UNIVERSITIES SOUTH AFRICA

AIMS African Institute for Mathematical Sciences SOUTH AFRICA

Technology and Human Interface in the Future University

Ulrich Paquet

Nobel Prize in Physics

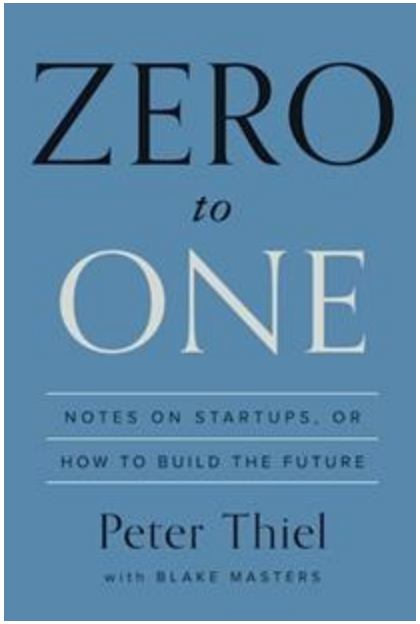
The 2024 physics laureates

The Nobel Prize in Physics 2024 was awarded to John J. Hopfield and Geoffrey E. Hinton "for foundational discoveries and inventions that enable machine learning with artificial neural networks."

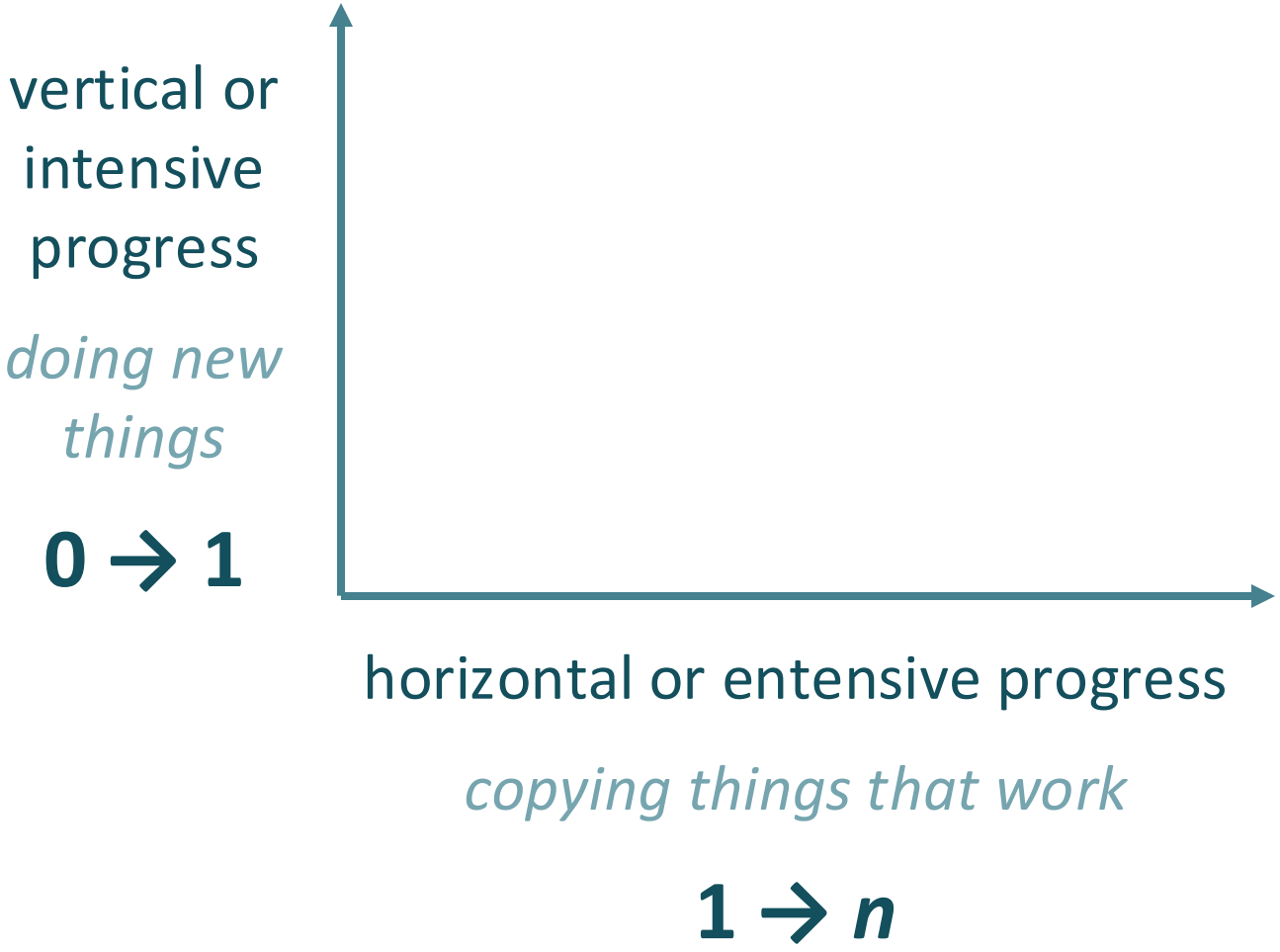
Hopfield created a structure that can store and reconstruct information. Hinton invented a method that can independently discover properties in data and which has become important for the large neural networks now in use.

Ill. Niklas Elmehed © Nobel Prize Outreach

How my lenses are tinted



What important truth do very few people agree with you on?


vertical or intensive progress
doing new things
0 → 1
horizontal or entensive progress
copying things that work
1 → n

What important truth do very few people agree with you on?

What important truth do very few people agree with you on?

“The university, in its present form, will be obsolete in two decades”

- Great resources freely available
- MOOCs
- Student debt
- Components are on the table to disrupt a traditional industry
- Rise of AI

“Lectures were once useful; but now, when all can read, and books are so numerous, lectures are unnecessary.”

Samuel Johnson (~1781)

“With good probability, the university, in its present form, will be obsolete in two decades”

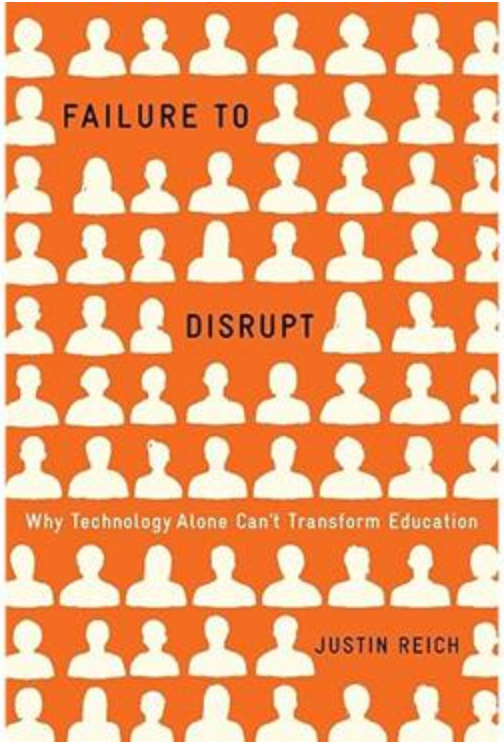
No!

Social

Extrinsic motivation

(peer pressure)

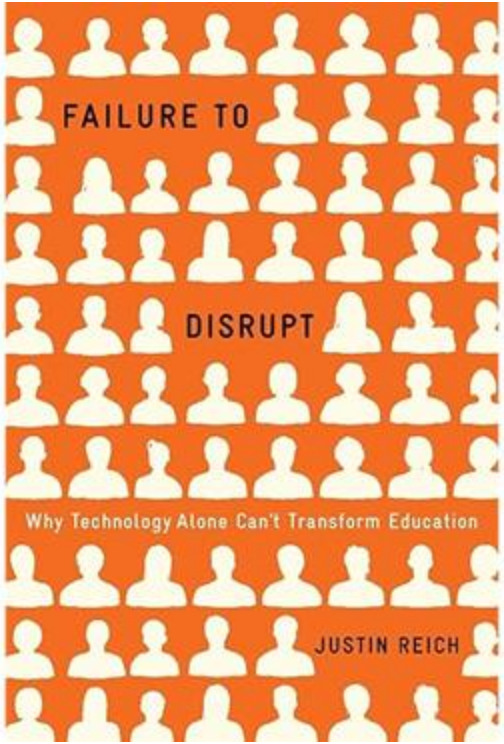
MOOCs = 🚀 ?



Three big promises:

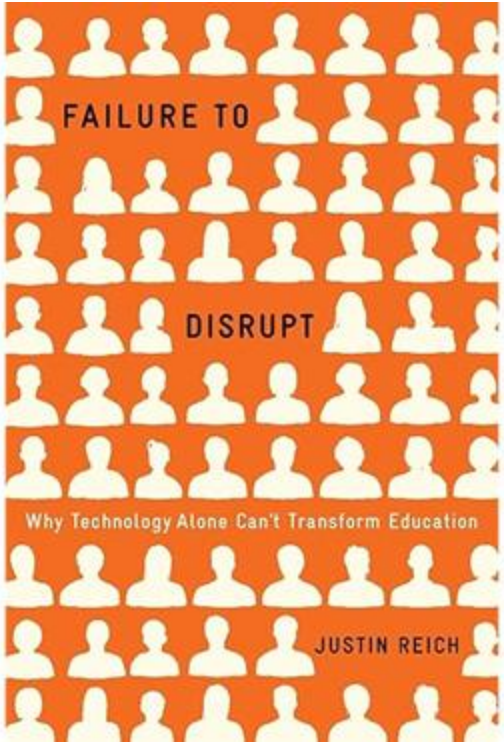
1. expanded access
2. transformed systems
3. better teaching from data driven research

MOOCs = 🚀 ?



Successful MOOC learners demonstrate effective use of self-regulated learning strategies: goal setting, time management, help seeking, self-monitoring…

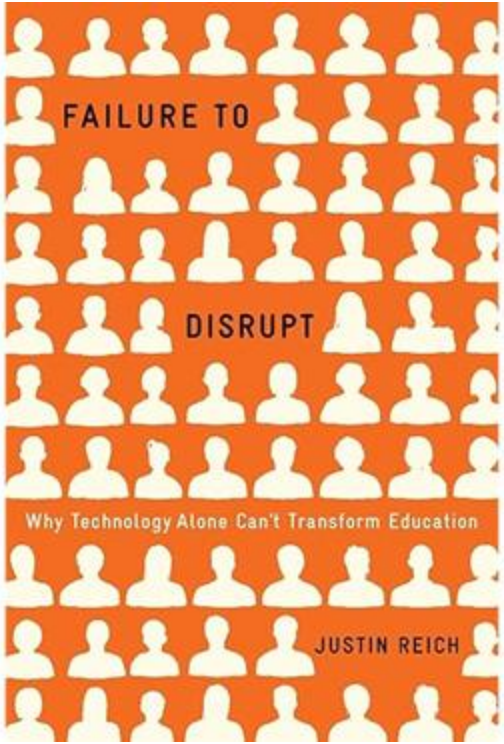
MOOCs = 🚀 ?



Successful MOOC learners demonstrate effective use of self-regulated learning strategies: goal setting, time management, help seeking, self-monitoring…

~80% have a bachelor's degree already

MOOCs = 🚀 ?



Successful MOOC learners demonstrate effective use of self-regulated learning strategies: goal setting, time management, help seeking, self-monitoring…

~80% have a bachelor's degree already

Best serves the already educated pursuing advanced postsecondary learning - this in turn determines in which disciplines most courses are developed

social inequality is a tenacious feature of educational systems

- Great resources freely available
- MOOCs
- Student debt
- Components are on the table to disrupt a traditional industry
- **Rise of AI**

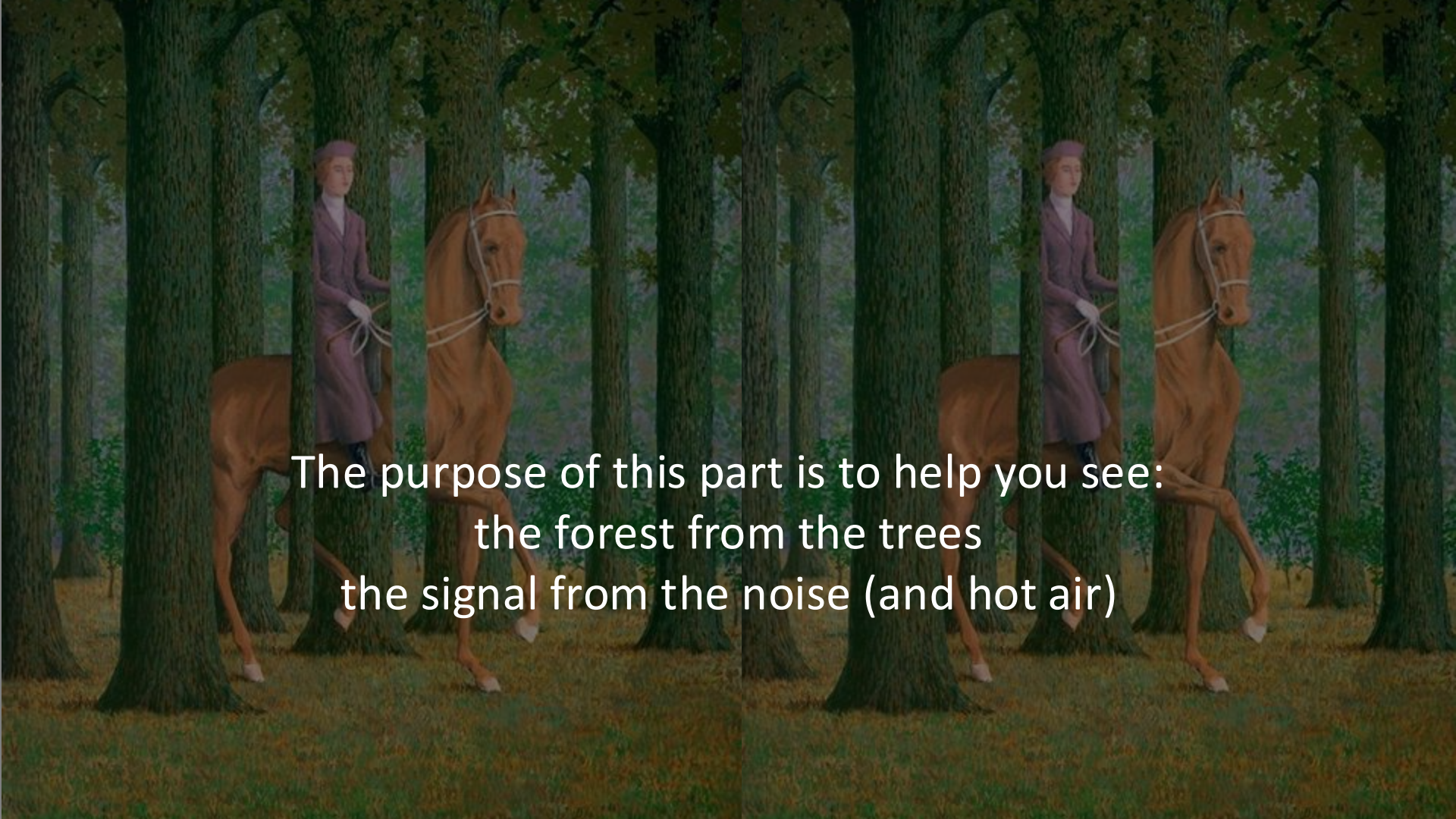
A new wind is blowing (fast)

A modern AI system* can have a >1 million token context window – enough to take these books as input and output a summary; an essay…

student query

you're welcome!

The fall of the Western Roman Empire was a long and complex process, not a single event. It was caused by a combination of factors, including economic decline, political instability, military overextension, and pressure from migrating Germanic tribes. The Empire gradually lost its ability to enforce its rule, leading to the division of its territory into various successor states. While the traditional date for the fall of the Western Roman Empire is 476 CE, when the last Roman emperor was deposed, the decline had been occurring for centuries prior. The Eastern Roman Empire, also known as the Byzantine Empire, continued to thrive for another thousand years.


The purpose of this part is to help you see:
the forest from the trees
the signal from the noise (and hot air)

A few challenges


A very
IMPORTNT
THING
TO DO

A few challenges

A few challenges

How many people were there?

How many presidents were there?

Did you see anything unusual?

What would each person eat?

From which direction was the sun shining?

Of what other things were you reminded?

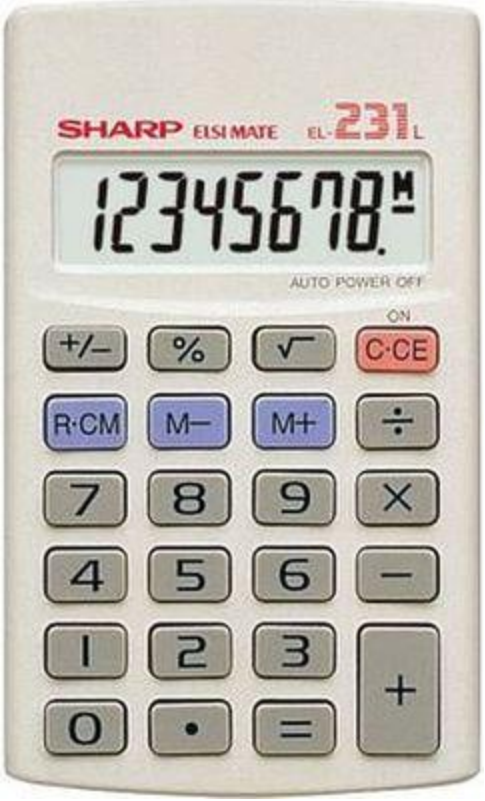
You were fast!

Another challenge

What is **0.4235 x 1.84**?

Anyone...?

?


SHARP ELSI MATE EL-231L
12345678.M
AUTO POWER OFF
ON
+/–
%
√
C·CE
R·CM
M–
M+
÷
7
8
9
×
4
5
6
–
1
2
3
+
0
•
=

What is artificial intelligence?

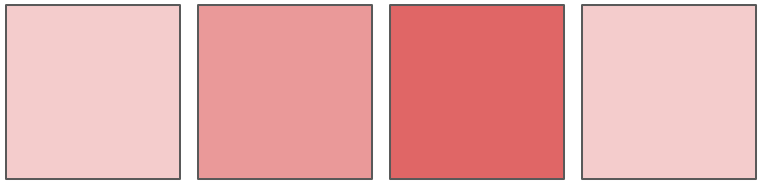
What is so hard?

What is so hard?

What is so hard?

What is so hard?

What is so hard?



What is so hard?

221 154 101 221

What is so hard?

221	154	101	221

A computer only sees a looooooong sequence of numbers!

Let’s try!

0	red pixels = 0
100	lighter red pixels = 100
200	light pink pixels = 200
255	white pixels = 255

Is this pixel white or light pink?

221

Is this pixel white or light pink?

221

It is closer to 255 (white) than 0 (red) so it must be light pink!

221 > 150 (a number that you chose)

Is this pixel white or light pink?

221

Do these three pixels look light pink from afar?

221

154

221

Do these three pixels look light pink from afar?

⅓ x 221 + ⅓ x 154 + ⅓ x 221

= 199

> 150 (the number that you chose)

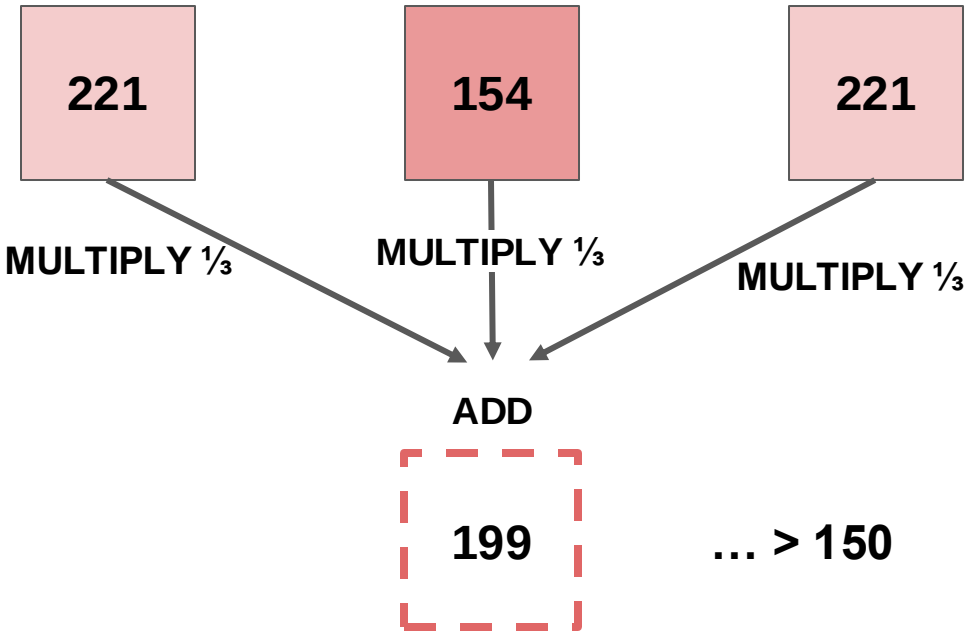
The building blocks of an *artificial* neural network!

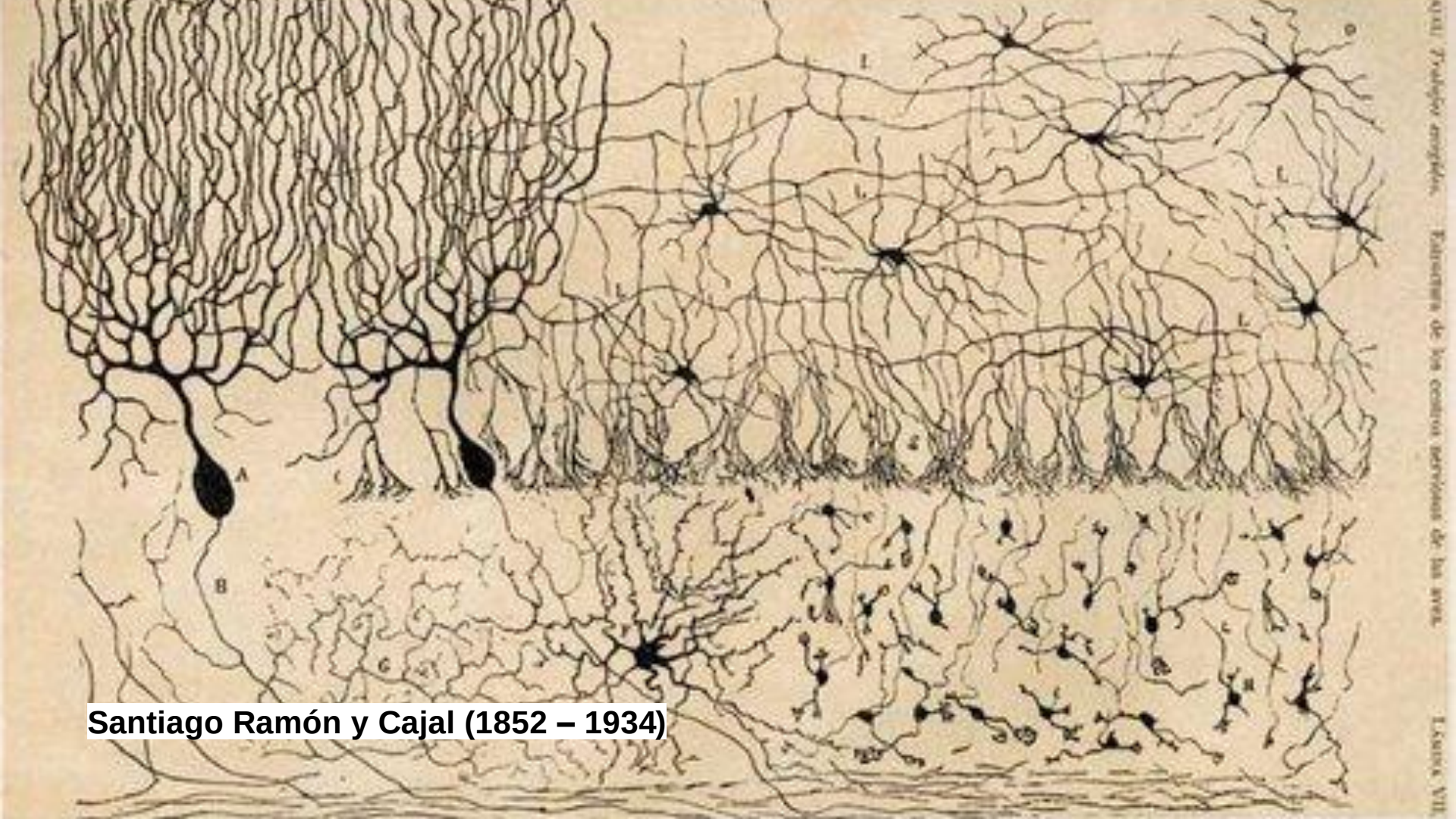
221

154

221

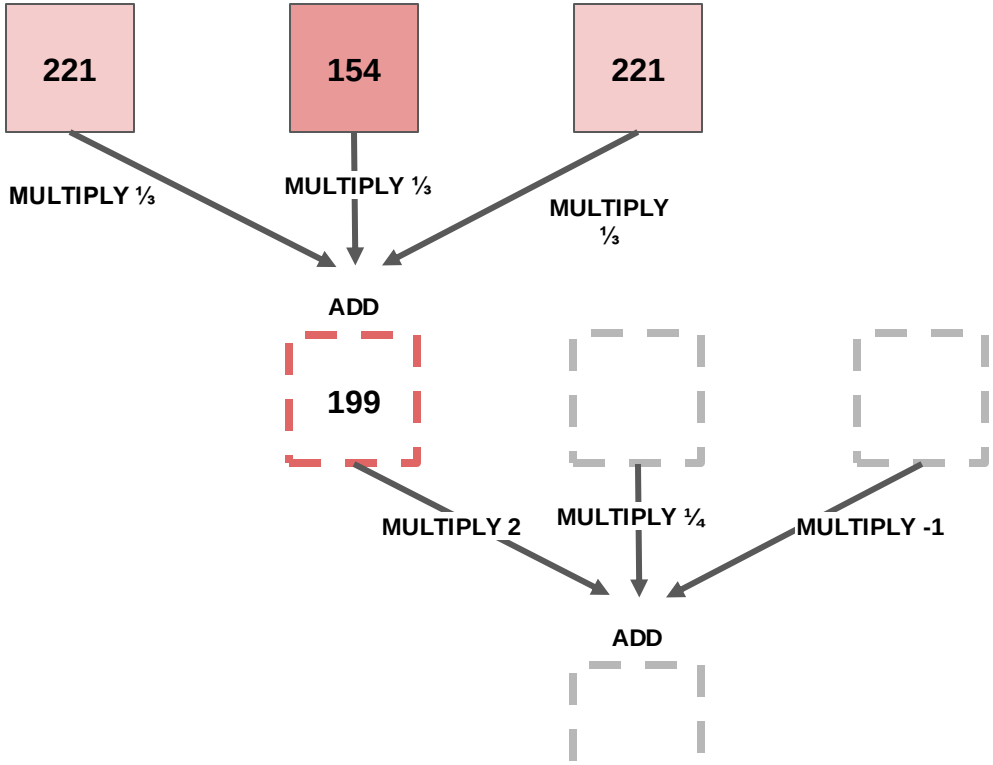
The building blocks of an *artificial* neural network!

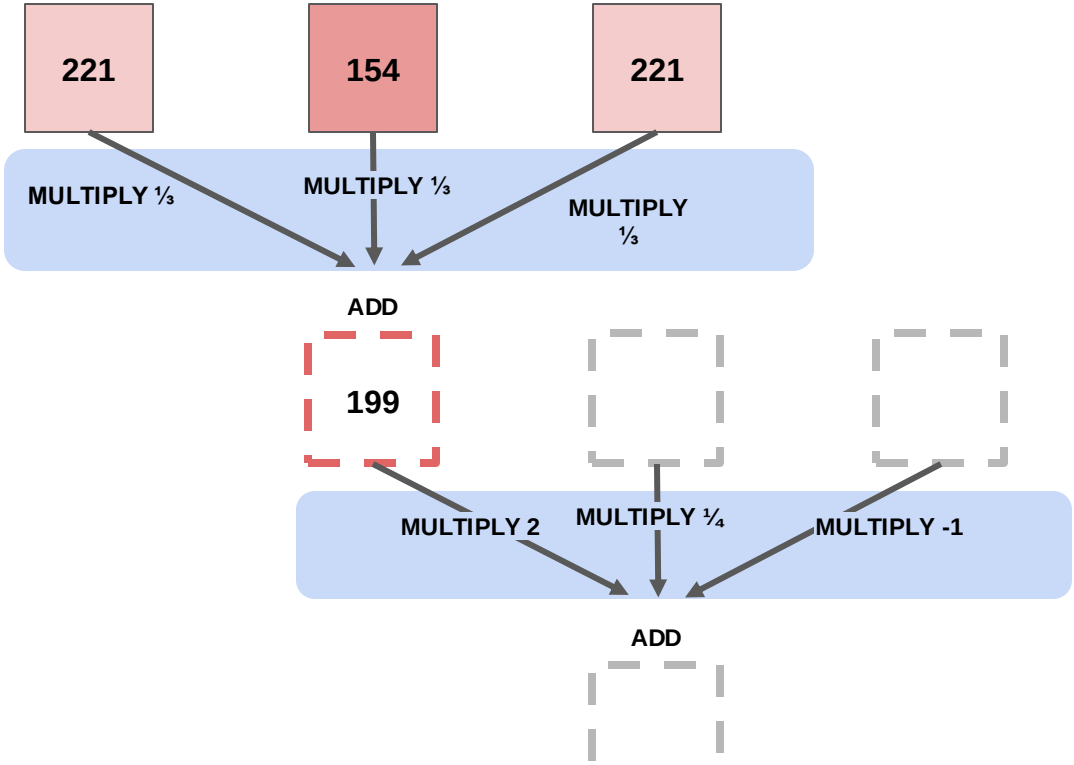




Santiago Ramón y Cajal (1852 – 1934)

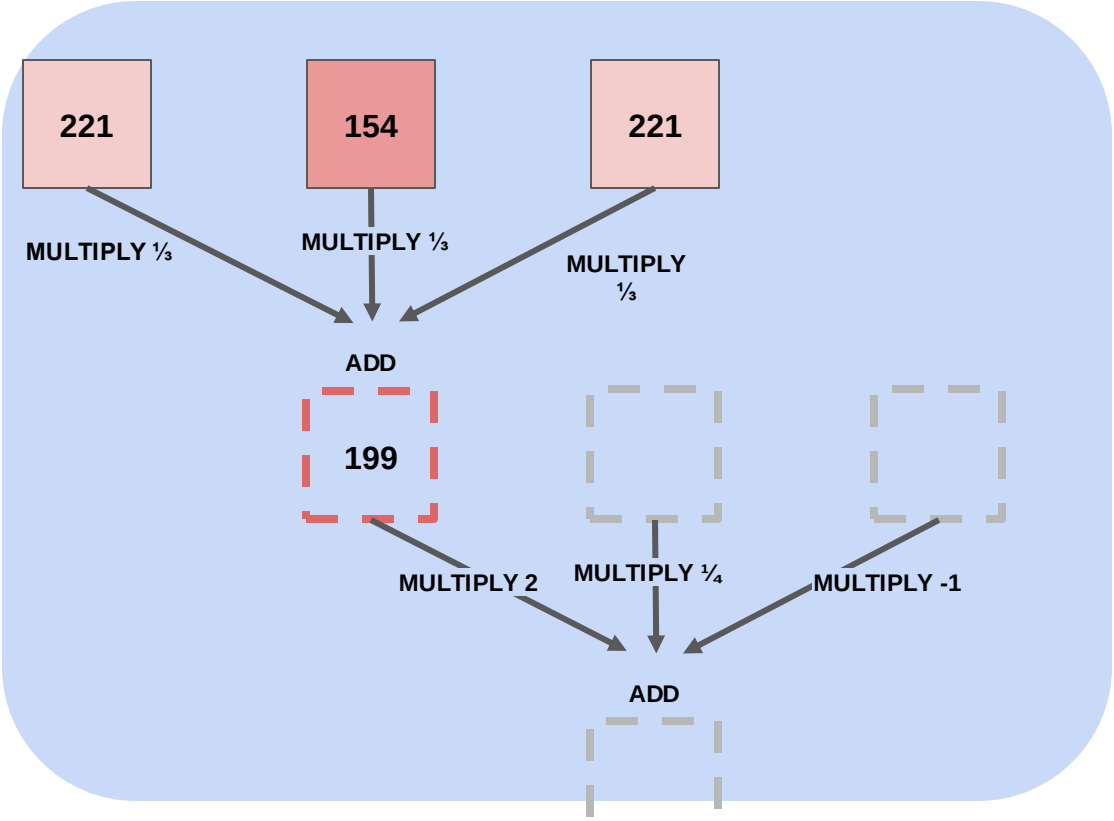
The building blocks of an *artificial* neural network!





Hey, we have six **parameters** here!

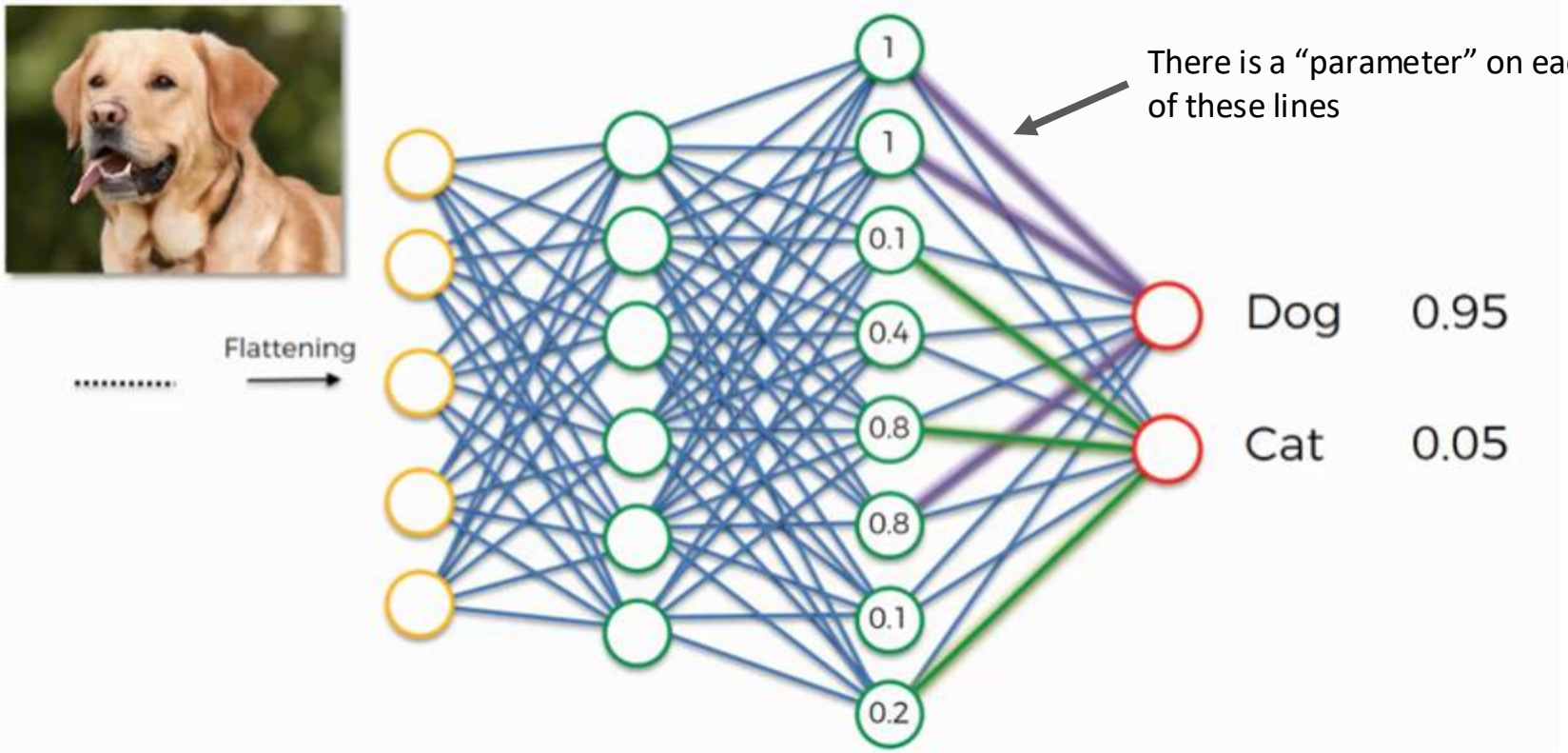
⅓, ⅓, ⅓, 2, ¼, -1



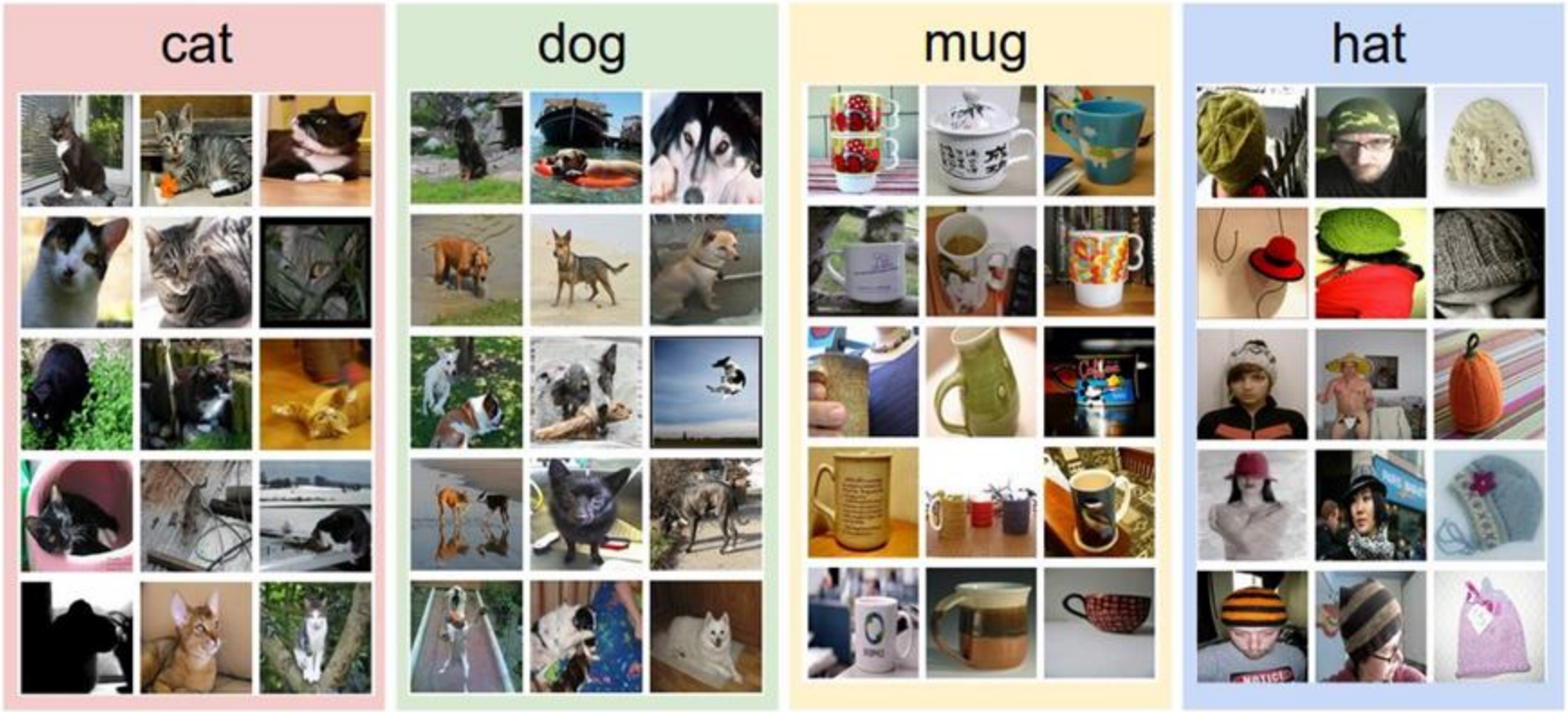
This configuration is referred to as a "**model**" or "**architecture**"

Last bit of terminology: “training” or “learning”

An *artificial* neural network!

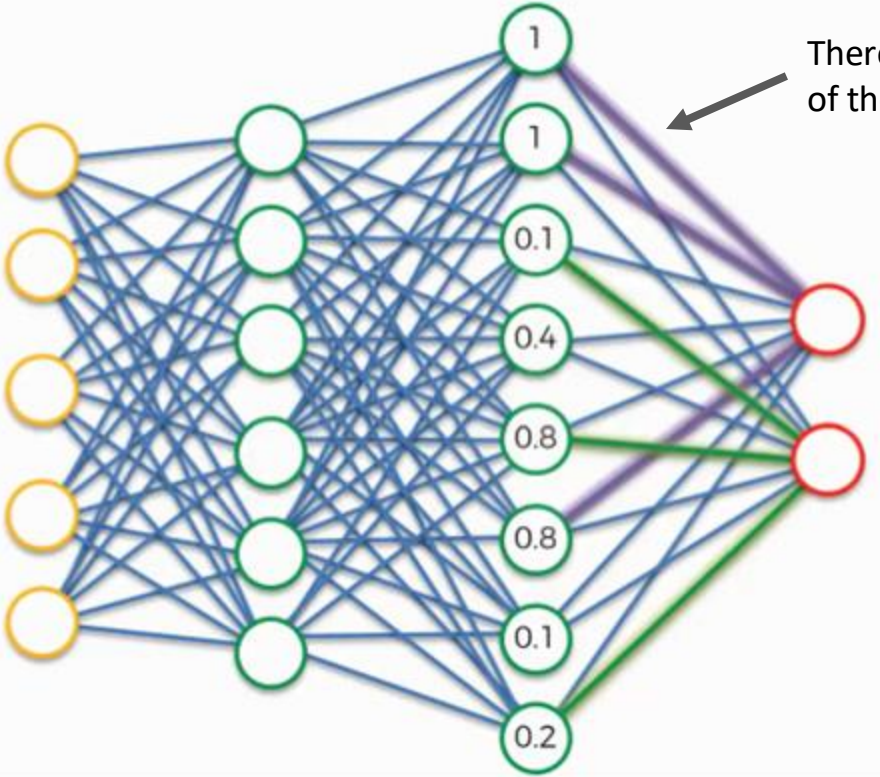


How does an *artificial* neural network learn?



How does an *artificial* neural network learn?

Flattening



There is a “parameter” on each of these lines

Dog

Cat

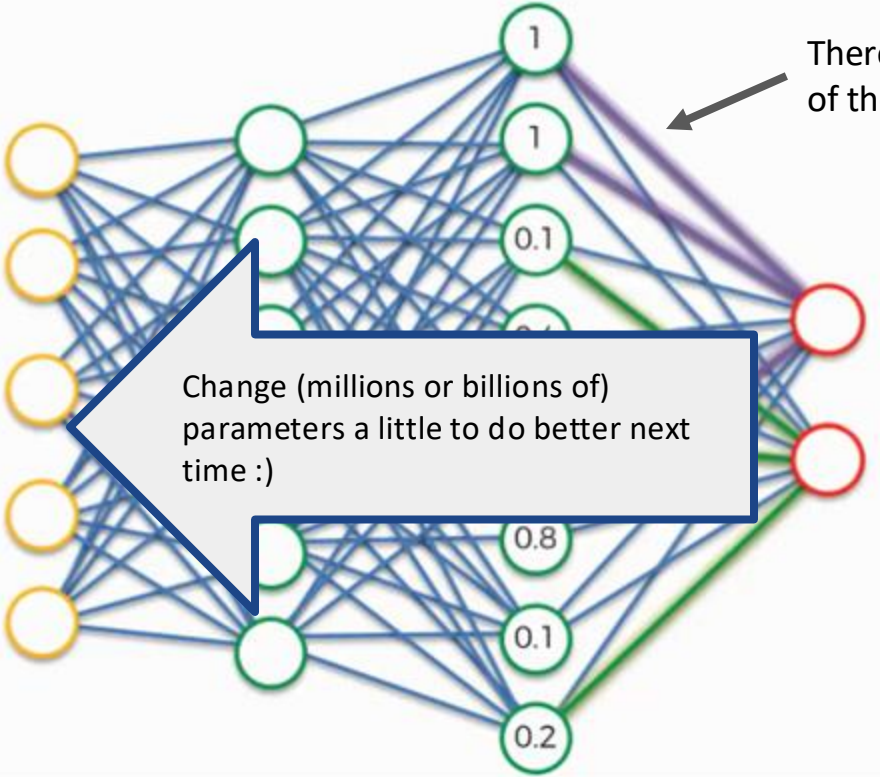
5%

95%

Oh no!!

How does an *artificial* neural network learn?

Flattening



There is a “parameter” on each of these lines

Change (millions or billions of) parameters a little to do better next time :)

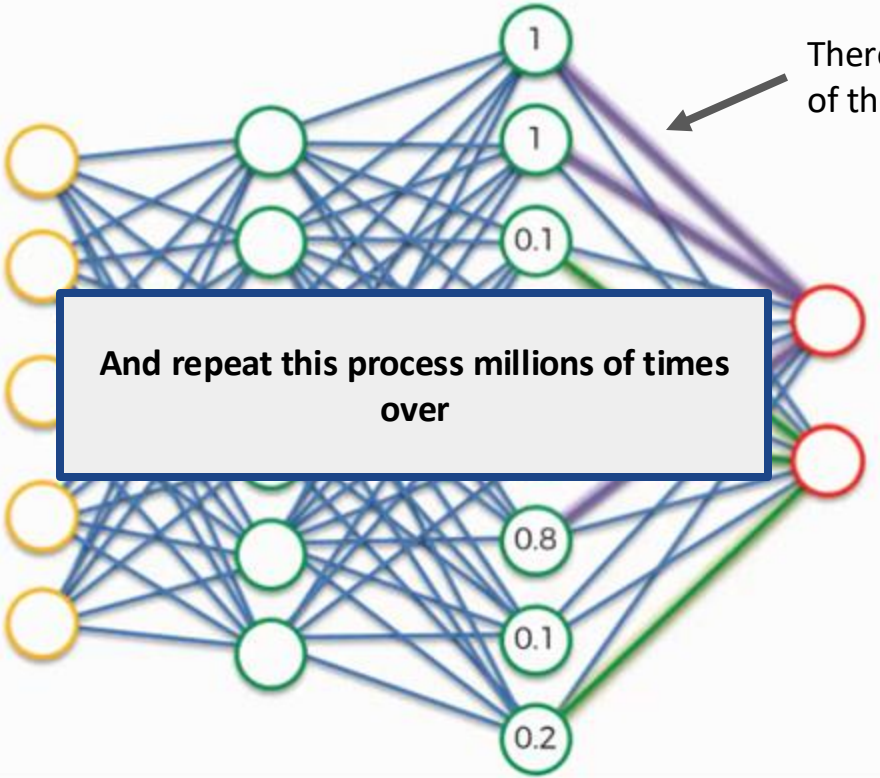
5%

95%

Oh no!!

How does an *artificial* neural network learn?

Flattening



There is a “parameter” on each of these lines

And repeat this process millions of times over

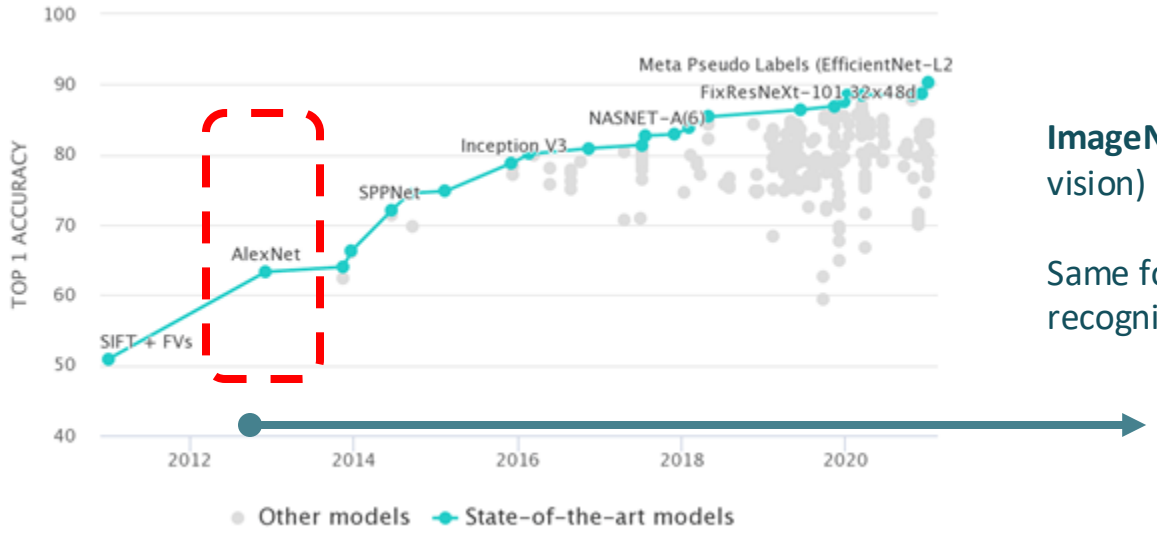
7 big deals

1. What's the big deal?

Big deal #1

This stuff started paying dividends ~2013. Before then, top techniques looked very different!

top results for a task (speech, computer vision, etc.)



ImageNet (computer vision)

Same for speech recognition (MSR)

2. What's the big deal?

Big deal #2

The "models" you read about have >100,000,000,000 **parameters**

The parameters can be "tuned" because of a combination of:

- progress in **computation**

2. What's the big deal?

Big deal #2

The "models" you read about have >100,000,000,000 **parameters**

The parameters can be "tuned" because of a combination of:

- progress in **computation**
- progress in "**model architectures**"

2. What's the big deal?

Big deal #2

The "models" you read about have >100,000,000,000 **parameters**

The parameters can be "tuned" because of a combination of:

- progress in **computation**
- progress in "**model architectures**"
- progress in **algorithms** to set them

2. What's the big deal?

Big deal #2

The "models" you read about have >100,000,000,000 **parameters**

The parameters can be "tuned" because of a combination of:

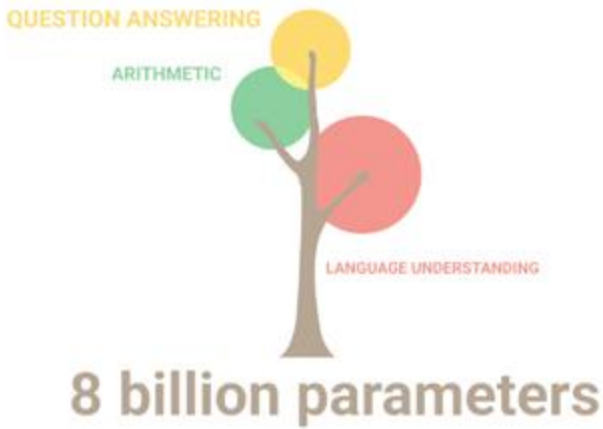
- progress in **computation**
- progress in "**model architectures**"
- progress in **algorithms** to set them
- increased **data** set sizes

} **technological advances in all 4 things had to happen**

3. What's the big deal?

Big deal #3

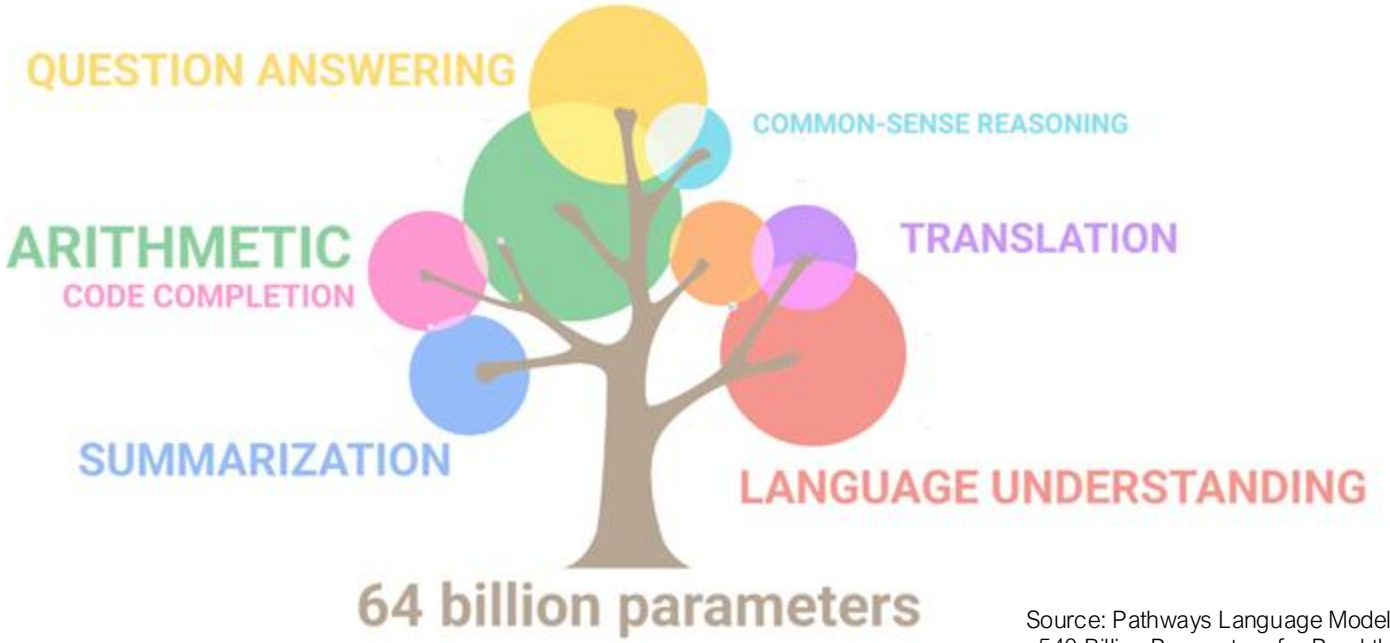
Beyond a certain point, **scaling unlocks emergent behaviour <u>that wasn't predictable</u>**



Source: Pathways Language Model (PaLM): Scaling to 540 Billion Parameters for Breakthrough Performance

3. What's the big deal?

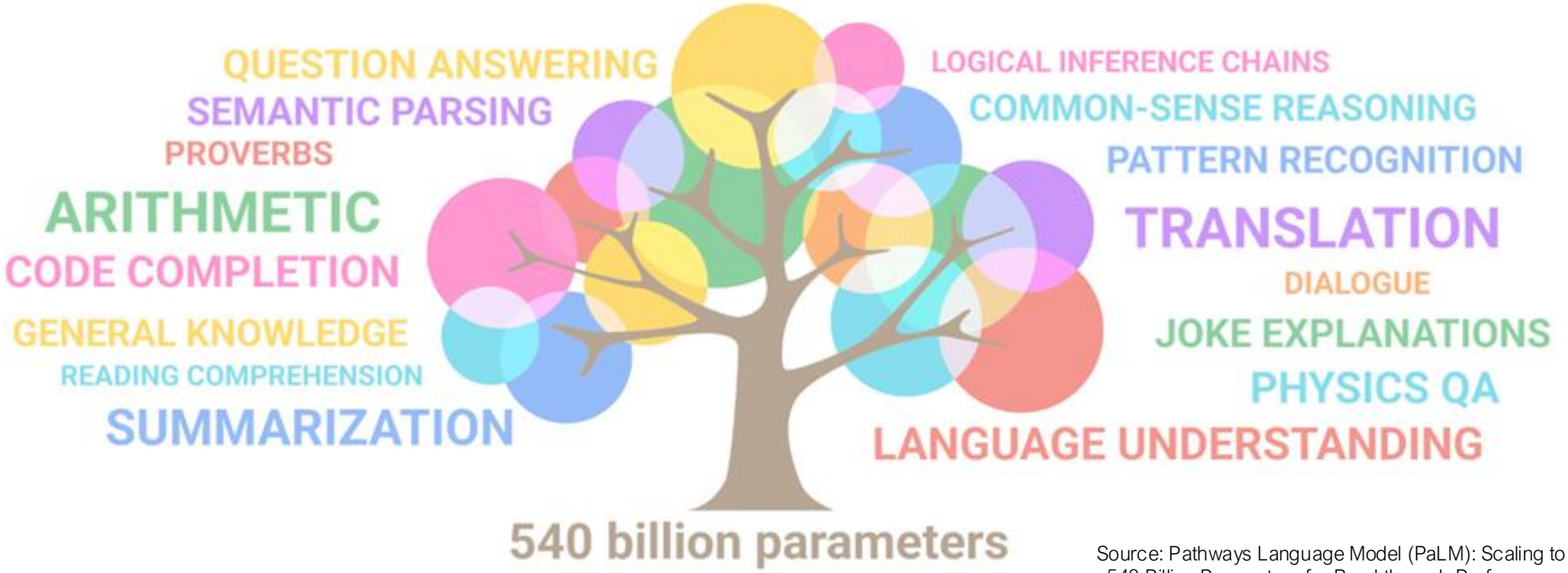
Big deal #3



Source: Pathways Language Model (PaLM): Scaling to 540 Billion Parameters for Breakthrough Performance

3. What's the big deal?

Big deal #3



Source: Pathways Language Model (PaLM): Scaling to 540 Billion Parameters for Breakthrough Performance

3. Unpredictable scaling and emergent behaviour

Big deal #3

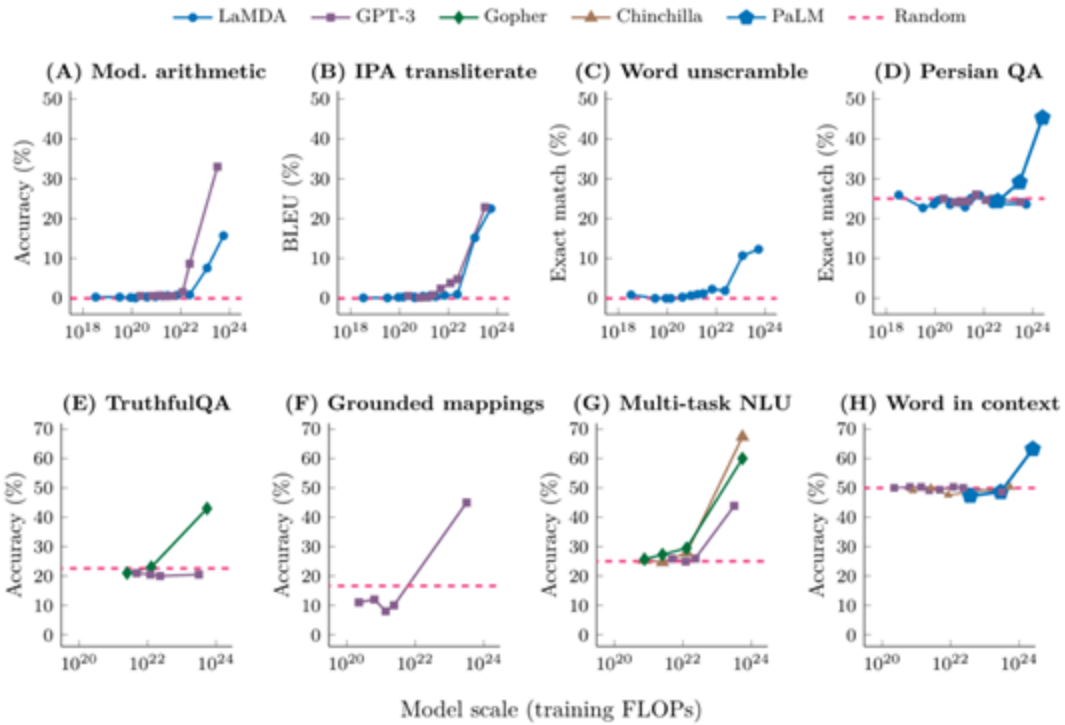


Figure 2: Eight examples of emergence in the few-shot prompting setting. Each point is a separate model. The ability to perform a task via few-shot prompting is emergent when a language model achieves random performance until a certain scale, after which performance significantly increases to well-above random. Note

3. Unpredictable scaling and emergent behaviour

Big deal #3

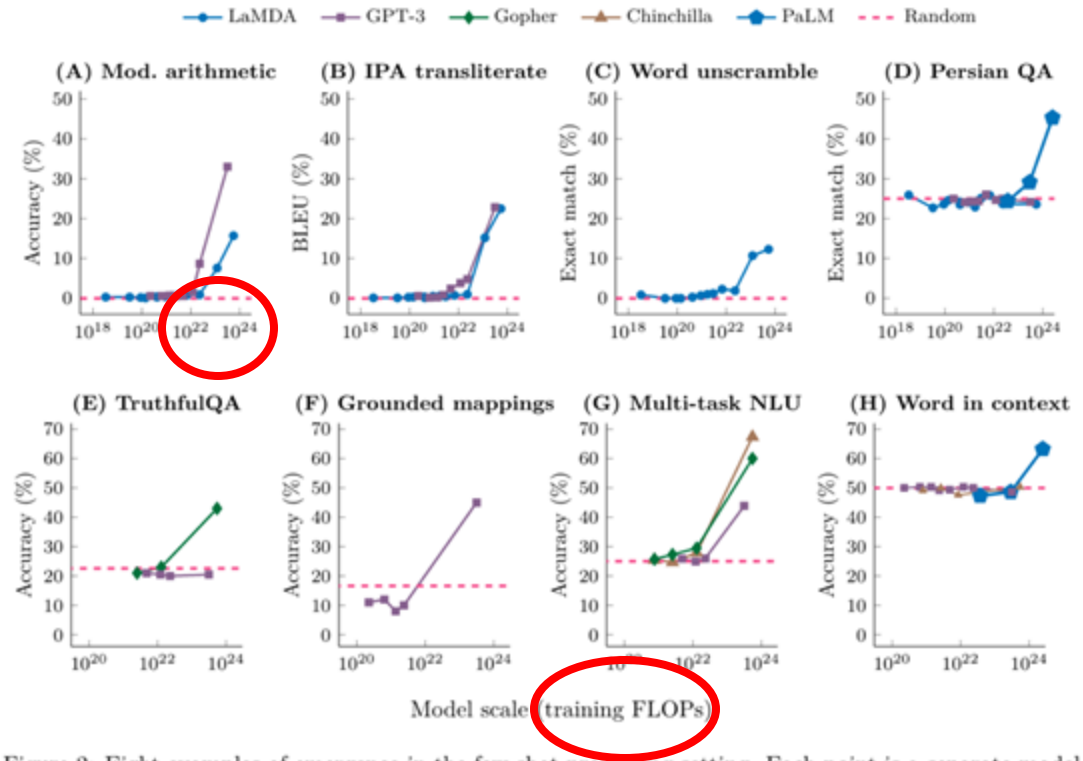


Figure 2: Eight examples of emergence in the few-shot prompting setting. Each point is a separate model. The ability to perform a task via few-shot prompting is emergent when a language model achieves random performance until a certain scale, after which performance significantly increases to well-above random. Note

TECHNOLOGY

End of the road for AI advisers 'blindsided' by latest tech

The sudden release of large-language model artifical intelligence caught many in the tech industry by surprise

FLORENCE LO/REUTERS

Share

Downing Street is overhauling its advisers on artificial intelligence amid criticism that they were caught out by the advances in large language models such as ChatGPT.

The AI Council, a committee of experts that advised ministers and The Alan Turing Institute (ATI), the national institute for data science and AI, have been accused of being blindsided by the importance of the technology. Both bodies deny the claims.

The government has now appointed Ian Hogarth, the tech entrepreneur, as chairman of the new foundation model taskforce, which will spearhead the adoption and regulation of the technology in the UK.

The taskforce will be the central forum of outside advice for the government, replacing the AI Council, and there will be a rotating panel of experts ministers can call on ad hoc.

Tech entrepreneur Ian Hogarth to lead UK's AI Foundation Model Taskforce

Artificial intelligence expert announced as chair of government's Foundation Model Taskforce.

From: Department for Science, Innovation and Technology, Chloe Smith MP, and The Rt Hon Rishi Sunak MP

Published 18 June 2023

- Leading tech entrepreneur and renowned investor and AI specialist will chair the Foundation Model Taskforce.

The renowned tech investor, entrepreneur and AI specialist Ian Hogarth has been announced as the chair of the Government's Foundation Model Taskforce, reporting directly to the Prime Minister and Technology Secretary.

4. What's the big deal?

Big deal #4

It is not a cheap exercise; innovation driven by **industry**

10^{24} "training floating point operations" =

10,000,000,000,000,000,000,000,000 of them

4. What's the big deal?

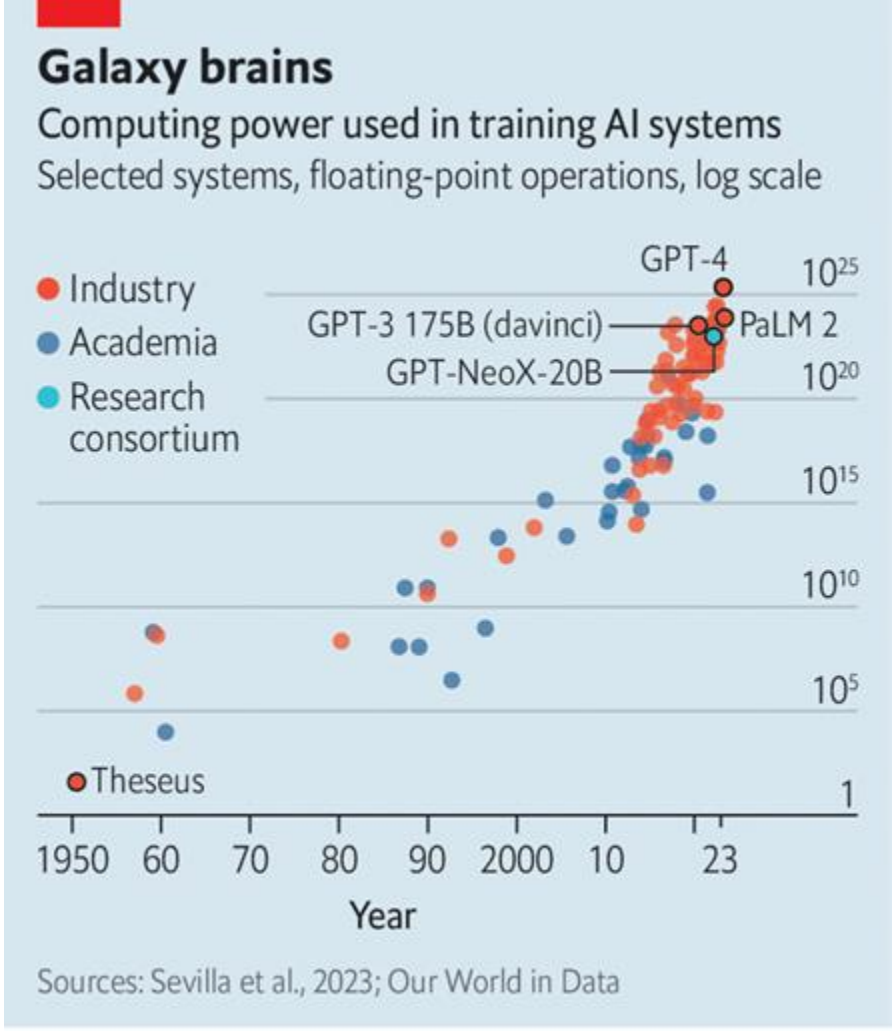
Big deal #4

It is not a cheap exercise; innovation driven by **industry**

10^{24} "training floating point operations" =

10,000,000,000,000,000,000,000,000 of them

Big deal #4



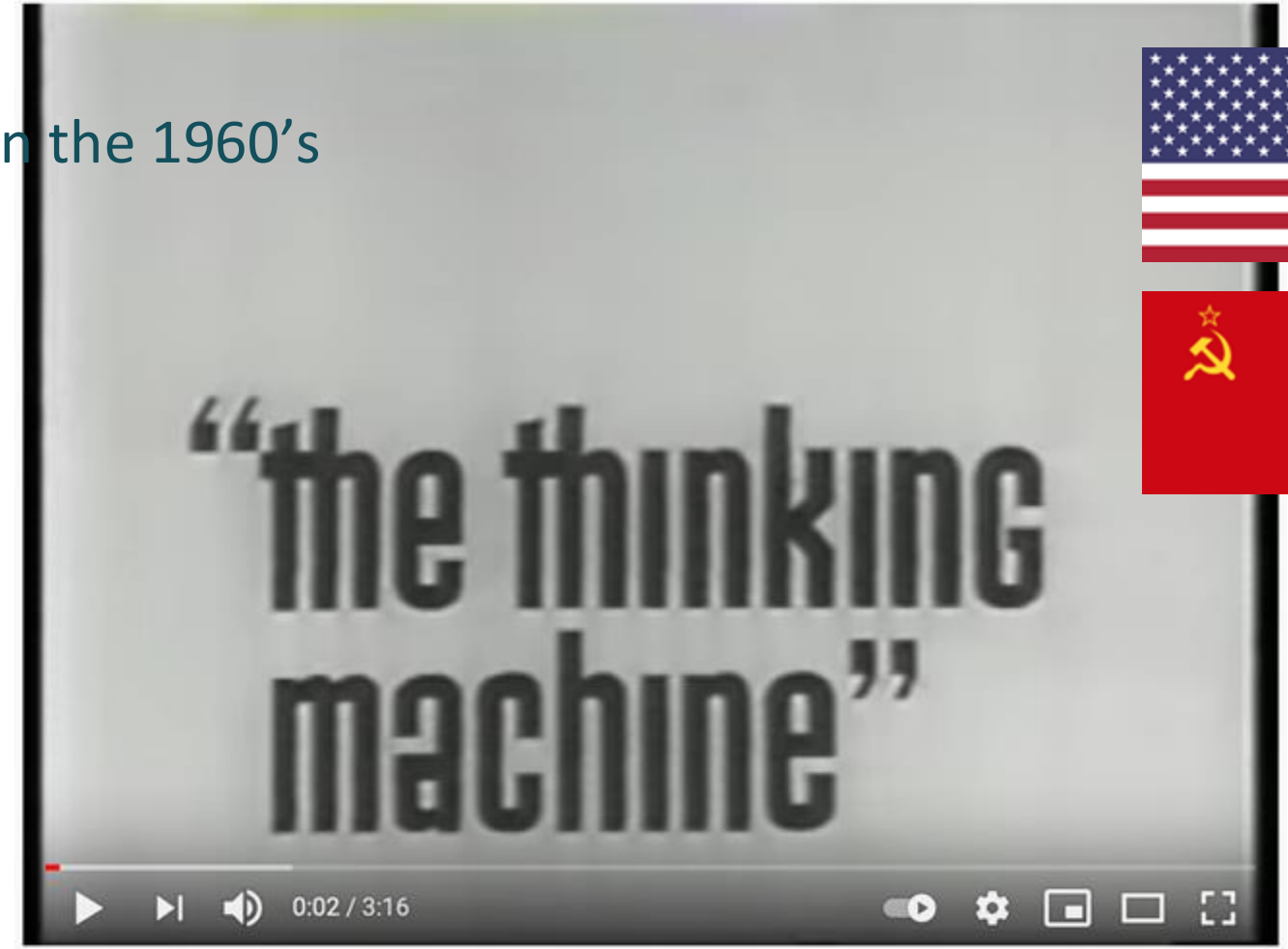
The Economist

5. What’s the big deal?

Big deal #5

Disruptive progress is (exponentially) faster

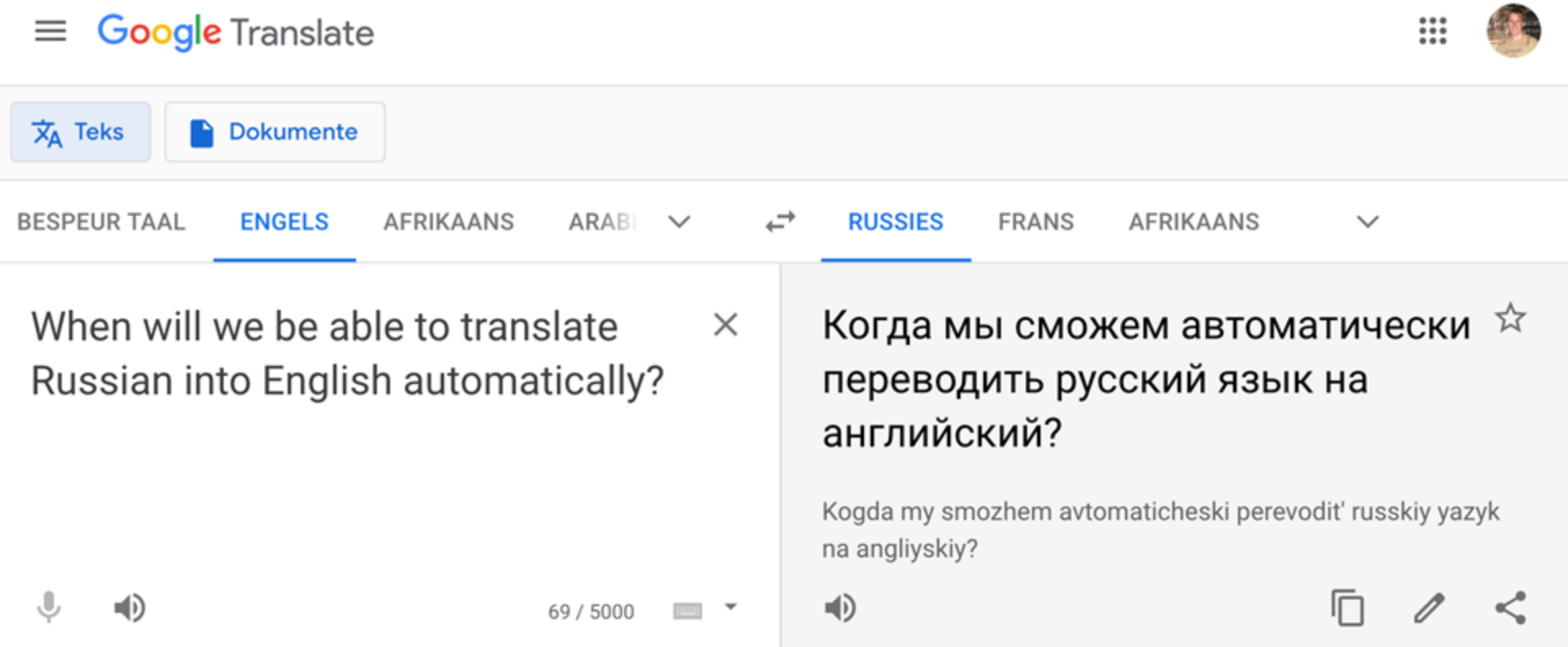
5. MIT in the 1960's



The Thinking Machine (Artificial Intelligence in the 1960s)

video link: https://youtu.be/aygSMgK3BEM?si=lCuWlZayXMRiw-GQ

5. Fifty years later...



5. A personal story

Spoken dialogue systems, around 2014-15

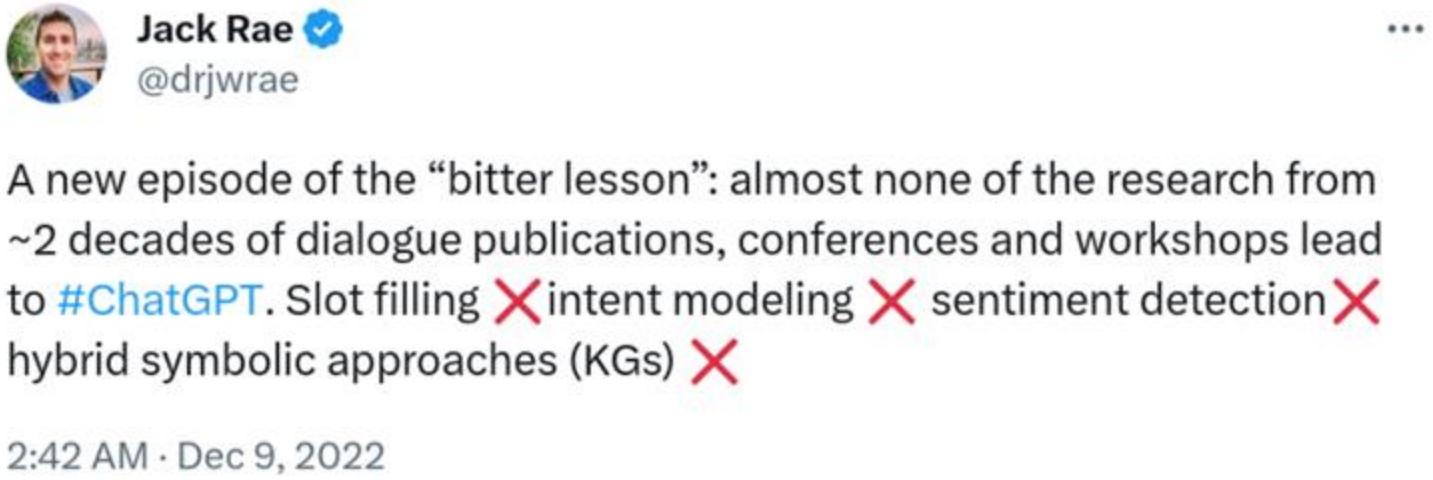
Apple buys UK-based speech technology start-up VocalIQ

Cambridge group's technology could be used to improve Siri, the iPhone maker's virtual assistant

© Getty

Tim Bradshaw in San Francisco, **Sally Davies** and **Murad Ahmed** in London OCTOBER 2 2015

5. Only 7 years later



Jack Rae
@drjwrae

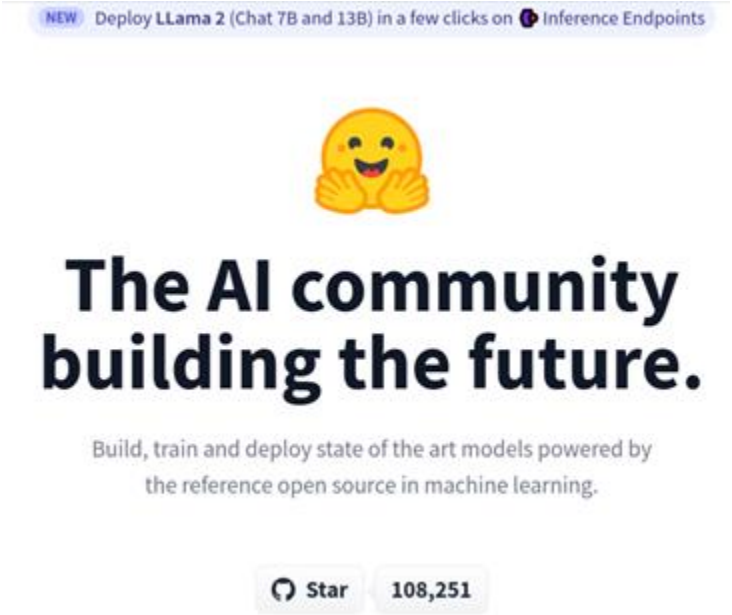
A new episode of the “bitter lesson”: almost none of the research from ~2 decades of dialogue publications, conferences and workshops lead to #ChatGPT. Slot filling ❌ intent modeling ❌ sentiment detection ❌ hybrid symbolic approaches (KGs) ❌

2:42 AM · Dec 9, 2022

6. What's the big deal?

Big deal #6

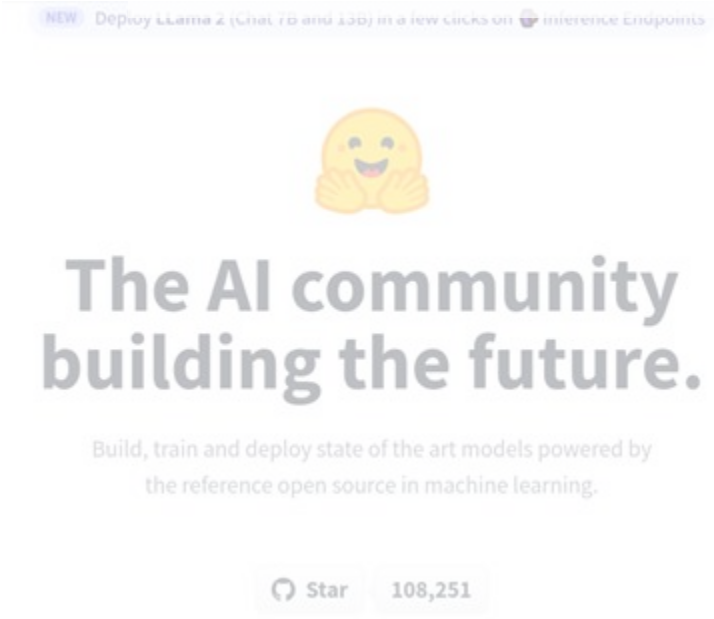
Startups and open models abound. It's not a stationary world



6. What's the big deal?

Big deal #6

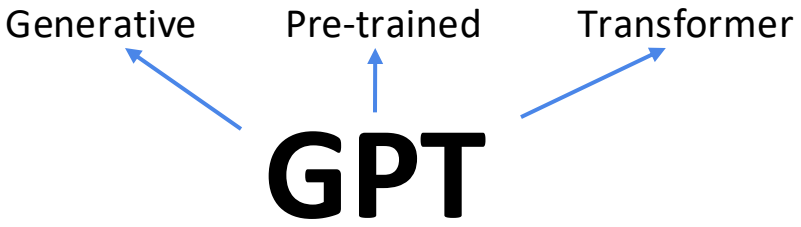
Startups and open models abound. It's not a stationary world



6. What's the big deal?

Big deal #6

Startups and open models abound. It's not a stationary world



Pre-trained means that the "big investment" in training is done. You get a copy and **fine-tune for a task**

7. What’s the big deal?

Big deal #7

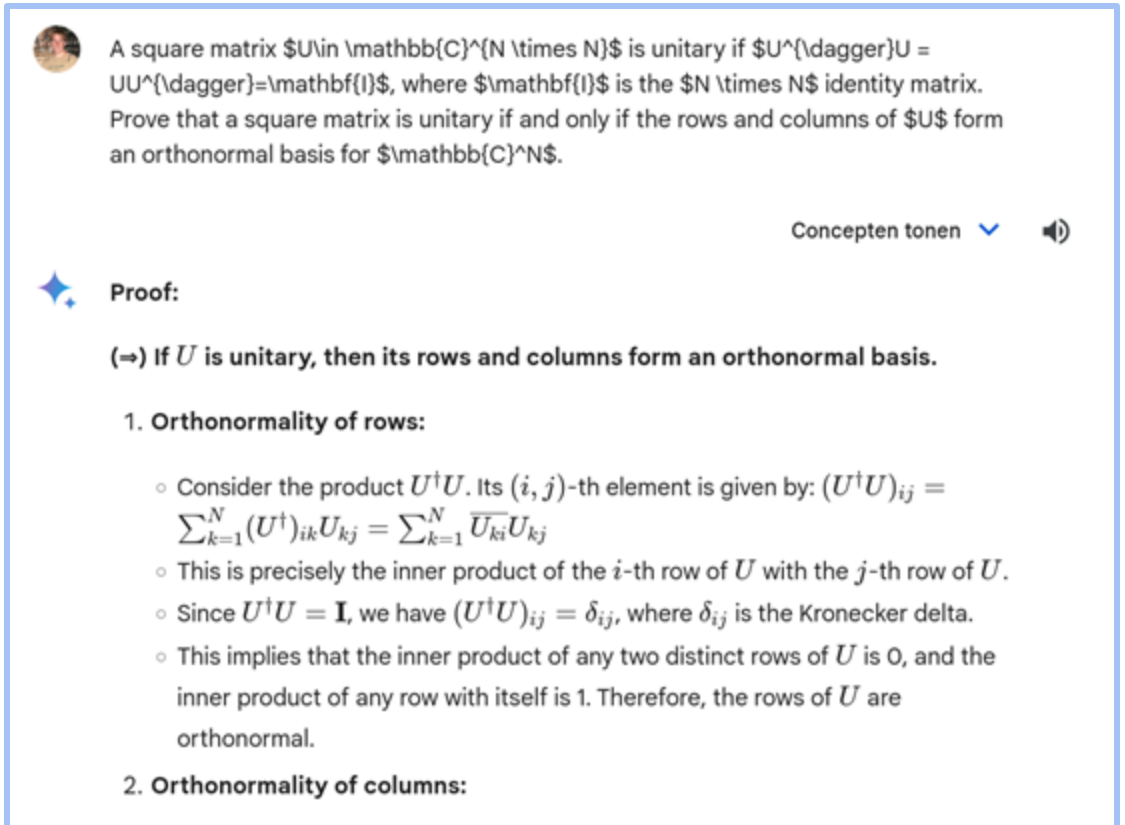
Huge disruption to education

Current LLMs pushing context widths > 1M tokens.

**A good historical textbook can be part of a query

7. What's the big deal?

Big deal #7



A square matrix $U\in \mathbb{C}^{N \times N}$ is unitary if $U^{\dagger}U = UU^{\dagger}=\mathbf{I}$, where $\mathbf{I}$ is the $N \times N$ identity matrix. Prove that a square matrix is unitary if and only if the rows and columns of U form an orthonormal basis for $\mathbb{C}^N$.

Concepten tonen

Proof:

(⇒) If U is unitary, then its rows and columns form an orthonormal basis.

1. **Orthonormality of rows:**
   - Consider the product $U^\dagger U$. Its (i, j)-th element is given by: $(U^\dagger U)_{ij} = \sum_{k=1}^{N} (U^\dagger)_{ik} U_{kj} = \sum_{k=1}^{N} \overline{U_{ki}} U_{kj}$
   - This is precisely the inner product of the i-th row of U with the j-th row of U.
   - Since $U^\dagger U = \mathbf{I}$, we have $(U^\dagger U)_{ij} = \delta_{ij}$, where δ_{ij} is the Kronecker delta.
   - This implies that the inner product of any two distinct rows of U is 0, and the inner product of any row with itself is 1. Therefore, the rows of U are orthonormal.
2. **Orthonormality of columns:**

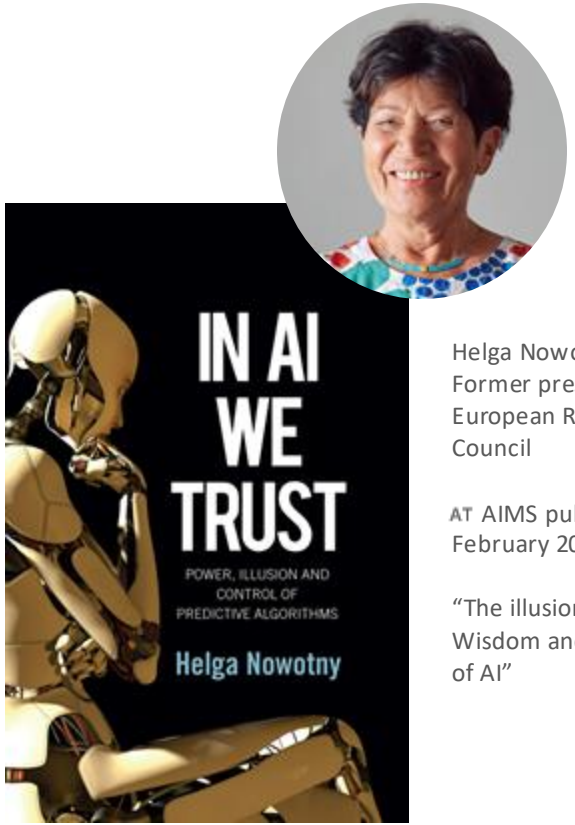
7. What's the big deal?

Big deal #7

Yes it's **disruptive**:

1. The way you assess students should change completely
2. Curriculums should be redesigned from scratch

Run an open competition to design new curriculums



Helga Nowotny
Former president of the European Research Council

AT AIMS public lecture, February 2024

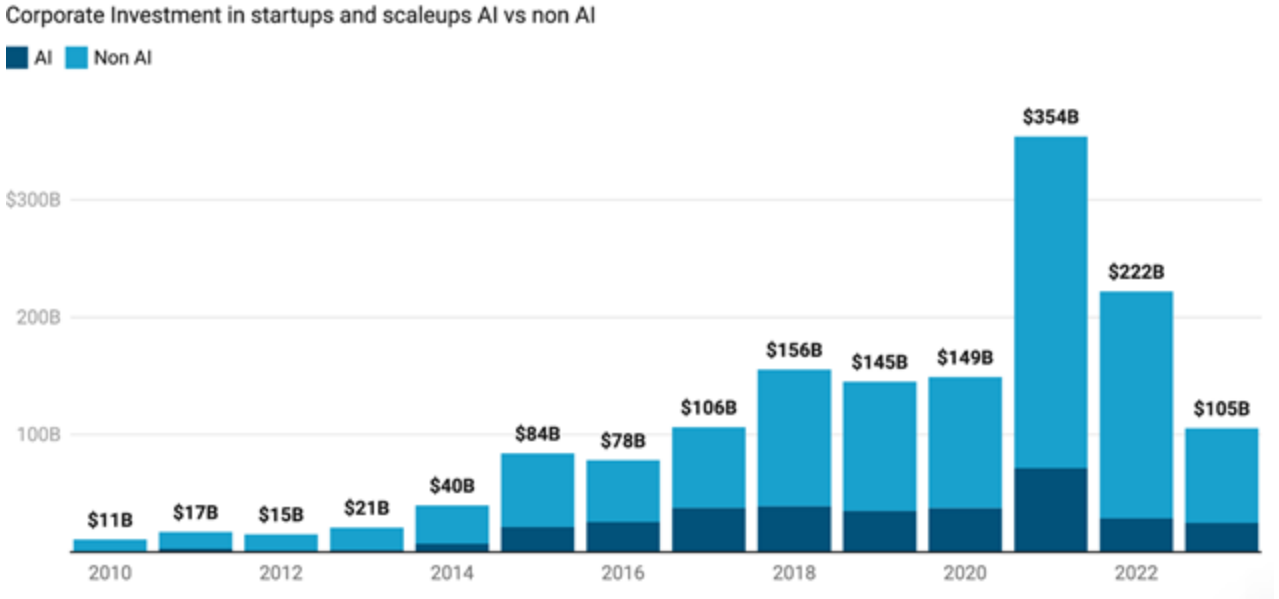
"The illusion of control: Wisdom and regulation of AI"

stateof.ai

{ the reports I study in detail }

24% of all corporate VC investments went into AI companies in 2023

- **In 2023, corporates refocused their investments towards GenAI. They cut investments into non-AI companies by 50% YoY while keeping AI investments roughly steady ($29B in '22 vs. $22B in '23).**



dealroom.co

Data as of 19 Sept 2023

South Africa’s challenge

24% of all VC funding to AI startups

- 10-15 in the graduate AI class that I teach
- comparatively, >1,000 at DTU

WHEN YOU TRAIN PREDICTIVE MODELS ON INPUT FROM YOUR USERS, IT CAN LEAK INFORMATION IN UNEXPECTED WAYS.