



Artificial Intelligence

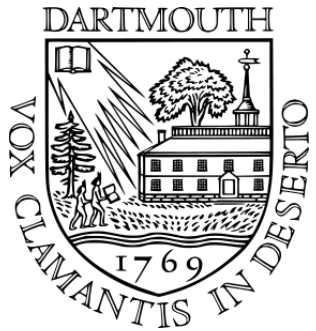
A historical perspective

Ann Nowé

Head of the AI Lab
Vrije Universiteit Brussel

The birth of the term AI

IJCAI 1956



"We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. 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**. 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."



John McCarthy



Marvin Minsky



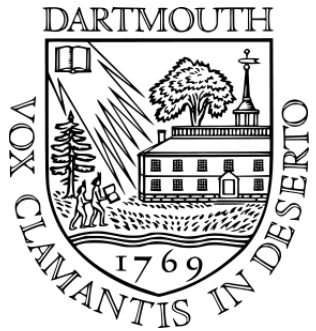
Nathaniel Rochester



Claude Shannon

The birth of the term AI

IJCAI 1956



"We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. 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. 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."



John McCarthy



Marvin Minsky

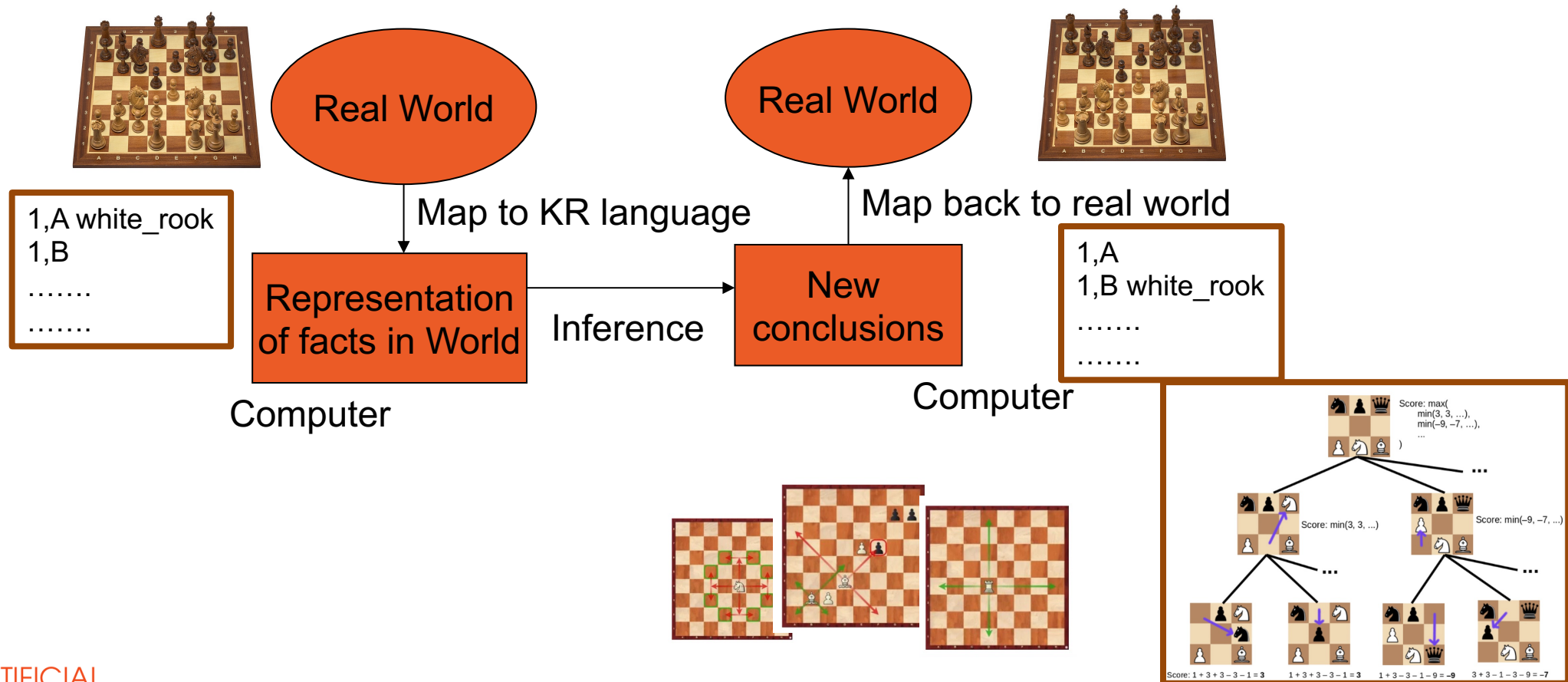


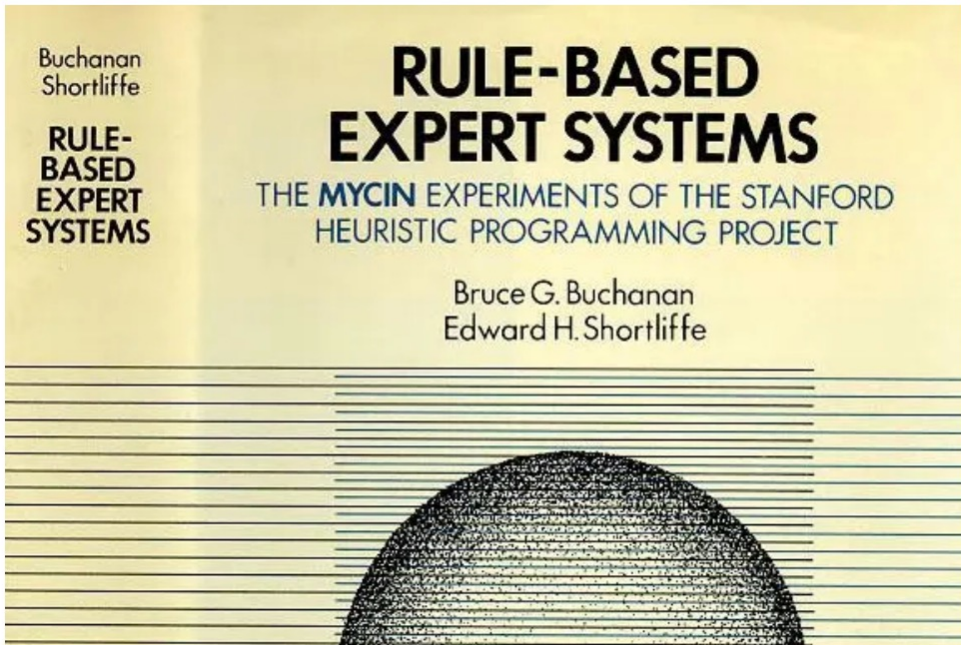
Nathaniel Rochester



Claude Shannon

The symbolic approach

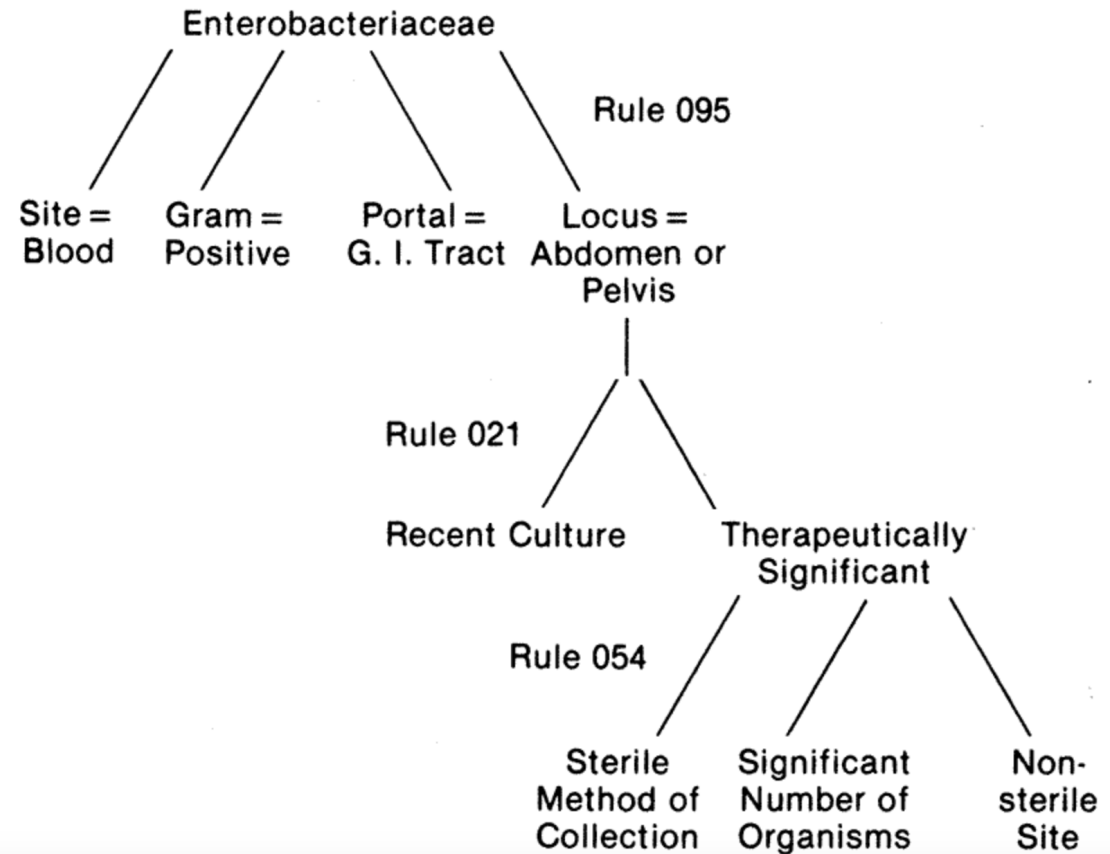


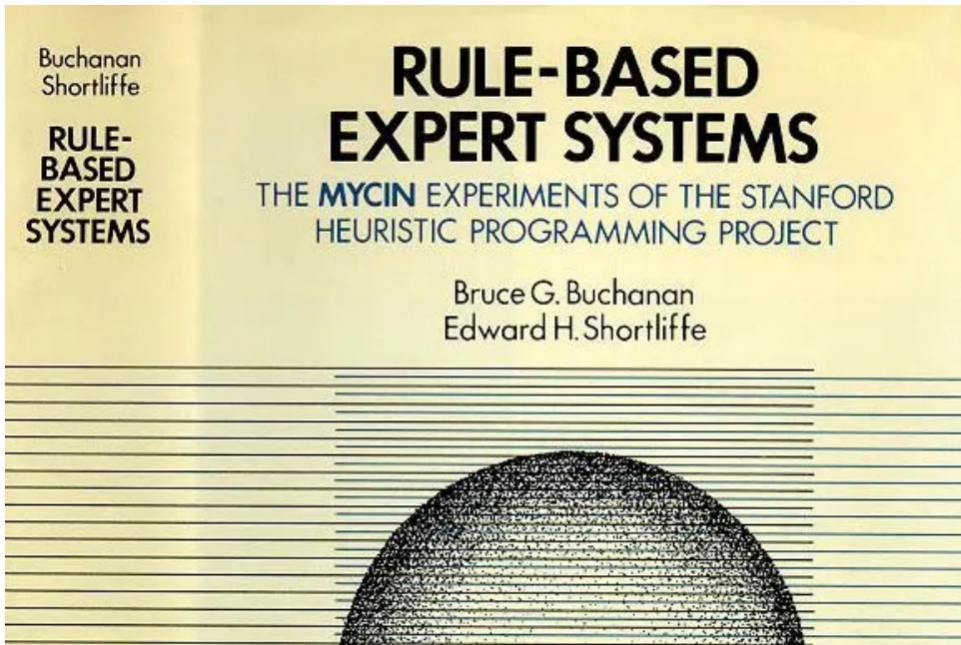


Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, 1984 BUCHANAN AND SHORTLIFFE

```
(defrule 52
  if (site culture is blood)
    (gram organism is neg)
    (morphology organism is rod)
    (burn patient is serious)
  then .4
    (identity organism is pseudomonas))
```

Rule 52:
 If
 1) THE SITE OF THE CULTURE IS BLOOD
 2) THE GRAM OF THE ORGANISM IS NEG
 3) THE MORPHOLOGY OF THE ORGANISM IS ROD
 4) THE BURN OF THE PATIENT IS SERIOUS
 Then there is weakly suggestive evidence (0.4) that
 1) THE IDENTITY OF THE ORGANISM IS PSEUDOMONAS





Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, 1984 BUCHANAN AND SHORTLIFFE

```
(defrule 52
  if (site culture is blood)
    (gram organism is neg)
    (morphology organism is rod)
    (burn patient is serious)
  then .4
    (identity organism is pseudomonas))
Rule 52:
If
  1) THE SITE OF THE CULTURE IS BLOOD
  2) THE GRAM OF THE ORGANISM IS NEG
  3) THE MORPHOLOGY OF THE ORGANISM IS ROD
  4) THE BURN OF THE PATIENT IS SERIOUS
Then there is weakly suggestive evidence (0.4) that
  1) THE IDENTITY OF THE ORGANISM IS PSEUDOMONAS
```

$$\text{OP1} \quad \int r \cdot f(x) \, dx \Rightarrow r \int f(x) \, dx$$

$$\text{OP2} \quad \text{Integration by parts:} \\ \int u \, dv \Rightarrow uv - \int v \, du \\ \text{(the precondition is internally represented as } \int f_1(x) f_2(x) \, dx, \text{ where } f_1(x) \text{ corresponds to } u \text{ and } f_2(x) \, dx \text{ corresponds to } dv \text{)}$$

$$\text{OP3} \quad 1 \cdot f(x) \Rightarrow f(x)$$

$$\text{OP4} \quad \int f_1(x) + f_2(x) \, dx \Rightarrow \int f_1(x) \, dx + \int f_2(x) \, dx$$

$$\text{OP5} \quad \int \sin(x) \, dx \Rightarrow -\cos(x) + C$$

$$\text{OP6} \quad \int \cos(x) \, dx \Rightarrow \sin(x) + C$$

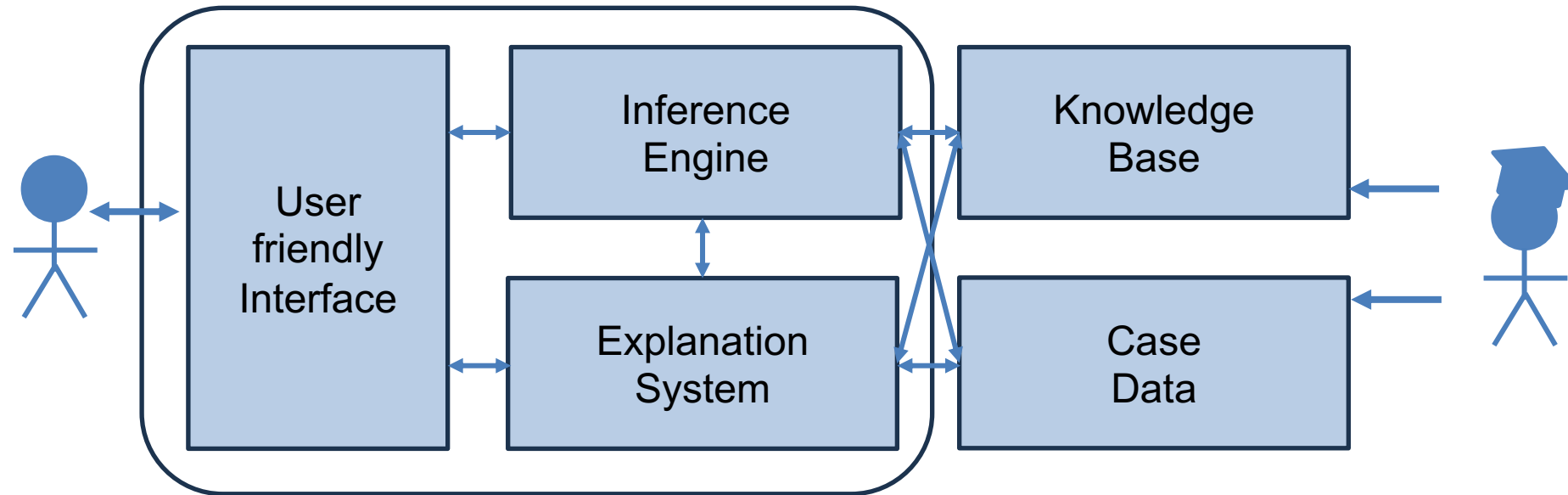
$$\text{OP7} \quad \int x^r \, dx \Rightarrow [x^{(r+1)}] / (r+1) + C$$

LEARNING BY EXPERIMENTATION: ACQUIRING AND REFINING PROBLEM-SOLVING HEURISTICS

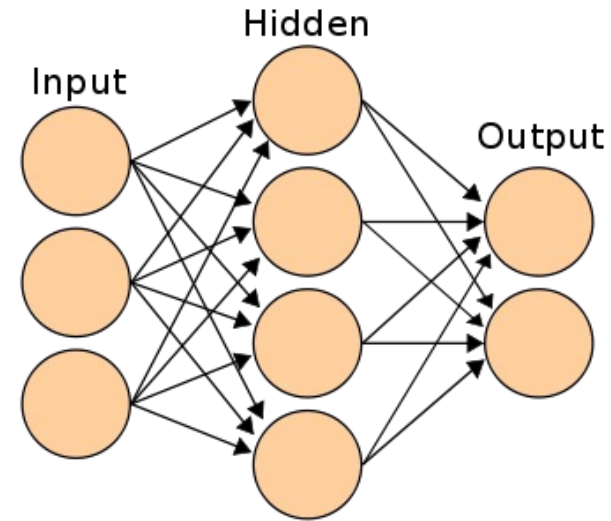
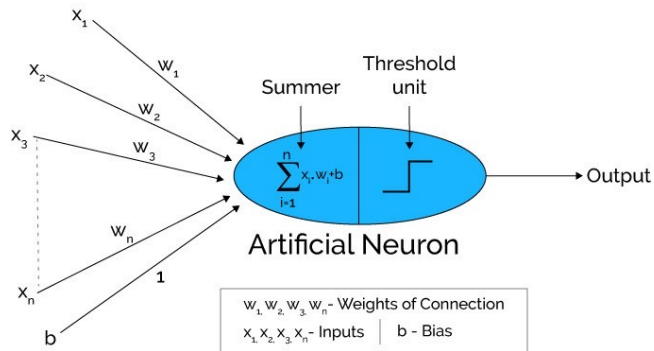
Tom M. Mitchell
Paul E. Utgoff
Rutgers University

Ranan Banerji
St. Joseph's University

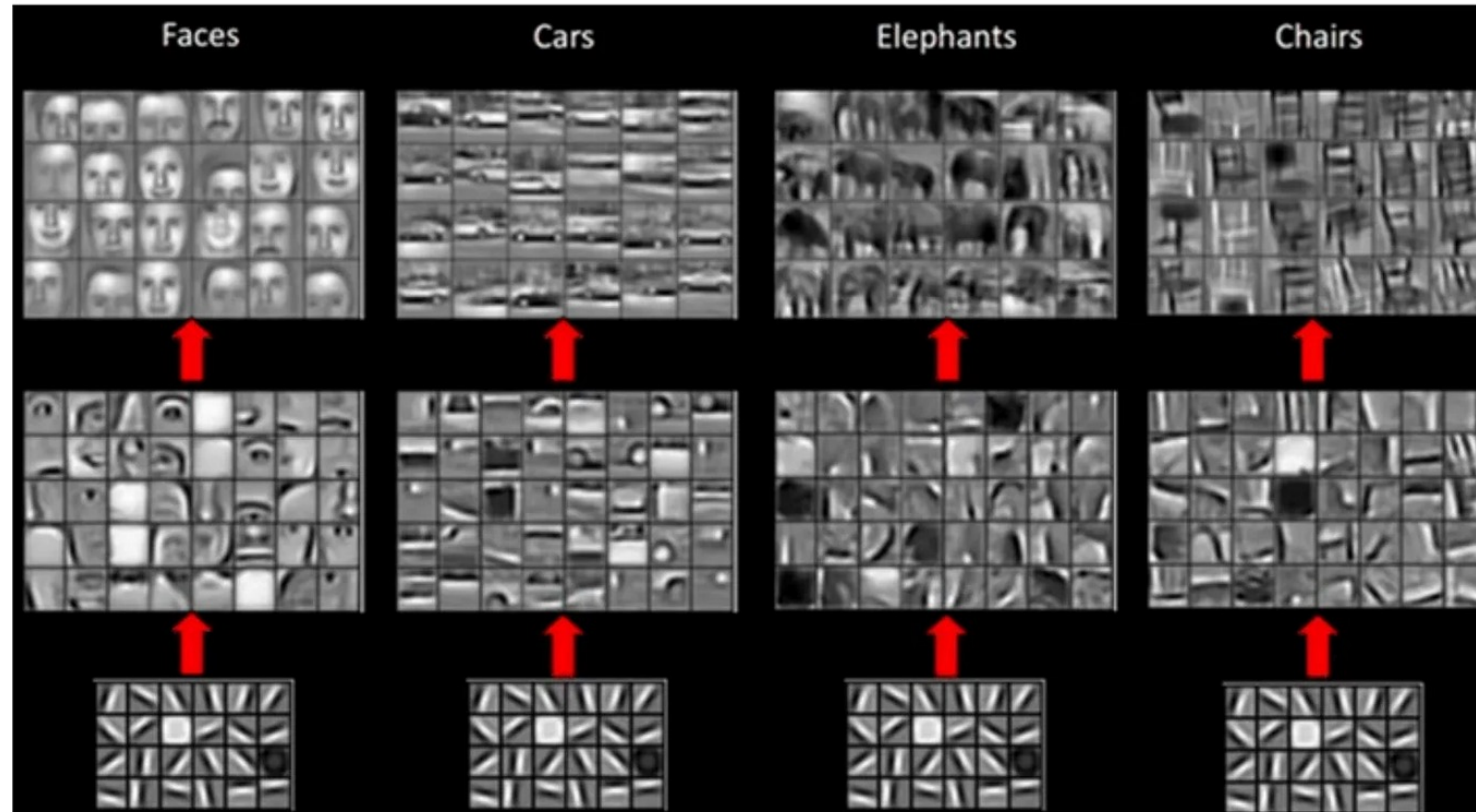
Expert systems



Sub-symbolic approach

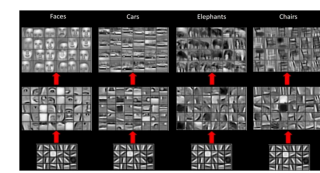
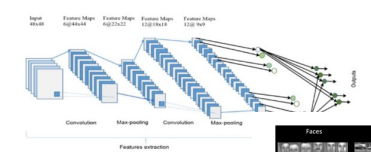
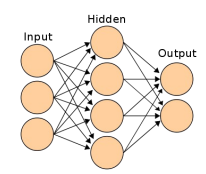
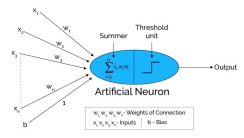
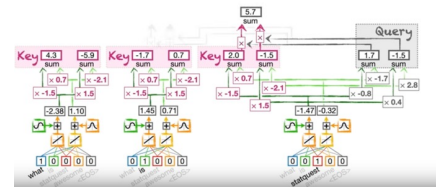
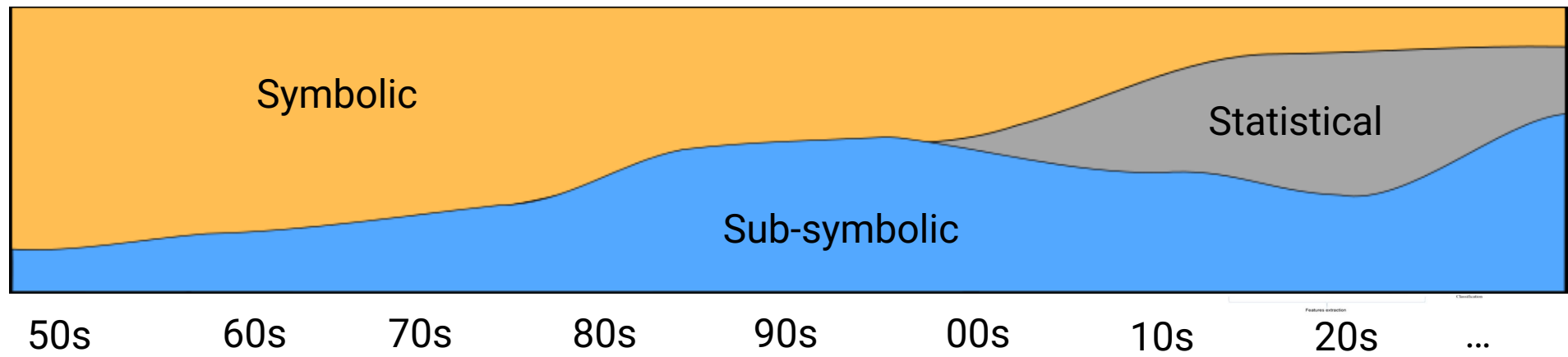
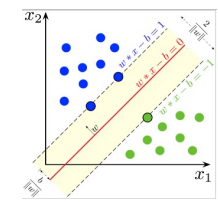
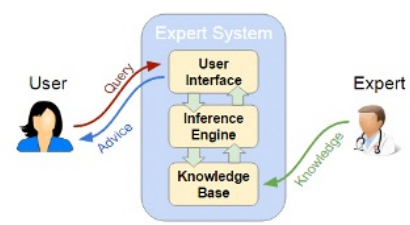


CNN Deep neural network



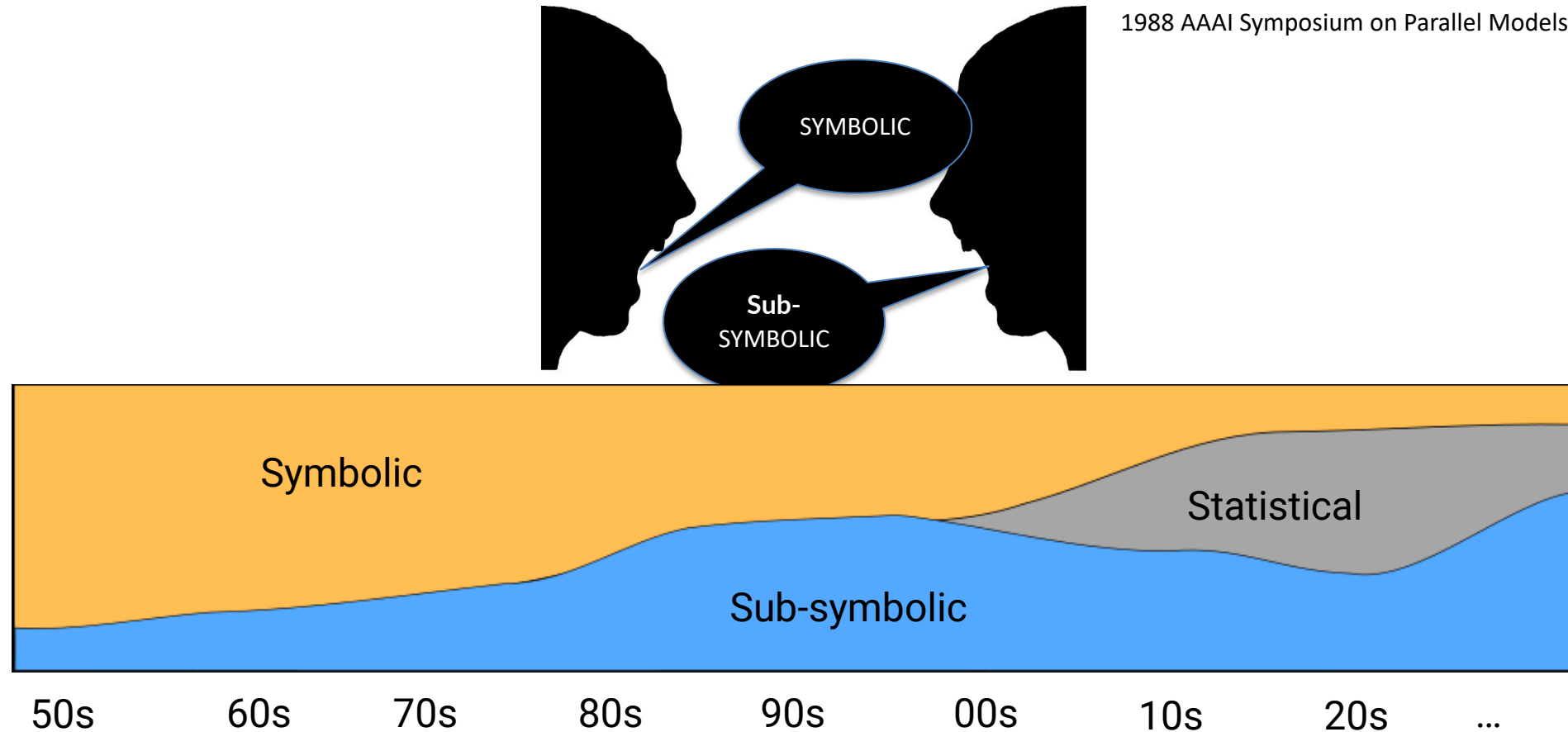
Examples of CNN's trained to recognize specific objects and their generated feature maps.

AI The different movements

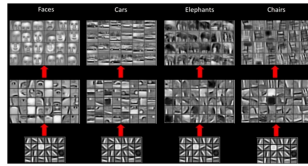
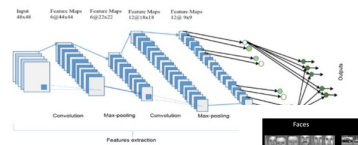
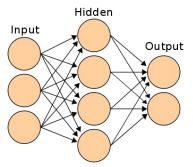
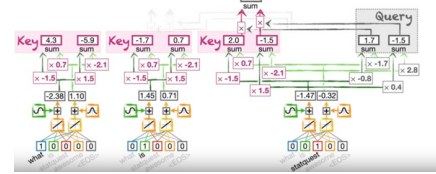
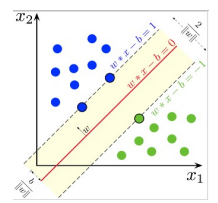
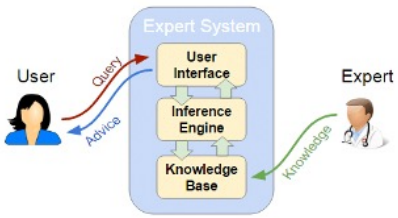
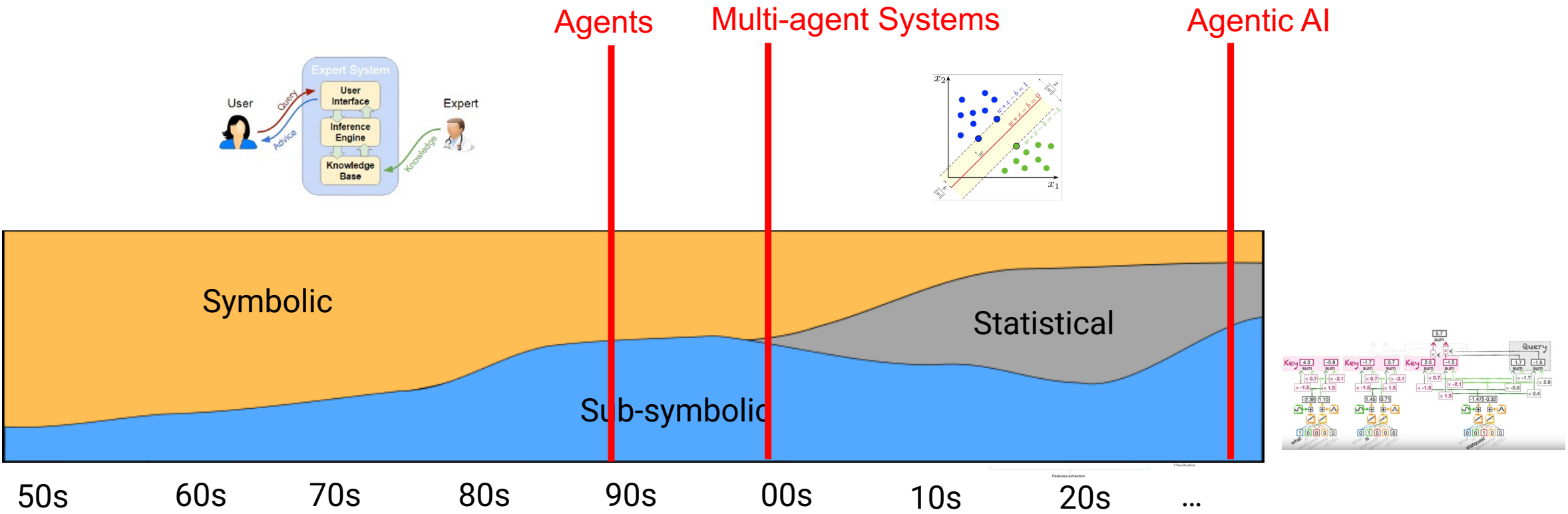


AI The different movements

1988 AAAI Symposium on Parallel Models of Intelligence



AI The different movements

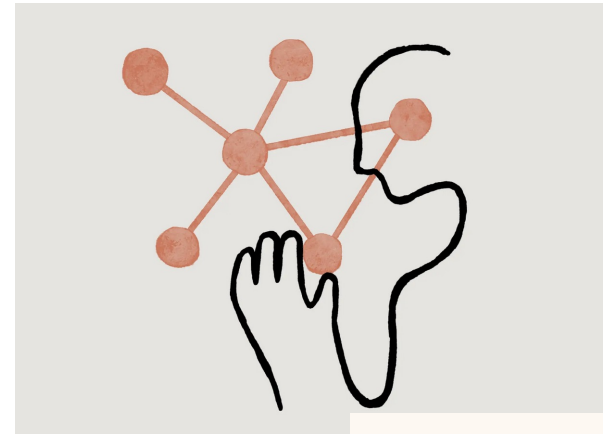


Large Language Models (LLMs)

Transformer



ChatGPT



Gemini



Copilot



Llama 3.3

Claude Shannon



1. Zero-order approximation (symbols independent and equiprobable).

XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZL-
HJQD.

A Mathematical Theory of Cryptography

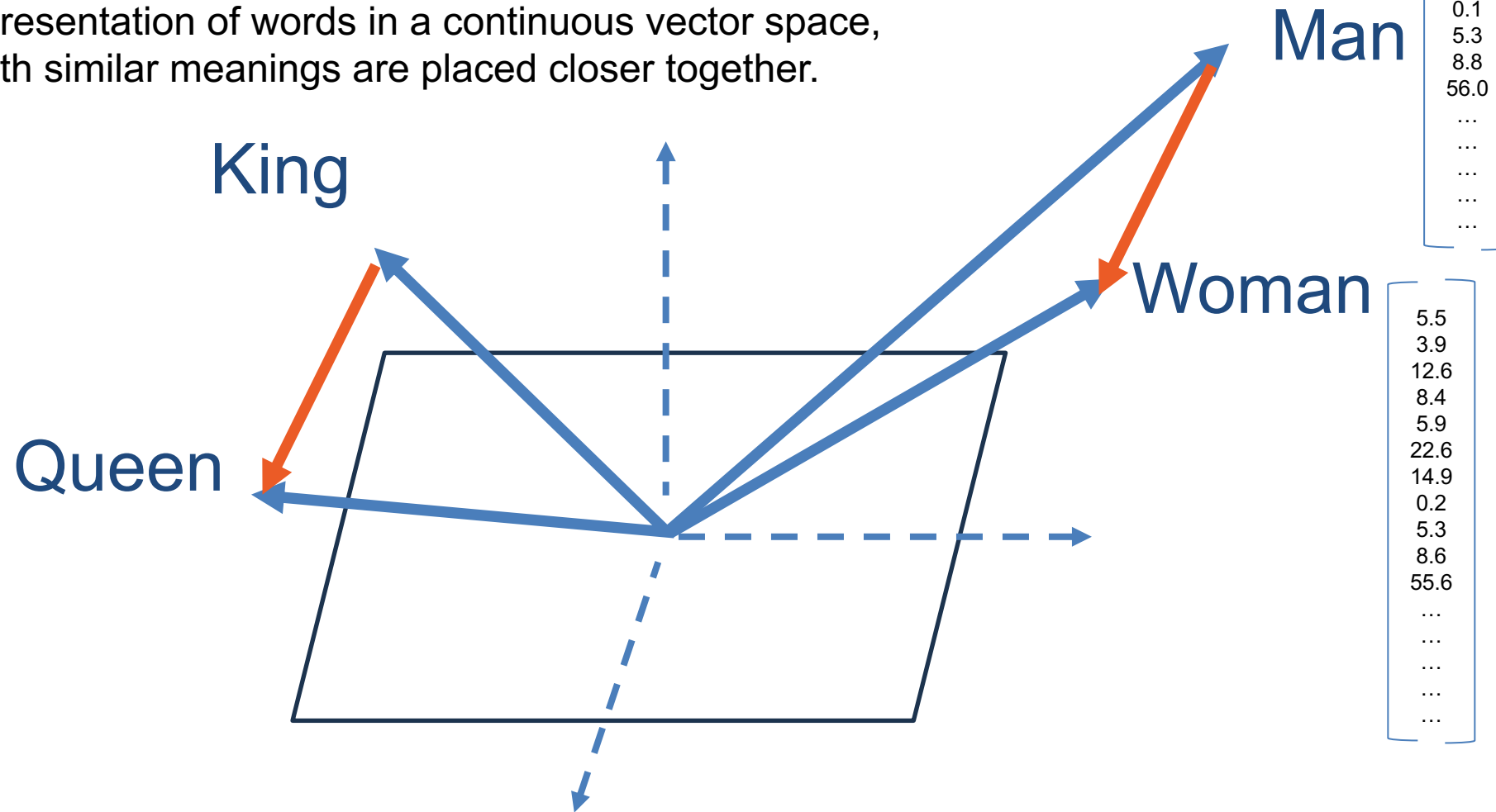
By C. E. Shannon

Reprinted with corrections from *System Technical Journal*,
Vol. 27, pp. 379–423, 623–656, July, October, 1948.

GPT3 # text equivalent to
reading 24 /7 2600 years

Word embeddings

a numerical representation of words in a continuous vector space, where words with similar meanings are placed closer together.



Attention



There is a **mole** in the garden

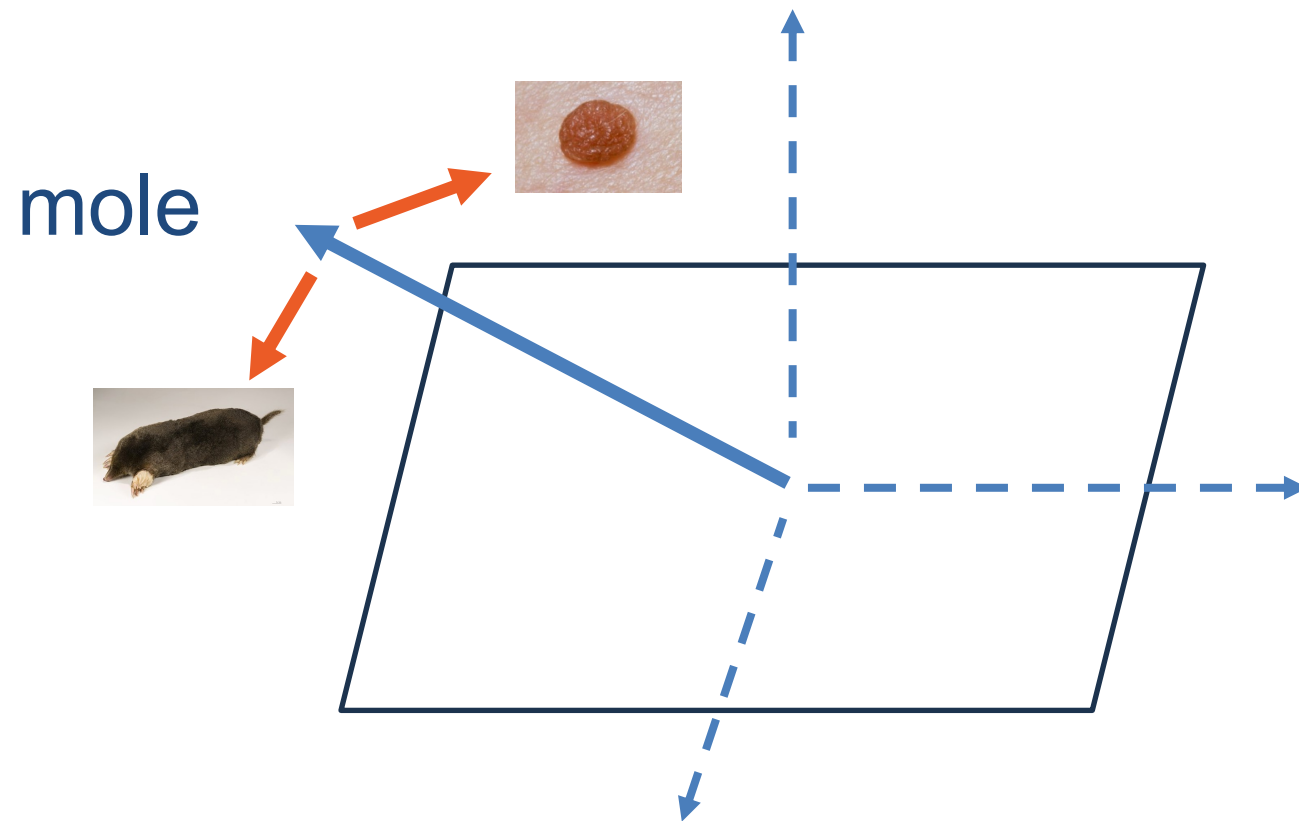
...	...	...	1.2	...	...	...
...	...	...	3.2	...	...	...
...	...	...	10.6	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...



The doctor took a biopsy of the **mole**

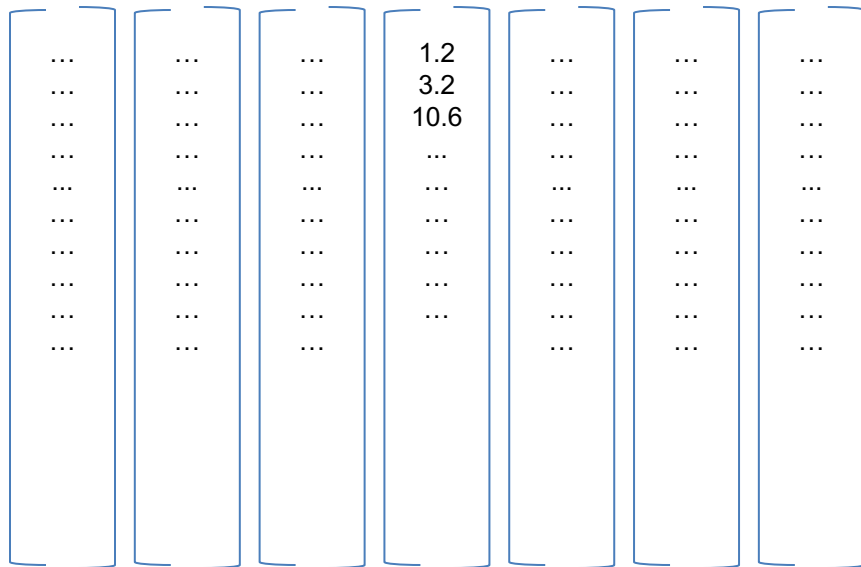
...	...	...	...	...	...	...	1.2
...	...	...	...	...	...	...	3.2
...	...	...	...	...	...	...	10.6
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

Attention

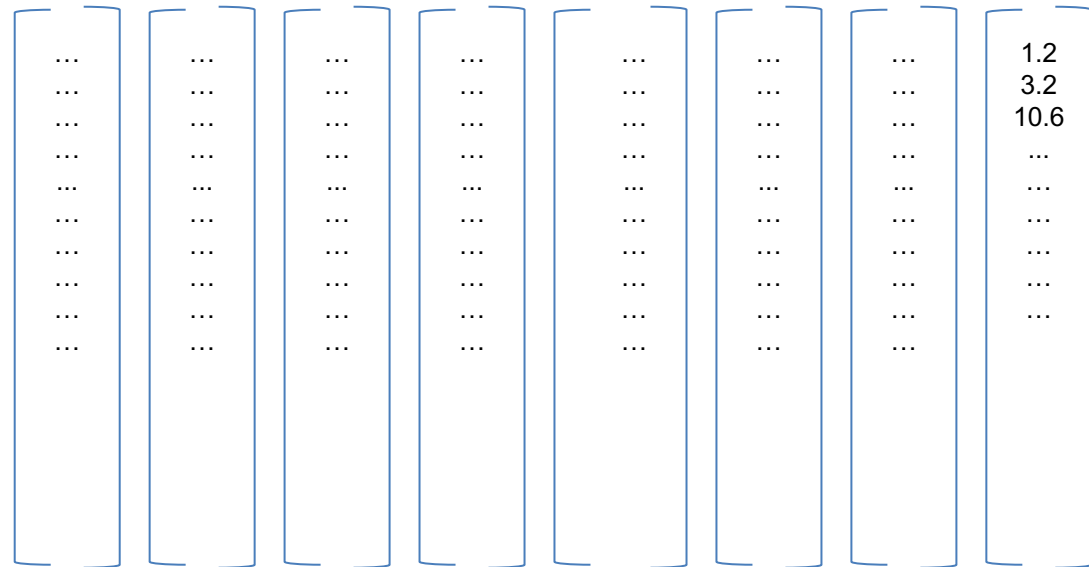


Attention

There is a **mole** in the garden

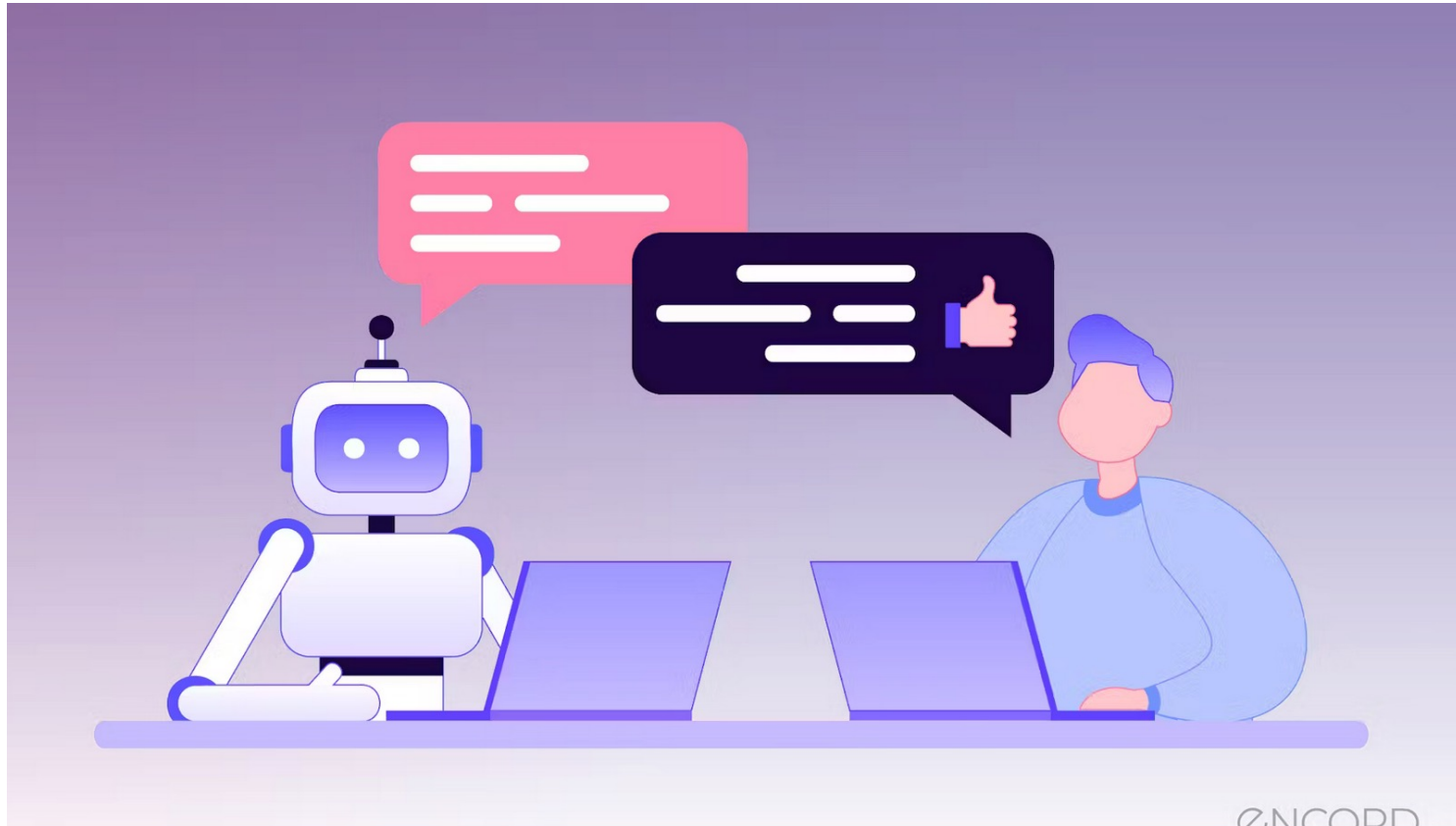


The doctor took a biopsy of the **mole**



Strenght of connections is trained
Pre-trained & domain specific

Reinforcement Learning from Human Feedback (RLHF)

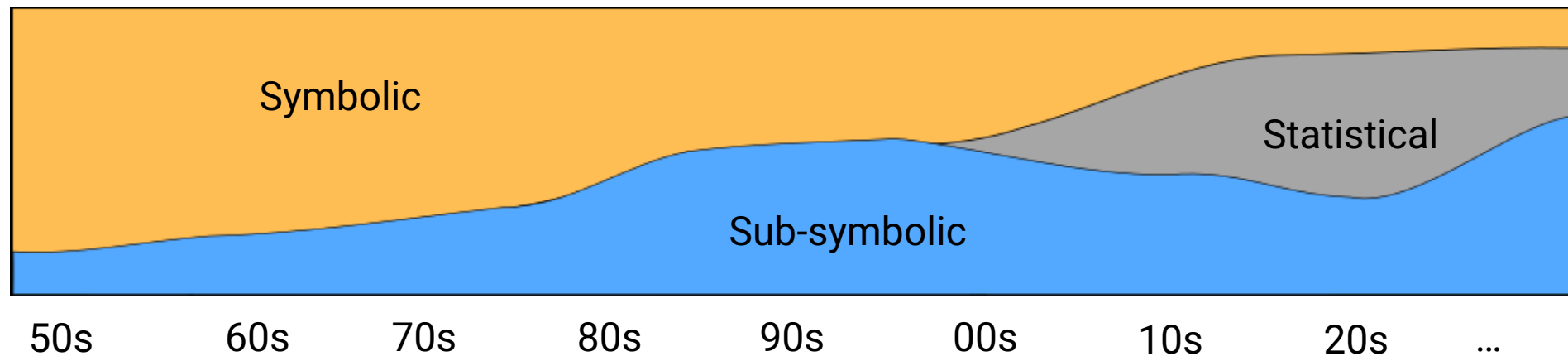


Taken from : <https://encord.com/blog/top-tools-rlhf/>

Human effort

Knowledge
intensive

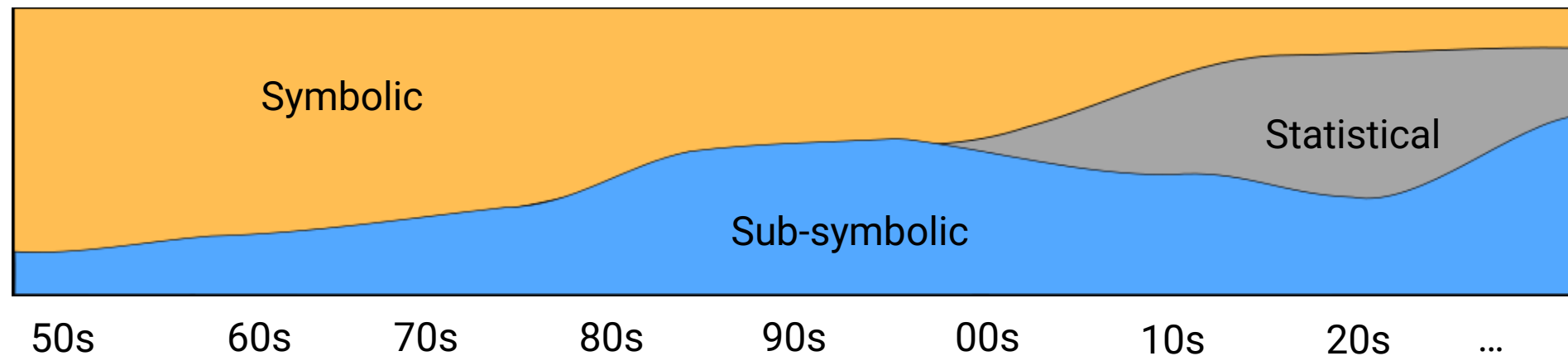
Data driven



Human AI interaction

Syntactic sugar

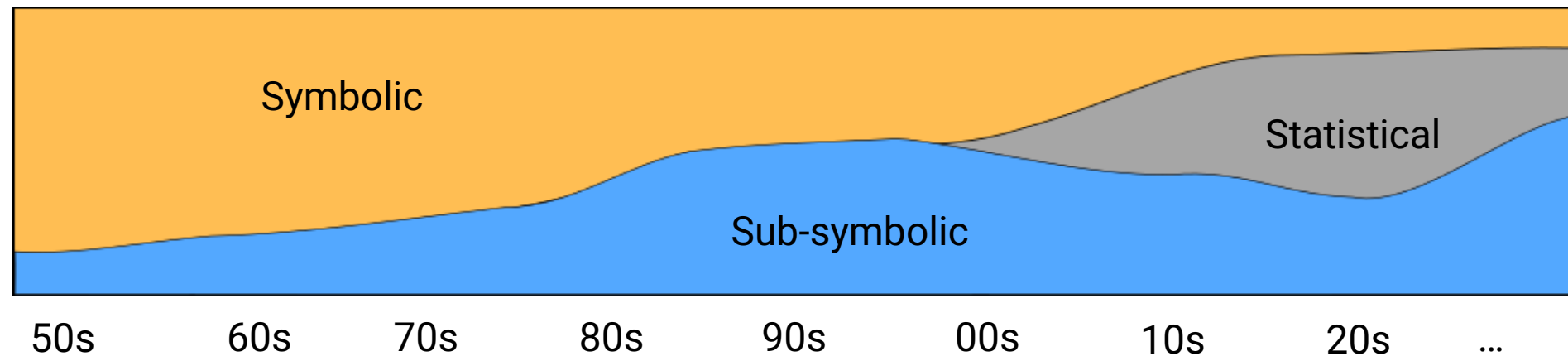
End-to-end LLM



Scope

Domain specific

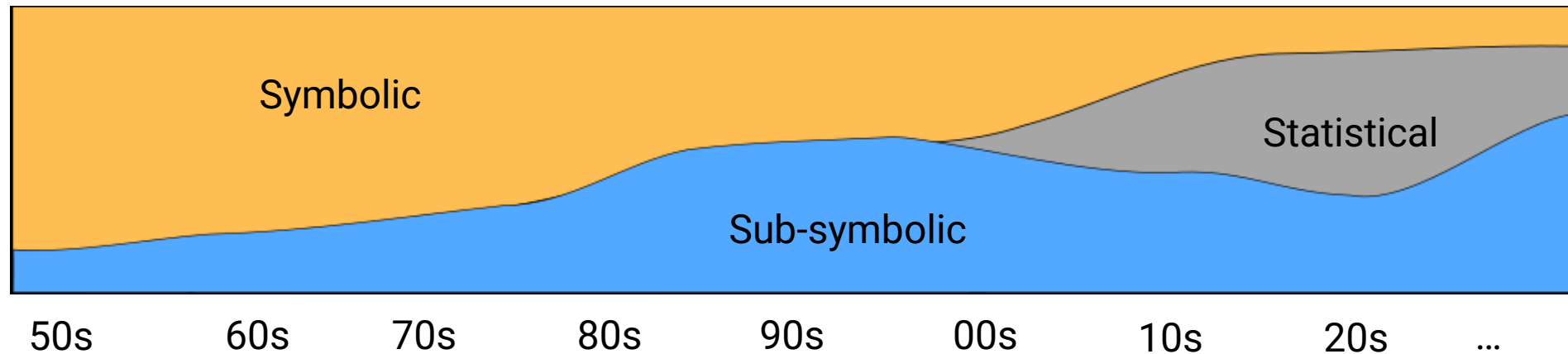
General purpose



Validation

Formal verification

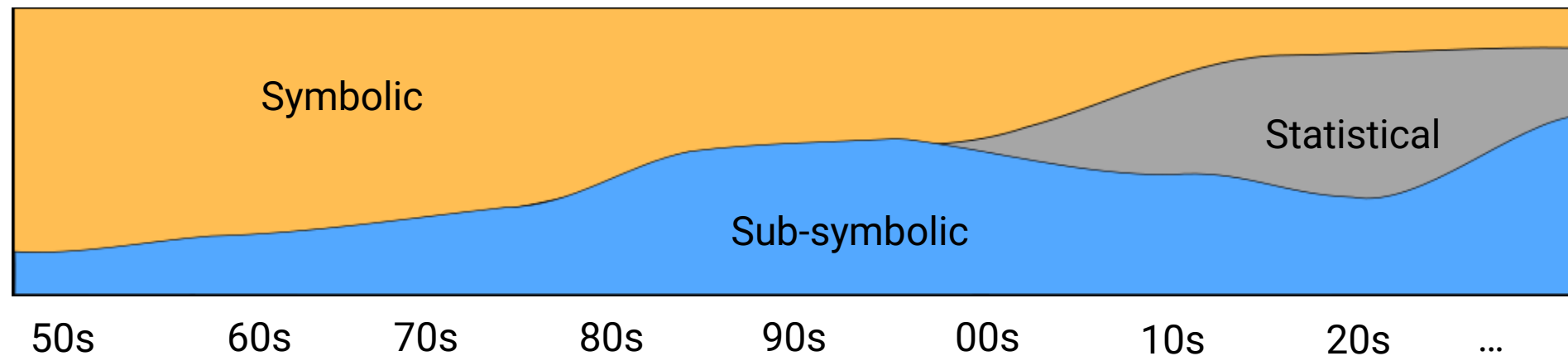
Experimental
validation



Energy consumption

Low , small models

High, large models



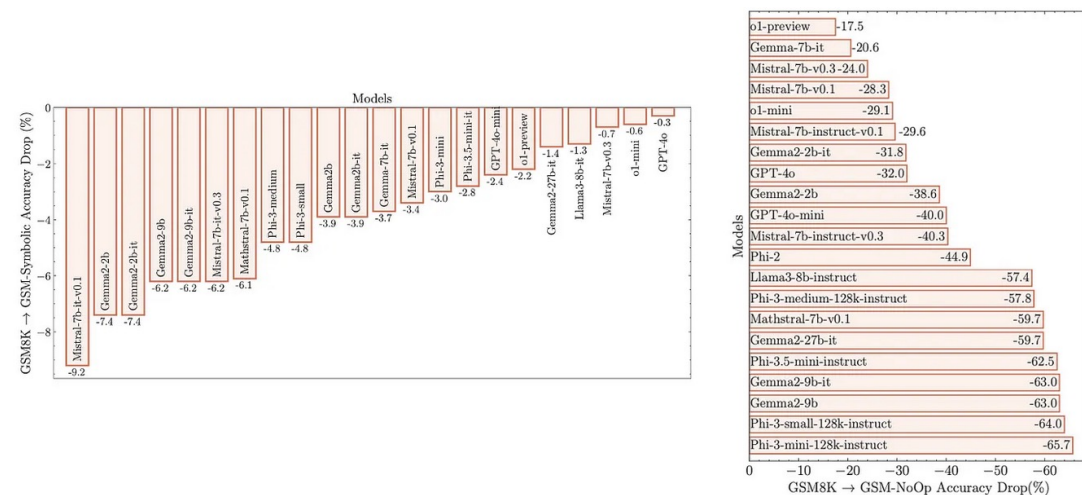
Can LLMs reason? No consensus

When Sophie watches her nephew, she gets out a variety of toys for him. The bag of building blocks has 31 blocks in it. The bin of stuffed animals has 8 stuffed animals inside. The tower of stacking rings has 9 multicolored rings on it. Sophie recently bought a tube of bouncy balls, bringing her total number of toys for her nephew up to 62. How many bouncy balls came in the tube?

Example of question from the GSM8K benchmark. The elements highlighted have been changed in the new benchmark GSM-Symbolic. From [1].

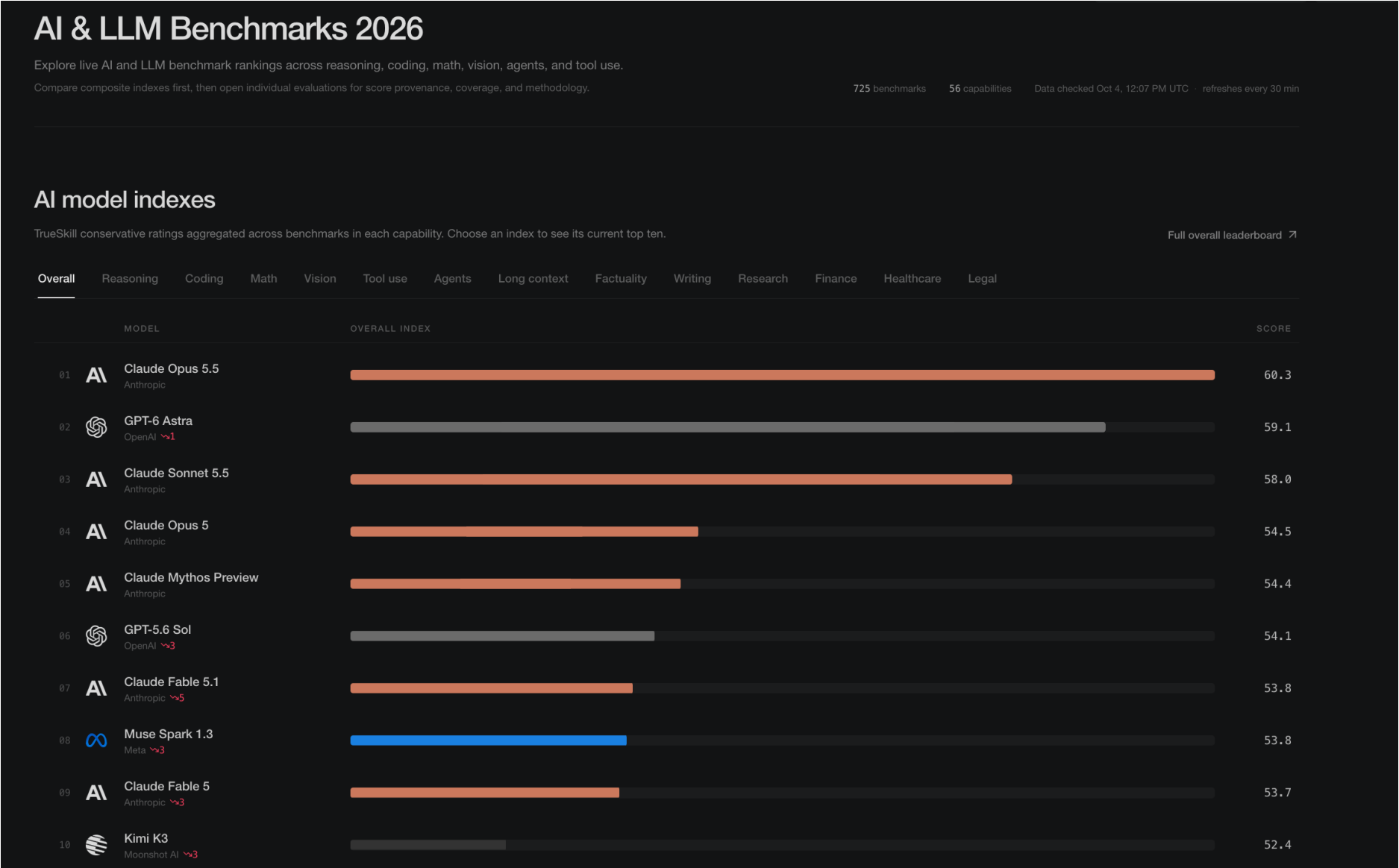
Oliver picks 44 kiwis on Friday. Then he picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, but five of them were a bit smaller than average. How many kiwis does Oliver have?

Example of question from the GSM8K benchmark. The element highlighted has been added in the new benchmark GSM-NoOp. From [1].



Accuracy decrease for (i) names and numbers changes (GSM-Symbolic) on the left, and (ii) adding unnecessary information (GSM-NoOp). From [1].

AI & LLM Benchmarks 2026

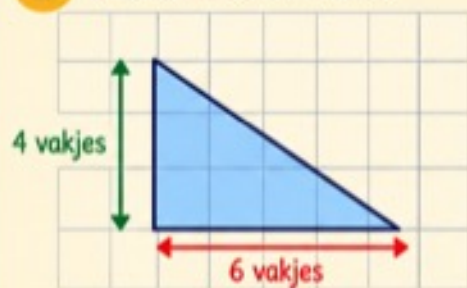


ChatGPT sept 2026

De oppervlakte van een driehoek

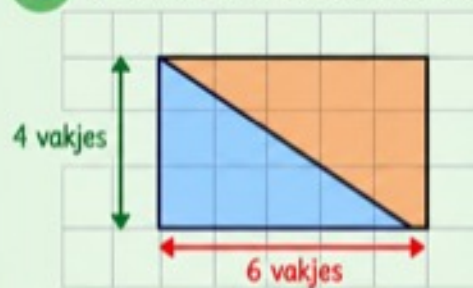
Een driehoek is de helft van een rechthoek met dezelfde basis en hoogte.

1 Teken een driehoek



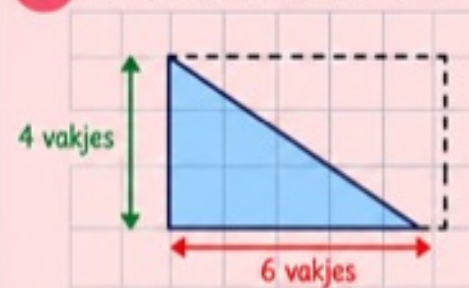
Deze driehoek heeft een basis van 6 vakjes en een hoogte van 4 vakjes.

2 Maak er een rechthoek van



Als je een gelijke driehoek ernaast legt, krijg je een rechthoek van $6 \times 4 = 24$ vierkantjes.

3 Eén driehoek is de helft



Onze driehoek is de helft van de rechthoek. Dus de oppervlakte is de helft van 24 vierkantjes, namelijk 12 vierkantjes.

4 De formule

$$\text{oppervlakte} = \frac{\text{basis} \times \text{hoogte}}{2}$$



De hoogte is de loodrechte afstand van de basis tot de overstaande hoek.

5 Voorbeeld



$$\begin{aligned}\text{Oppervlakte} &= \frac{6 \times 4}{2} \\ &= \frac{24}{2} \\ &= 12 \text{ cm}^2\end{aligned}$$

Een driehoek neemt de helft van de ruimte in!



De oppervlakte vertelt hoeveel vlakke ruimte een figuur inneemt. Bij een driehoek is dat de helft van $\text{basis} \times \text{hoogte}$.

New York City Schools Ban
ChatGPT Amid Cheating Worries

-CNET

*Alarmed by A.I. Chatbots, Universities
Start Revamping How They Teach*

-New York Times

Professor catches student cheating
with ChatGPT: 'I feel abject terror'

-New York Post

ChatGPT and the Death of Education

0:18 / 15:36 • Intro >

How AI Could Save (Not Destroy) Education | Sal Khan | TED



TED ✓

24.3M subscribers

Subscribe

1.4M views 1 year ago #TED #TEDTalks #khanacademy





Applications can make AI models better

GPT-4 clearly has the capability, but the context you give it matters a lot!

We have:

- Used "AI Thoughts"
- Give it the context of the problem (including human generated hints)
- Spent 6 months prompt engineering for tutoring with an emphasis on math
- Spent a lot of time helping to fine-tune the model for this use case

<internal khanmigo thoughts>

Original equation: $6(-2g-1) = -(13g+2)$

Step 1: Distribute the 6 to both terms inside the parentheses on the left side:

→ $-12g - 6 = -(13g + 2)$

Step 2: Distribute the negative sign to both terms inside the parentheses on the right side:

→ $-12g - 6 = -13g - 2$

Now the student's equation should look like this:

$-12g - 6 = -13g - 2$

The student got a different answer than I did, but do not tell them they made a mistake.

Instead, ask them to explain how they got to that step.

</internal khanmigo thoughts>

there was a lot of work
behind the scenes to make that happen.



12:13 / 15:36 • Magic >



HD



Alpha Geometry

AlphaGeometry: An Olympiad-level AI system for geometry

RESEARCH

17 JANUARY 2024

Trieu Trinh and Thang Luong

[Share](#)



AlphaGeometry

[Article](#) [Talk](#)

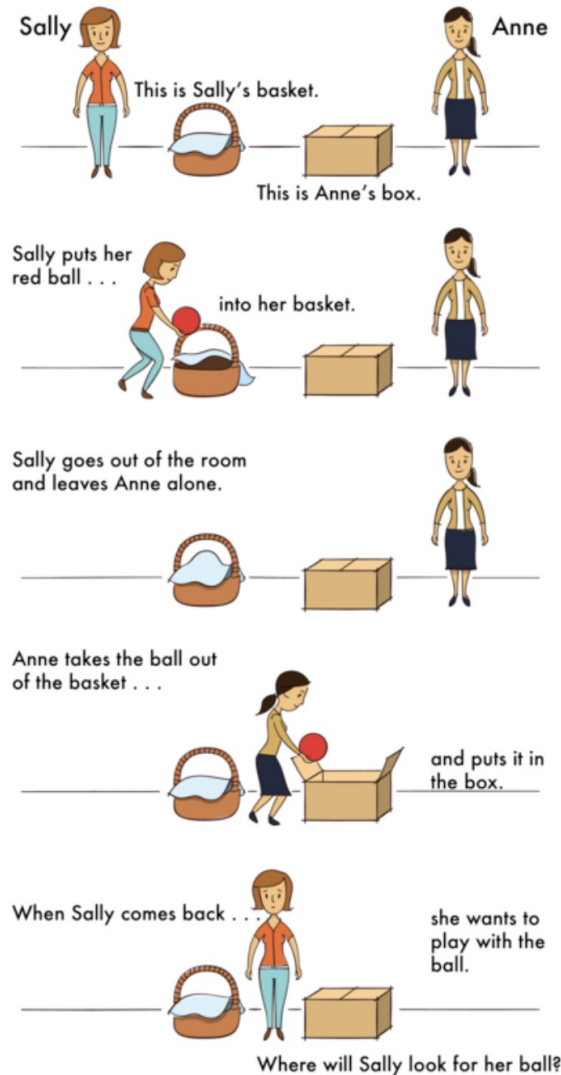
[Read](#) [Edit](#) [View history](#) [Tools](#)

[2 languages](#)

From Wikipedia, the free encyclopedia

AlphaGeometry is an [artificial intelligence](#) (AI) program that can solve hard problems in [Euclidean geometry](#). The system comprises [a data-driven large language model \(LLM\)](#) and [a rule-based symbolic engine](#) (*Deductive Database Arithmetic Reasoning*). It was developed by [DeepMind](#), a subsidiary of [Google](#). The program solved 25 geometry problems out of 30 from the [International Mathematical Olympiad](#) (IMO) under competition time limits—a performance almost as good as the average human gold medallist. For comparison, the previous AI program, called [Wu's method](#), managed to solve only 10 problems.^{[1][2]}

LLMs and Theory of Mind



AI : the good, the bad and the ugly

Emerging COVID-19 variants identified with AI

Health | 12th March 2024



© shutterstock/PopTika

Scientists at the Universities of Manchester and Oxford have developed an AI framework that can track new COVID-19 variants.

Nature Public Health Emergency Collection

Public Health Emergency COVID-19 Initiative

[Ethics Inf Technol.](#) 2021 Feb 9 : 1–7.

doi: [10.1007/s10676-020-09567-7](https://doi.org/10.1007/s10676-020-09567-7) [Epub ahead of print]

PMCID: PMC7871022

PMID: [33584129](https://pubmed.ncbi.nlm.nih.gov/33584129/)

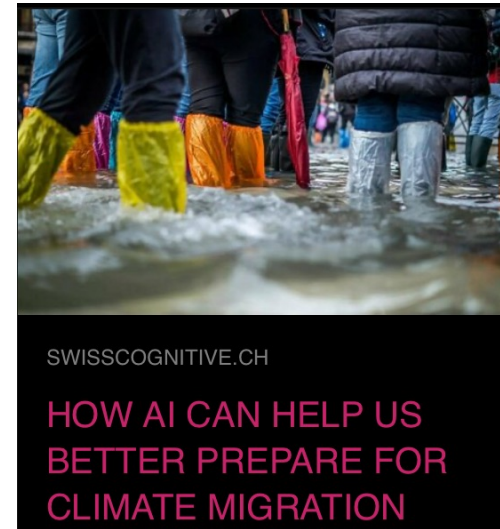
The CLAIRE COVID-19 initiative: approach, experiences and recommendations

[Gianluca Bontempi](#),¹ [Ricardo Chavarriaga](#),² [Hans eD Canck](#),³ [Emanuela Girardi](#),⁴ [Holger H larla Kilbane-Dawe](#),⁵ [Tonio Ball](#),⁷ [Ann Nowé](#),⁸ [Jose Sousa](#),⁹ [Davide Bacciu](#),¹⁰ [Marco Aldin Manlio eD Domenico](#),¹² [Alessandro Saffiotti](#),¹³ and [Marco Maratea](#)¹⁴

DataNews Nieuws Analyse Opinies Reviews Carrière Events Dit waren de Data News A



Het menselijke microbloom. © Getty Images



DataNews

FR

Nieuws Analyse Opinies Reviews Carrière Events Generative AI ICT Guide

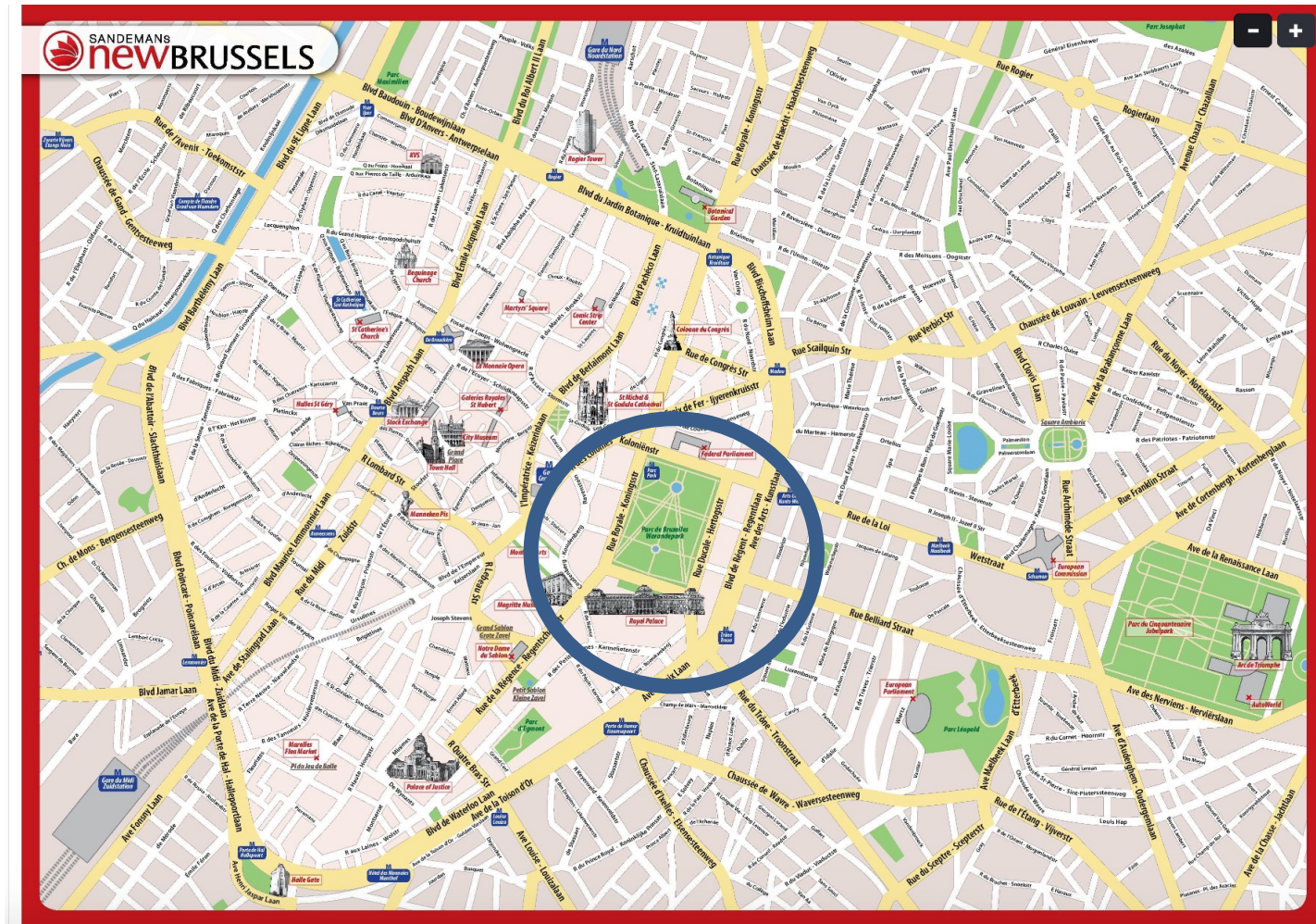
1



AI-tools kunnen vandaag al met 96 procent nauwkeurigheid de signalen van een exoplaneet herkennen.

<https://datanews.knack.be/nieuws/innovatie/ruimtevaart/kennis-over-ons-heelal-groeit-sneller-dan-ooit-met-dank-aan-ai>

The bad



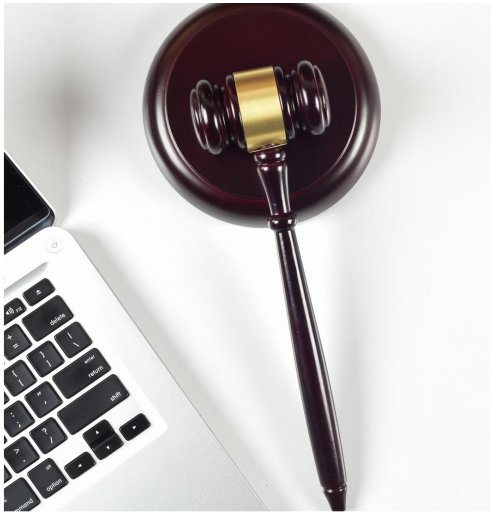
The ugly

The Washington Post
Democracy Dies in Darkness

MONKEY CAGE

A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear.

By Sam Corbett-Davies, Emma Pierson, Avi Feller and Shari
October 17, 2016 at 5:00 a.m. EDT



Is the algoritme fair?

*All models are wrong
but some are useful*



George E.P. Box

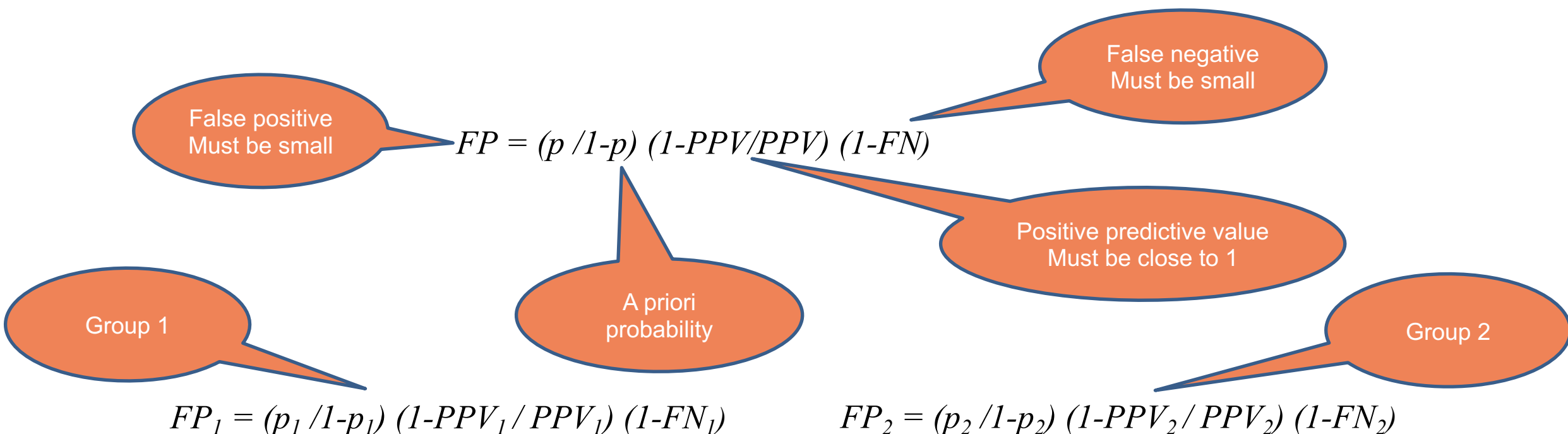
Table 15: Classification of fairness notions. (* notion newly defined in this paper)

Fairness Notion	Ref.	Formulation	Classification	Type
Statistical parity	[29]	$P(\hat{Y} \mid A = 0) = P(\hat{Y} \mid A = 1)$	Independence (equivalent or relaxed*)	Group
Conditional statistical parity	[21]	$P(\hat{Y} = 1 \mid E = e, A = 0) = P(\hat{Y} = 1 \mid E = e, A = 1)^*$		
Equalized odds	[44]	$P(\hat{Y} = 1 \mid Y = y, A = 0) = P(\hat{Y} = 1 \mid Y = y, A = 1) \quad \forall y \in \{0, 1\}$	Separation (equivalent or relaxed*)	
Equal opportunity		$P(\hat{Y} = 1 \mid Y = 1, A = 0) = P(\hat{Y} = 1 \mid Y = 1, A = 1)^*$		
Predictive equality	[21]	$P(\hat{Y} = 1 \mid Y = 0, A = 0) = P(\hat{Y} = 1 \mid Y = 0, A = 1)^*$		
Balance for positive class	[58]	$E[S \mid Y = 1, A = 0] = E[S \mid Y = 1, A = 1]^*$		
Balance for negative class		$E[S \mid Y = 0, A = 0] = E[S \mid Y = 0, A = 1]^*$		
Overall balance	*	$E[S \mid Y = y, A = 0] = E[S \mid Y = y, A = 1] \quad \forall y \in \{0, 1\}$		
Conditional use acc. equality	[10]	$P(Y = y \mid \hat{Y} = y, A = 0) = P(Y = y \mid \hat{Y} = y, A = 1) \quad \forall y \in \{0, 1\}$	Sufficiency (equivalent or relaxed*)	
Predictive parity	[18]	$P(Y = 1 \mid \hat{Y} = 1, A = 0) = P(Y = 1 \mid \hat{Y} = 1, A = 1)^*$		
Negative predictive parity	*	$P(Y = 1 \mid \hat{Y} = 0, A = 0) = P(Y = 1 \mid \hat{Y} = 0, A = 1)^*$		
Calibration	[18]	$P(Y = 1 \mid S = s, A = 0) = P(Y = 1 \mid S = s, A = 1) \quad \forall s \in [0, 1]$		
Well-calibration	[58]	$P(Y = 1 \mid S = s, A = 0) = P(Y = 1 \mid S = s, A = 1) = s \quad \forall s \in [0, 1]$		
Overall accuracy equality	[10]	$P(\hat{Y} = Y \mid A = 0) = P(\hat{Y} = Y \mid A = 1)$	Other metrics from confusion matrix	
Treatment equality		$\frac{FN}{FP}(A=0) = \frac{FN}{FP}(A=1)$	Independence, Separation and Sufficiency	
Total fairness		—		
Total effect	[76]	$TE_{a_1, a_0}(\hat{y}) = P(\hat{y}_{A \leftarrow a_1}) - P(\hat{y}_{A \leftarrow a_0})$	Causality	
Effect of treatment on treated		$ETT_{a_1, a_0}(\hat{y}) = P(\hat{y}_{A \leftarrow a_1} \mid a_0) - P(\hat{y} \mid a_0)$		
No unresolved discrimination	[53]	—		
No proxy discrimination		$P(\hat{Y} \mid do(P_x = p)) = P(\hat{Y} \mid do(P_x = p')) \quad \forall P_x \text{ and } \forall p, p'$		
Counterfactual fairness	[59]	$P(\hat{Y}_{A \leftarrow a}(U) = y \mid X = x, A = a) = P(\hat{Y}_{A \leftarrow a'}(U) = y \mid X = x, A = a)$	Similarity Metric	
Causal discrimination	[38]	$X_{(A=0)} = X_{(A=1)} \wedge A_{(A=0)} \neq A_{(A=1)} \Rightarrow \hat{y}_{(A=0)} = \hat{y}_{(A=1)}$		
Fairness through awareness	[29]	$D(M(v_i), M(v_j)) < d(v_i, v_j)$		

27

Karima Makhlouf, Sami Zhioua, and Catuscia Palamidessi. 2021.
On the Applicability of Machine Learning Fairness Notions.
SIGKDD Explor. Newsl. 23, 1 (June 2021), 14–23.
<https://doi.org/10.1145/3468507.3468511>

Can an algorithm (model) be fully fair?



Model is fair when

$$FP_1 = FP_2$$

$$FN_1 = FN_2$$

$$PPV_1 = PPV_2$$

Implies

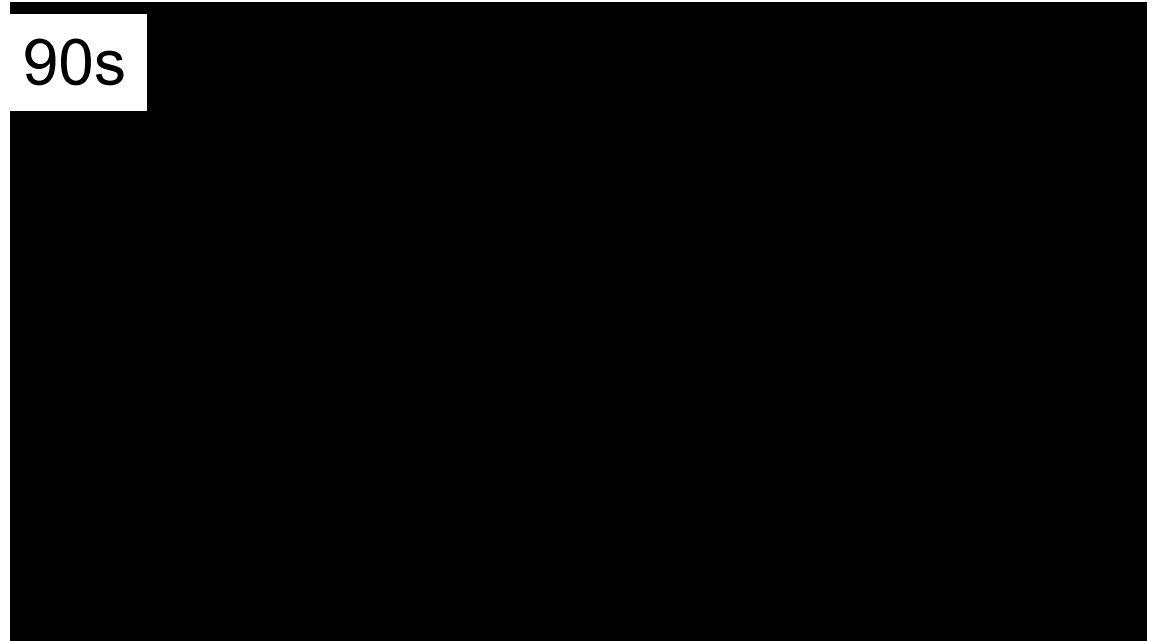
$$p_1 = p_2$$

Self driving cars

80s



90s



AI experience center

