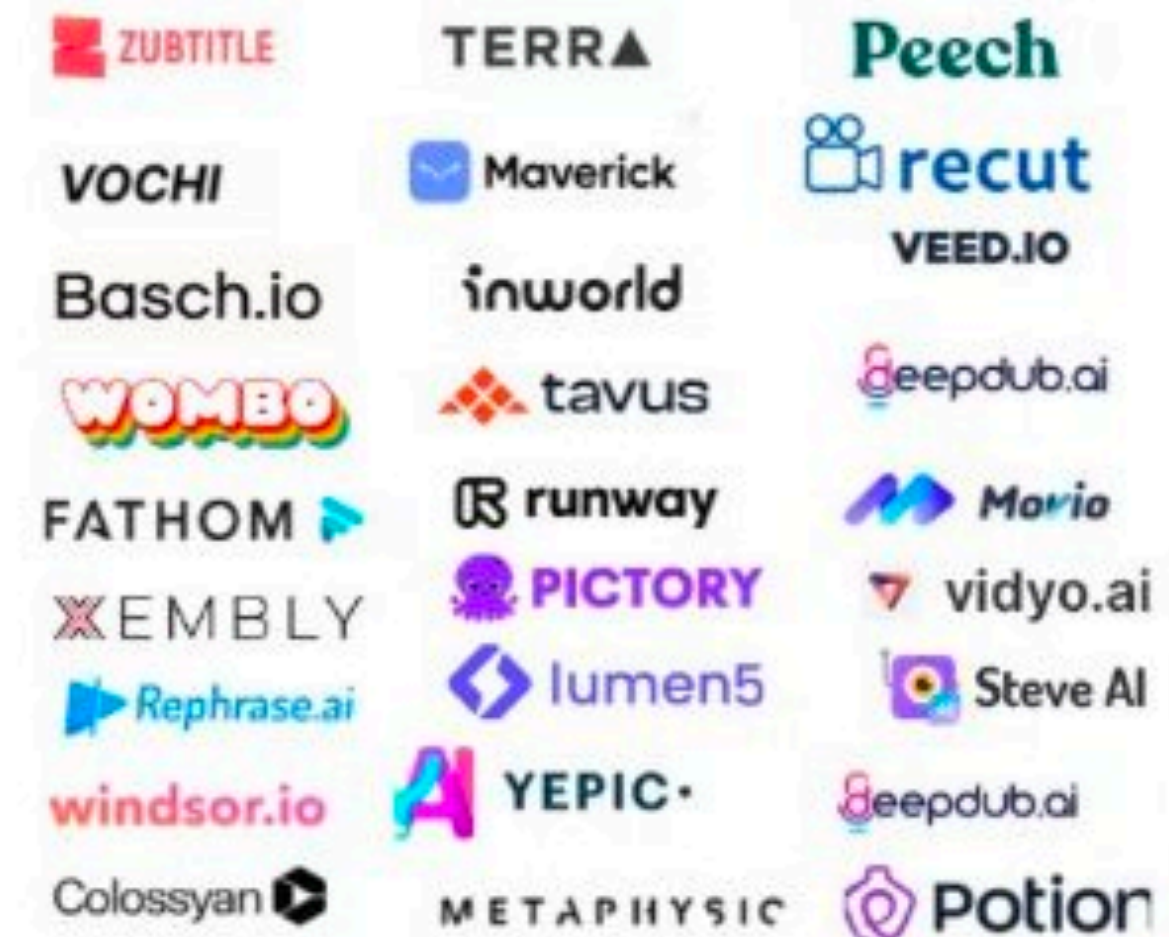
## CODE



## CHATBOTS



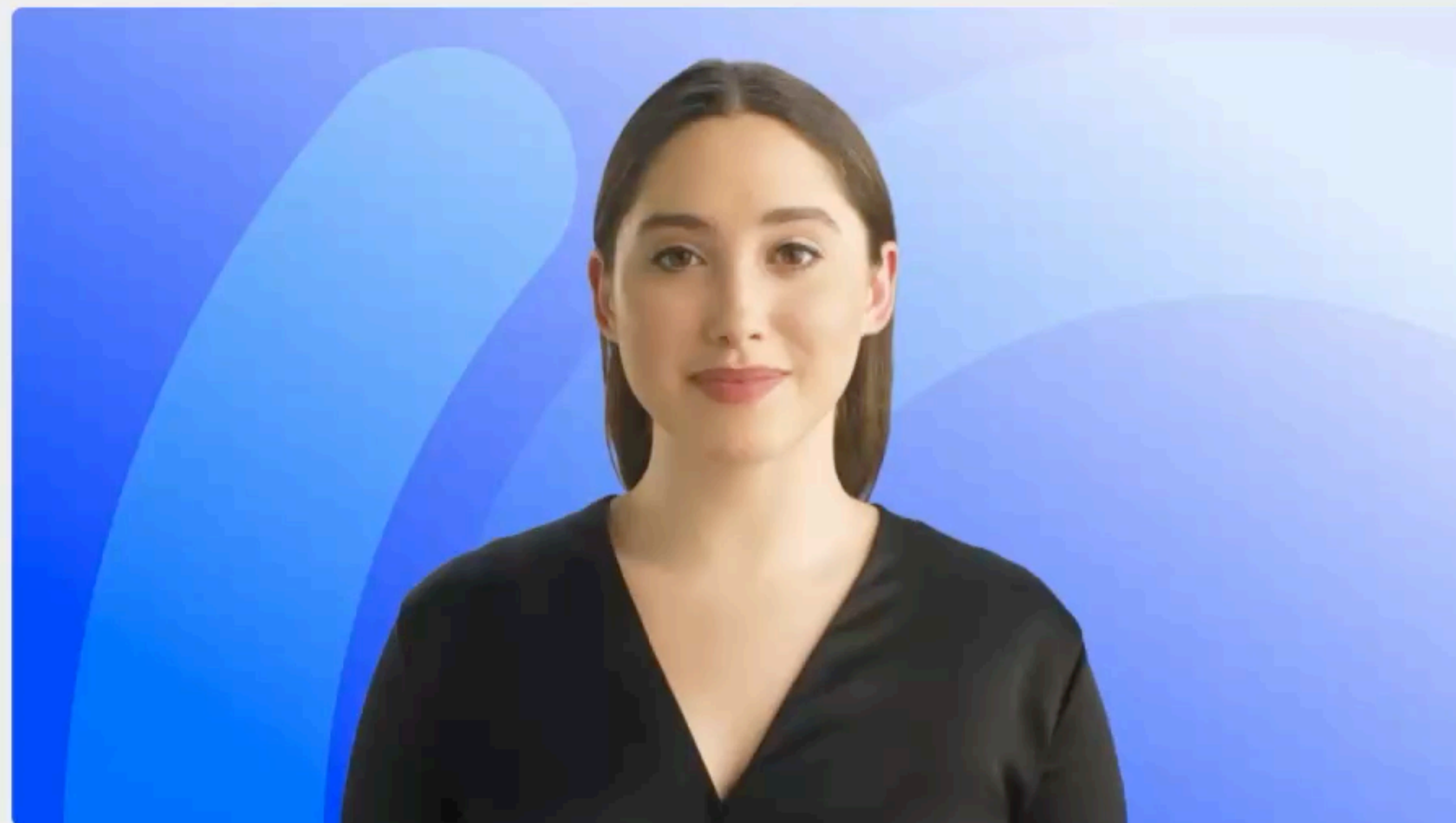
## VIDEO



Scenes



SCENE 1



Scene 1

Replace layout

Background color

#FFFFFF

Background media

Replace

Crop

Opacity 100 %

Blur 0 %

Music

English (US) - Engaging

Gesture Marker Pause Diction

00:00 / 01:22

Hey! I'm Alex, and here's a quick Synthesia product demo.

To begin, choose from a wide variety of fully customizable video templates and select your preferred AI avatar.

# AI Apps for IT

# Code Assistants

- GitHub Copilot (VSCode only)
- AWS CodeWhisperer
- bito.ai
- ChatGPT (web only)
- CodeGen (open source)
- K.Explorer
- Dodiga
- Replit
- SinCode AI
- Kite
- Tabnine
- SudoLang (pseudo code)
- Build AI (English = code)
- More...

# ChatGPT Demo

# GitHub Copilot Demo

- Write a comment

```
# q:
```

```
# Write a function that
```

```
def squareRoo
```

- Tab
- Arrow keys
- Cntl-enter

# AWS Code Whisperer Demo

- Keyboard shortcuts
  - Option-c
  - Tab
  - Arrow Keys
- Click on AWS in the sidebar

# Public Service Announcement

- <https://platform.openai.com/account/usage>



# iTerm



- Get the Beta “Test Release” <https://iterm2.com/downloads.html>
- Preferences -> General -> Magic
  - “OpenAI API Key” & “AI Prompt”
- Preferences -> Advanced -> Experimental Features
  - “Maximum tokens for OpenAI” & “OpenAI Model name”

# Installing cli tools

- Each one has it's own instructions or you can use tea
  - Install tea: ``sh <(curl https://tea.xyz)`` and enable “magic”
  - ``export OPENAI_API_KEY=sk-secret123ABCsecret123ABC``
  - Then just run the command

# OpenAI API on the CLI

```
aiac get "Ansible playbook that downloads and installs homebrew"
```

```
chatblade "sed command to print the output of the first column of ps"
```

```
openai api chat_completions.create -m gpt-3.5-turbo -g user "Hello world"
```

- The following isn't in tea

```
python engshell.py
```

# Python OpenAI Library

```
import openai  
  
openai.api_key = "sk-..."  
  
completion = openai.Completion.create(model="ada", prompt="Hello world")  
print(completion.choices[0].text)
```

# LangChain

- Generic LLM interface
- Prompts
- Chains
- Memory
- Indexes
- Agents & Tools

<https://www.youtube.com/watch?v=RoR4XJw8wIc>

# LangChain

## Generic Model Interface

```
from langchain.llms import OpenAI
from langchain import HuggingFaceHub
from langchain.llms import Cohere

llm = OpenAI(model_name="text-ada-001")

llm = HuggingFaceHub(repo_id="google/flan-t5-xl")

llm = Cohere()

llm("Tell me a joke")
```

# LangChain

## Prompt Templates

```
from langchain import PromptTemplate

template = """Question: {question}
Let's think step by step.
Answer: """

prompt = PromptTemplate(template=template, input_variables=["question"])

user_input = input("What's your question? ")
prompt.format(question=user_input)
```

# LangChain

## Chains

```
from langchain.prompts import PromptTemplate
from langchain.llms import OpenAI
from langchain.chains import LLMChain
llm = OpenAI(temperature=0.9)
template = "What is a good name for a company that makes {product}?"

prompt = PromptTemplate(input_variables=["product"], template=template)

chain = LLMChain(llm=llm, prompt=prompt)

chain.run("colorful socks")
```



# LangChain

## Memory

```
from langchain.memory import ChatMessageHistory
```

```
history = ChatMessageHistory()
```

```
history.add_user_message("hi!")
```

```
history.add_ai_message("whats up?")
```

# LangChain

## Indexes

```
from langchain.document_loaders import NotionDirectoryLoader
from langchain.document_loaders import PyPDFLoader
from langchain.document_loaders import UnstructuredEmailLoader
```

```
loader = NotionDirectoryLoader("Notton_DB")
```

```
loader = PyPDFLoader("d21-en.pdf")
```

```
loader = UnstructuredEmailLoader("example-email.eml")
```

```
data = loader.load()
```

# LangChain

## Agents & Tools

```
from langchain.agents import load_tools
from langchain.agents import initialize_agent
from langchain.llms import OpenAI
llm = OpenAI(temperature=0)

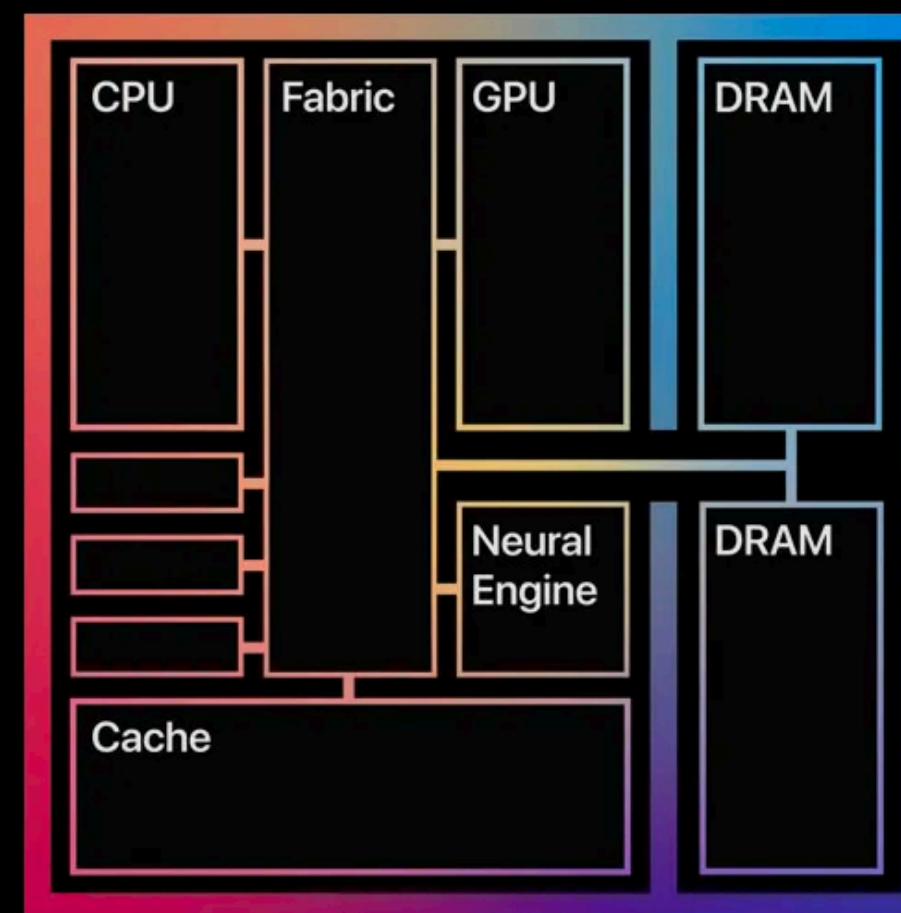
tools = load_tools(["google-search", "wikipedia", "llm-math"], llm=llm)

agent = initialize_agent(tools, llm, agent="zero-shot-react-description")

result = agent.run("Who Is Leo DiCaprio's girlfriend? What is her current
age raised to the 0.43 power?")
```

# About Apple Silicon

- Shared CPU and GPU memory
- PyTorch supports Apple Silicon Metal Performance Shader
- Apple support (<https://github.com/apple/ml-stable-diffusion>)



# How to run Stable Diffusion on your Mac

- <https://github.com/isekaidev/stable.art> (Photoshop plugin)

```
git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui  
tea --cd stable-diffusion-webui +python.org~3.10 +pip.pypa.io \  
+gnu.org/wget +protobuf.dev +rust-lang.org ./webui.sh
```

- Wait for all the model to download

# How to run LLaMA on your Mac

- With tea installed, just type:

```
llama.cpp -p "Hello World"
```

- Wait for all the model to download (can take a looong time)
- It's this: <https://github.com/ggerganov/llama.cpp/discussions>

# Stelfie Time Traveller



Questions?