

# Language, Language Models, and ChatGPT

---

Leipzig · July 6, 2023

- ① Demo and Motivation
- ② A Short History of Language Models
- ③ More on ChatGPT



*“While the functionality of ChatGPT is not brand new, the public interface—including layout, templating for code and related outputs, and general user experience—is new and innovative.”*

[Alan D. Thompson, Feb. 2023]

*“You shall know a word by the company it keeps.”*

[John Rupert Firth, 1957]

*“You shall **know a word** by the company it keeps.”*

[John Rupert Firth, 1957]

We **interpret words** (give them meaning) through their context.

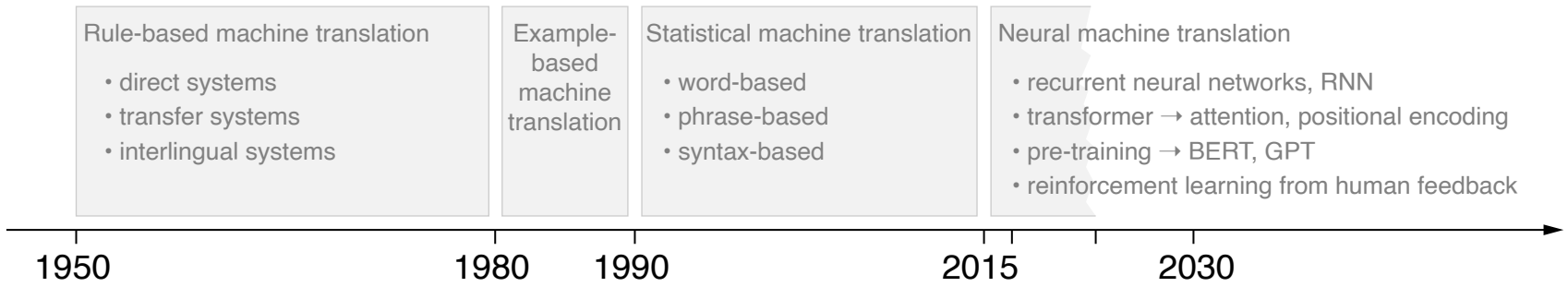
Illustration:

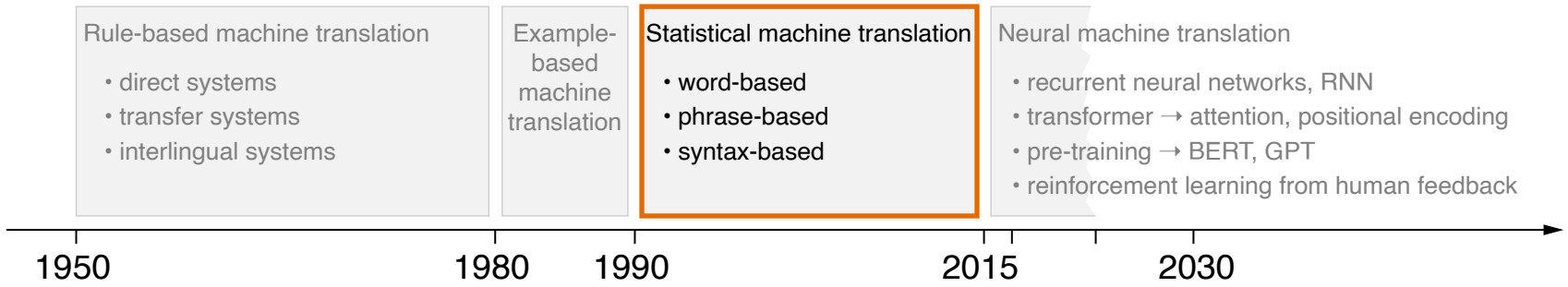
(a) I saw a **jaguar** in the zoo.

(b) The **jaguar** won the formula 1 race.



## A Short History of Language Models





A statistical language model  
is a probability distribution over all possible sentences.

### Rule-based machine translation

- direct systems
- transfer systems
- interlingual systems

### Example-based machine translation

### Statistical machine translation

- word-based
- phrase-based
- syntax-based

### Neural machine translation

- recurrent neural networks, RNN
- transformer → attention, positional encoding
- pre-training → BERT, GPT
- reinforcement learning from human feedback

1950

1980

1990

2015

2030

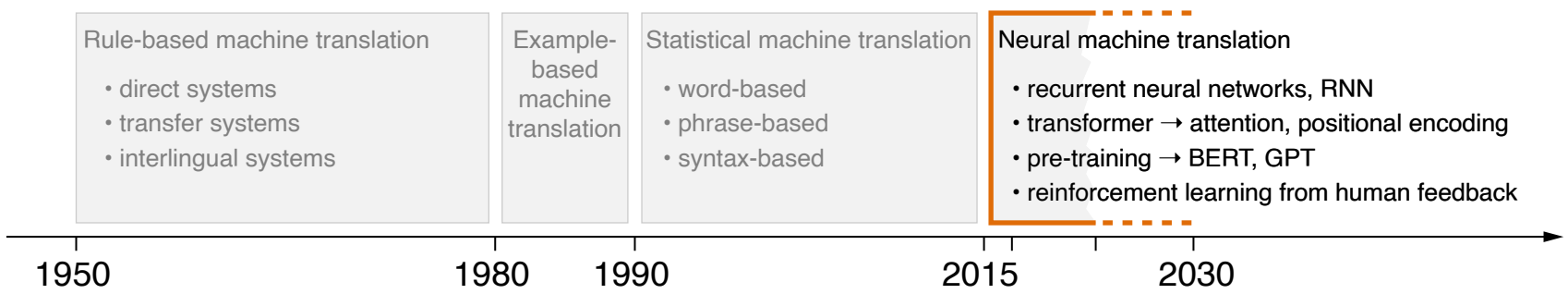
A statistical language model  
is a probability distribution over all possible sentences.

### Illustration:

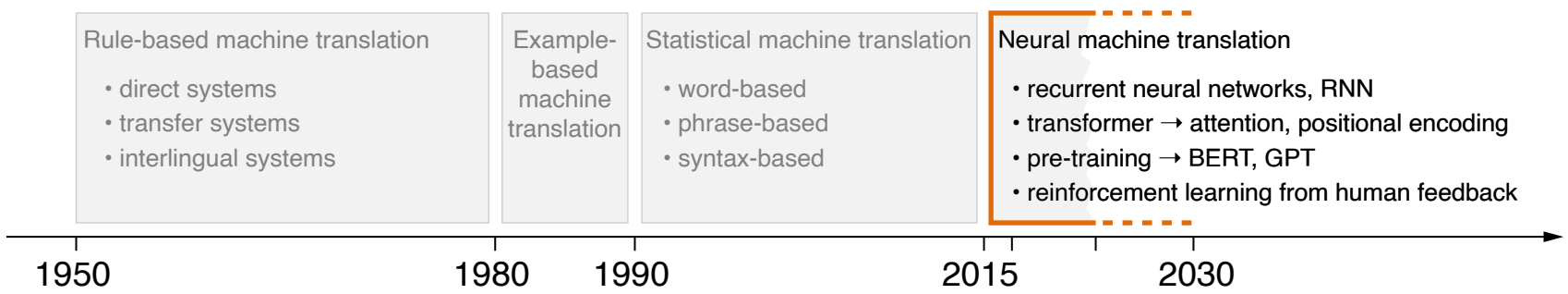
(1) i love my ? N N

(2) see ... works N N



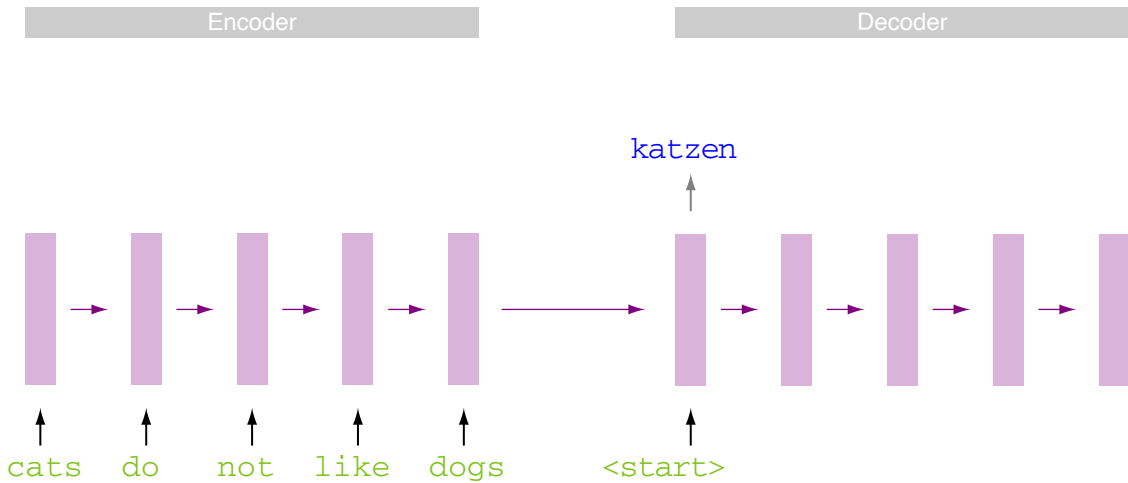


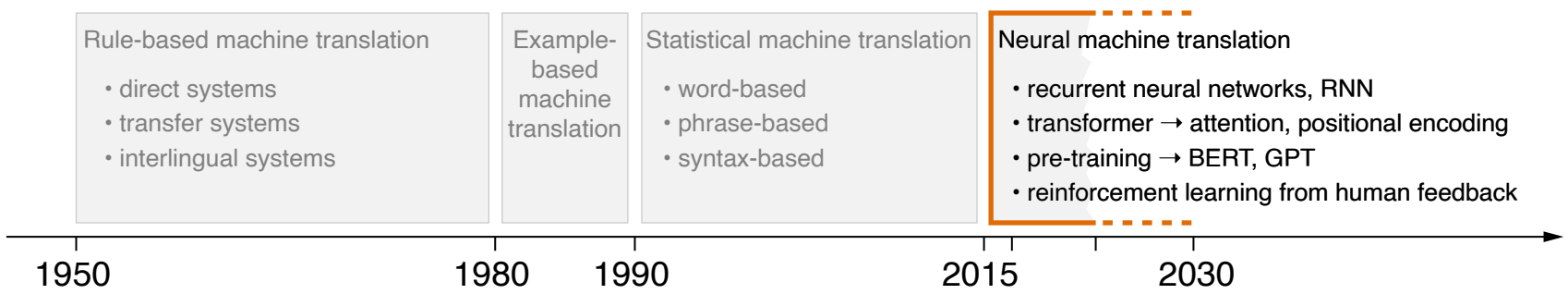
A neural language model  
tackles the optimization problem via loss minimization.



A neural language model  
tackles the optimization problem via loss minimization.

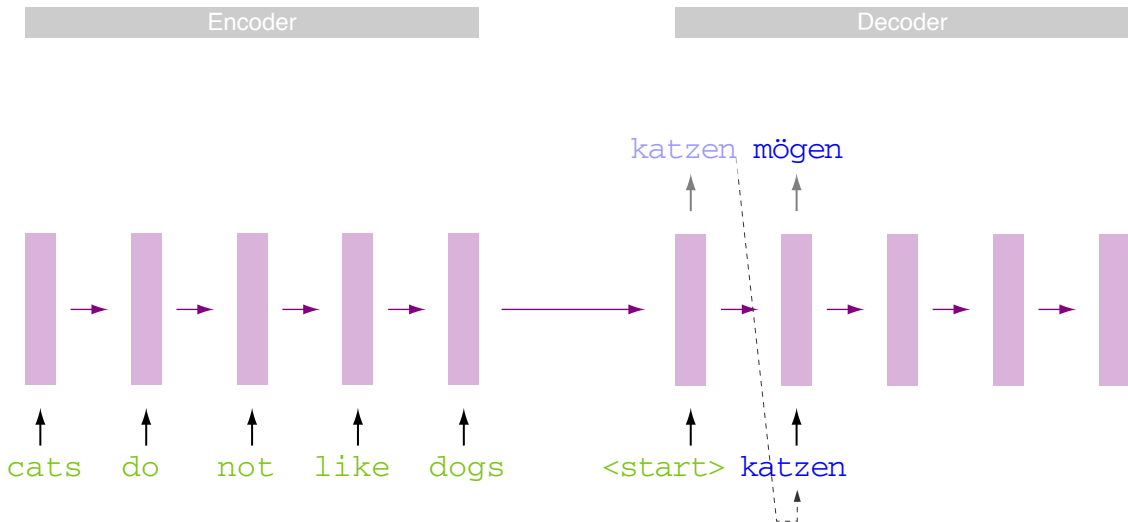
Illustration:

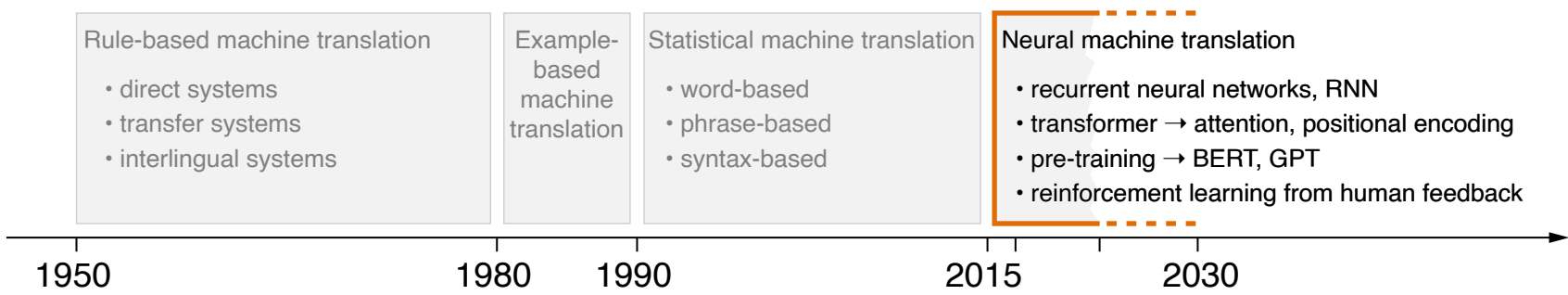




A neural language model tackles the optimization problem via loss minimization.

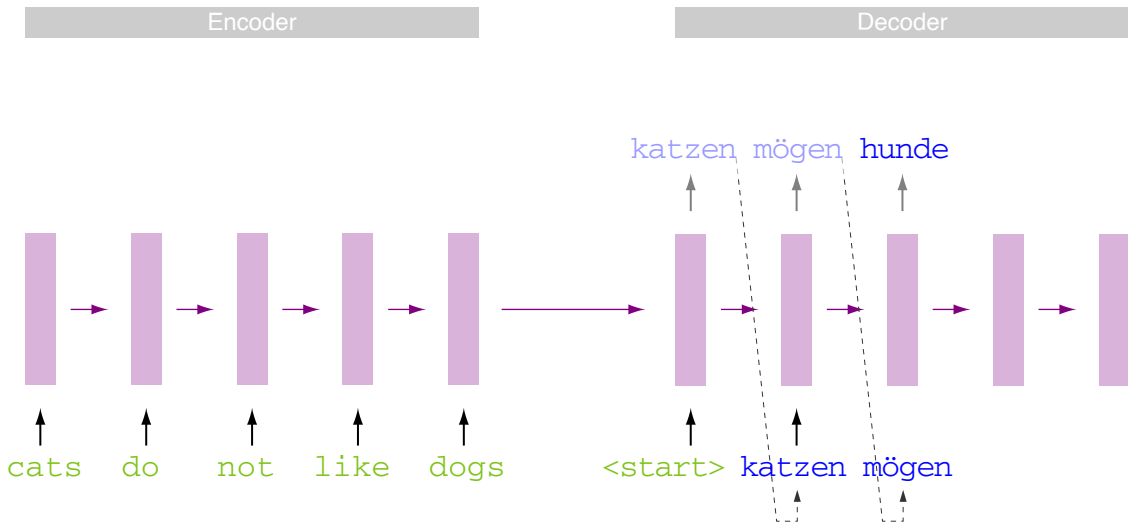
Illustration:

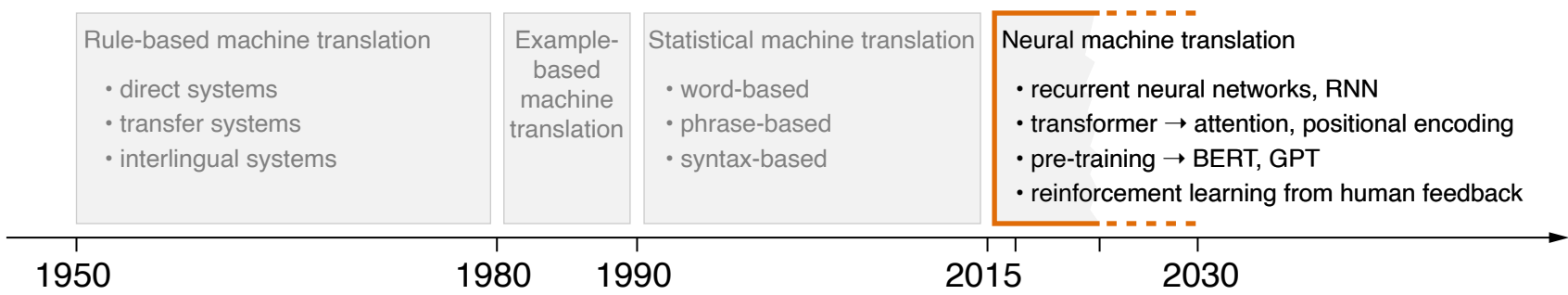




A neural language model tackles the optimization problem via loss minimization.

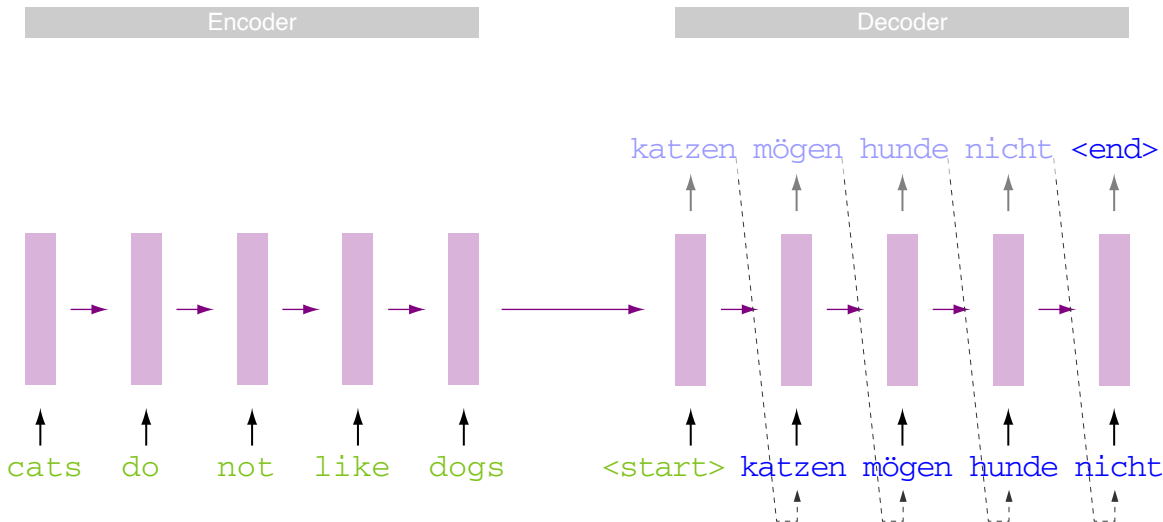
Illustration:

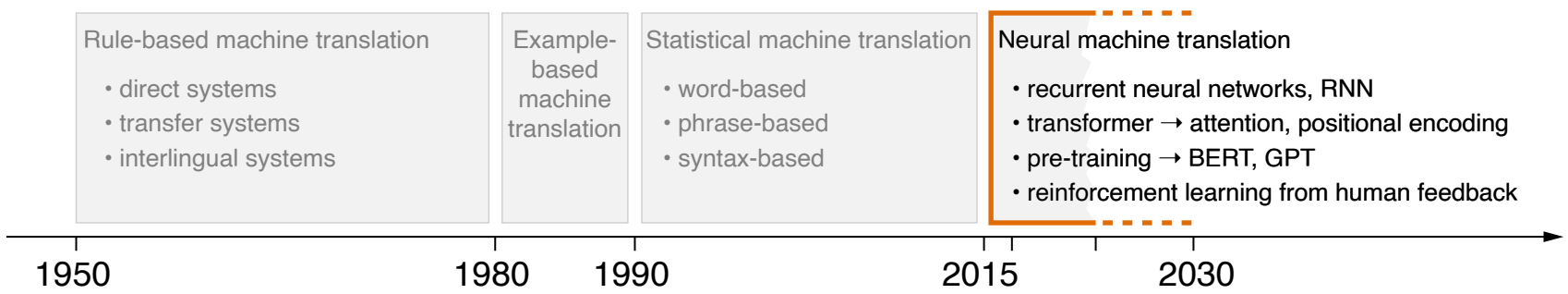




A neural language model tackles the optimization problem via loss minimization.

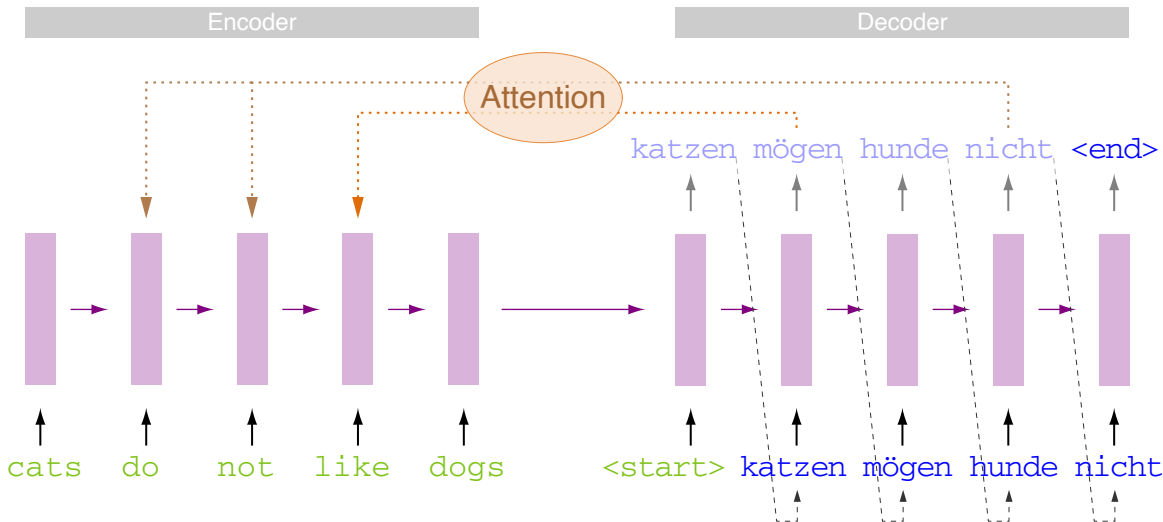
Illustration:

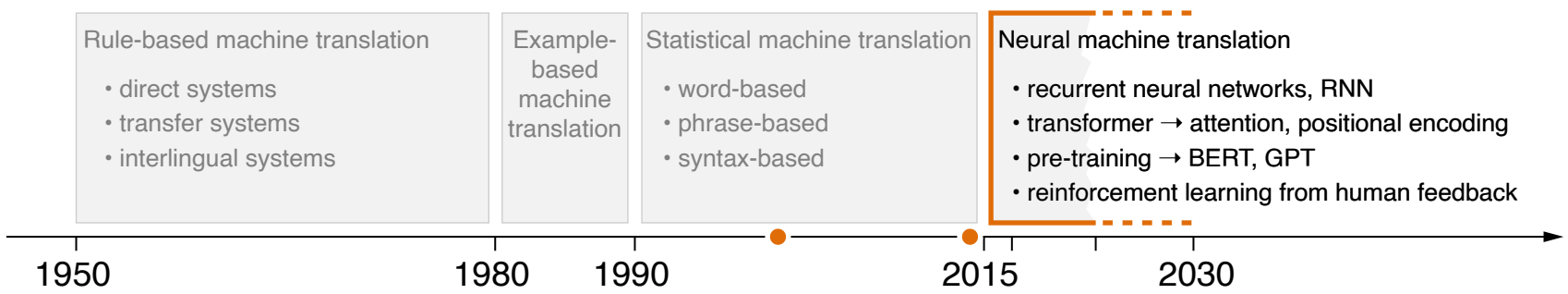




A neural language model  
tackles the optimization problem via loss minimization.

Illustration:





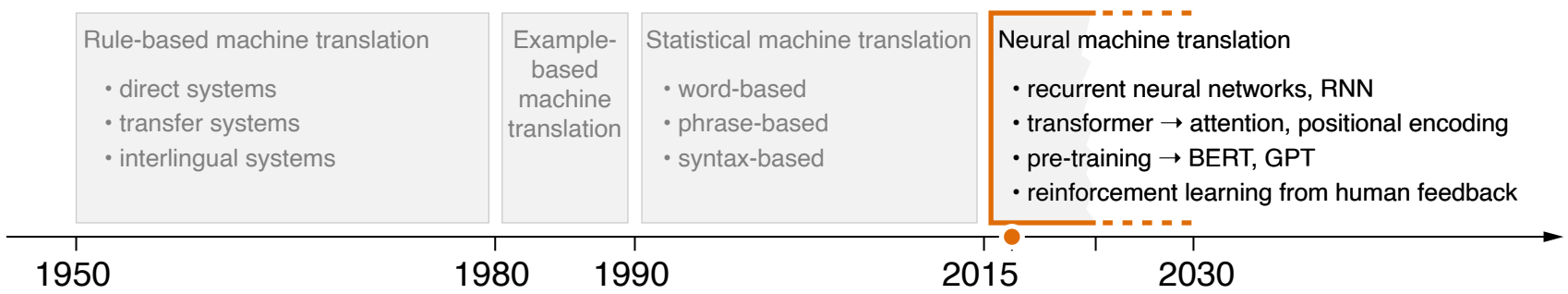
A neural language model  
tackles the optimization problem via loss minimization.

2000 The first (probabilistic) neural language model [Yoshua Bengio et al.]

⋮

2014 Recurrent neural language models with attention [Dzmitry Bahdanau et al.]

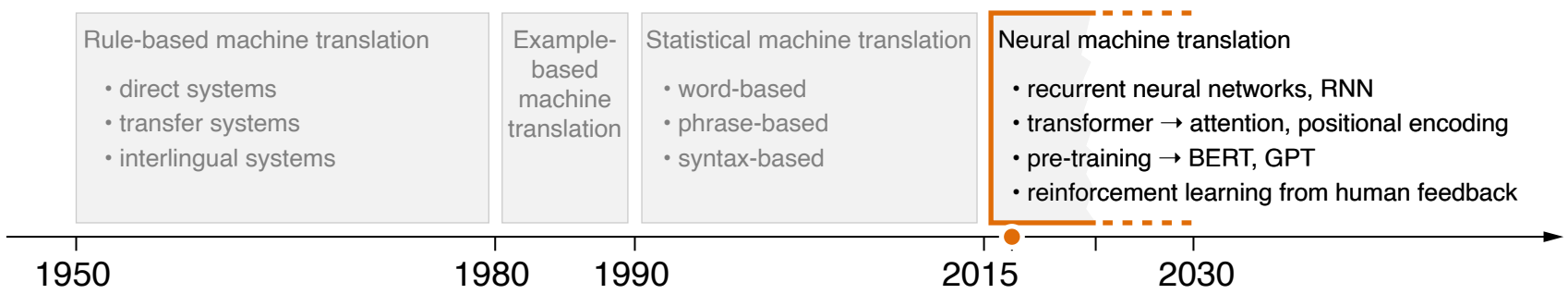
⋮



## 2017 The Transformer (fixed-length encoder / decoder) [Ashish Vaswani et al., Google]

- parallelization of computation through **positional encoding**
- distributional semantics through **self-attention**
- stack of multiple encoders and decoders





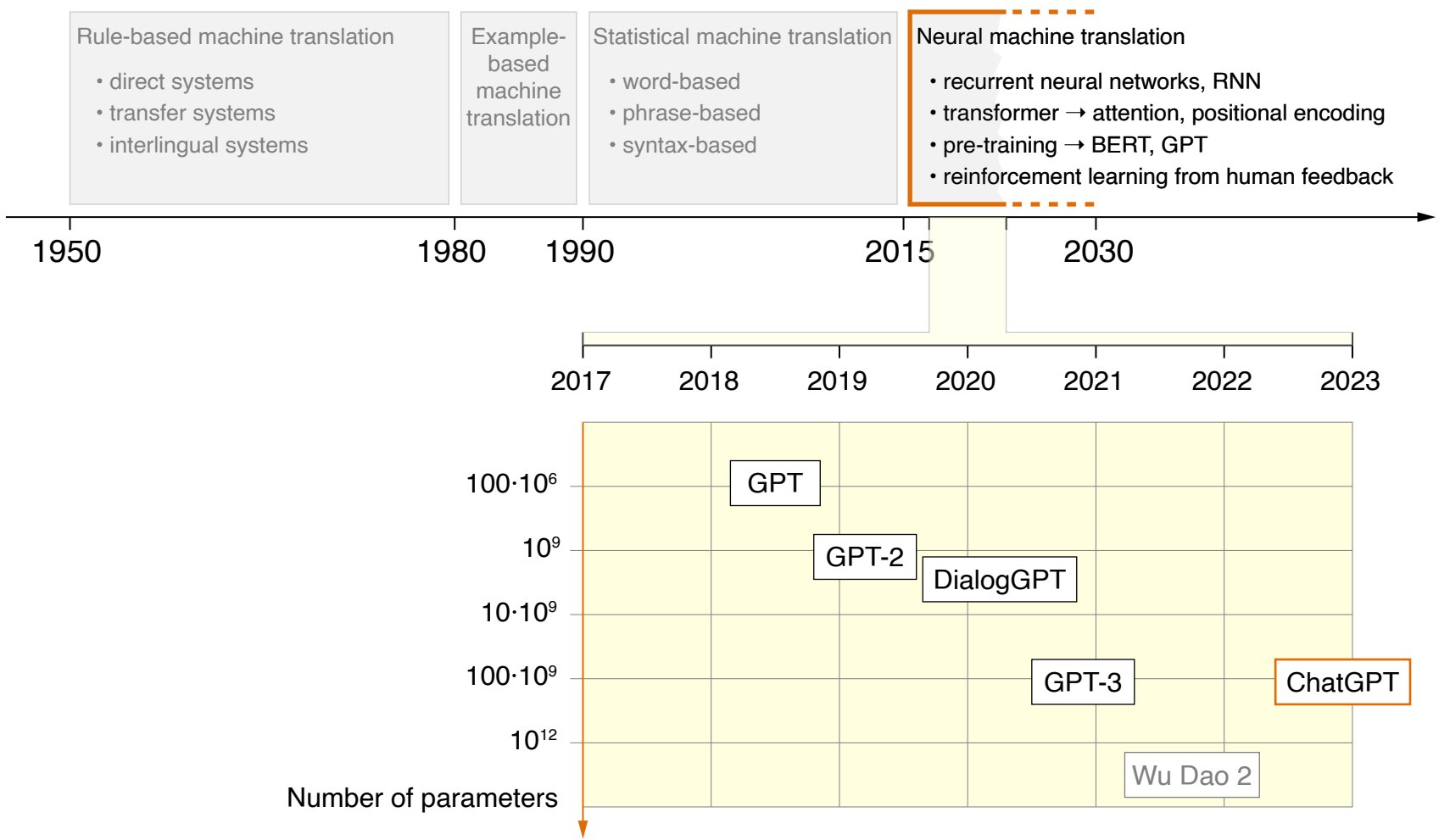
2017 **The Transformer** (fixed-length encoder / decoder) [Ashish Vaswani et al., Google]

- parallelization of computation through **positional encoding**
- distributional semantics through **self-attention**
- stack of multiple encoders and decoders

6/2018 **GPT** (generative pre-trained transformer) [Alec Radford et al., OpenAI]

- next word prediction

⋮





More on ChatGPT

### Training Corpora Sources

Wikipedia	11GB	Books	21GB
Journals	101GB	Reddit	50GB
CommonCrawl	570GB		

### Parameters

175,000,000,000  
( $175 \cdot 10^9$ )

### Computing / Training

- 355 years on a single Tesla V100 GPU.
- $\approx$  34 days on 1,024 x A100 GPUs.
- \$4.6M costs a single training run.

**GPT-3** [Jun. 2020]

### Training Corpora Sources

Wikipedia	11GB	Books	21GB
Journals	101GB	Reddit	50GB
CommonCrawl	570GB		

### Parameters

175,000,000,000  
( $175 \cdot 10^9$ )

### Computing / Training

- 355 years on a single Tesla V100 GPU.
- $\approx$  34 days on 1,024 x A100 GPUs.
- \$4.6M costs a single training run.

## GPT-3 [Jun. 2020]

### World Knowledge

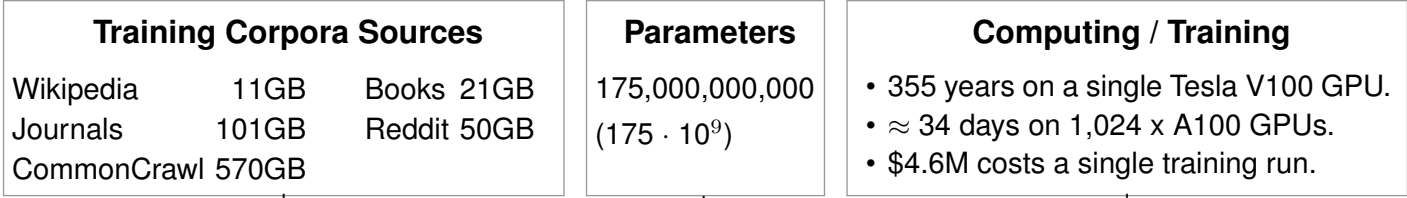
- 🧑 What city is in the northwest corner of Ohio?
- 🌀 Toledo is in the northwest corner of Ohio.

### Common Sense

- 🧑 Why don't have animals three legs?
- 🌀 Animals don't have three legs because they would fall over.

### Logical Reasoning

- 🧑 If I put a pencil in a box, then put another pencil in the box, what is in the box?
- 🌀 Two pencils.



**GPT-3** [Jun. 2020]

**World Knowledge**

🧑 What city is in the northwest corner of Ohio?

🌀 Toledo is in the northwest corner of Ohio.

**Common Sense**

🧑 Why don't have animals three legs?

🌀 Animals don't have three legs because they would fall over.

**Logical Reasoning**

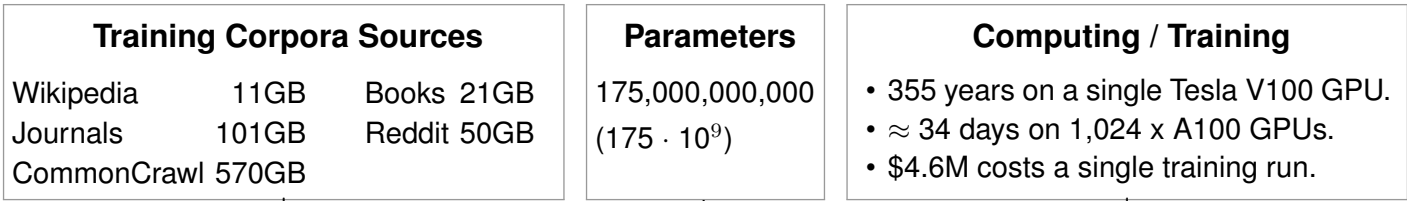
🧑 If I put a pencil in a box, then put another pencil in the box, what is in the box?

🌀 Two pencils.

**+ Learn to follow instructions and to comply with answer policies.**

- (1) Fine-tuning of GPT-3 to follow instructions: 13,000 popular prompts with hand-written answers.
- (2) Training of a reward model: 33,000 prompts with 4-9 answers, ranked from best to worse.
- (3) Training of the fine-tuned GPT-3 model from Step (1) to follow the reward policy.

**GPT-3.5 (InstructGPT)** [Jan. 2022]



**GPT-3** [Jun. 2020]

**World Knowledge**

🧑 What city is in the northwest corner of Ohio?

🌀 Toledo is in the northwest corner of Ohio.

**Common Sense**

🧑 Why don't have animals three legs?

🌀 Animals don't have three legs because they would fall over.

**Logical Reasoning**

🧑 If I put a pencil in a box, then put another pencil in the box, what is in the box?

🌀 Two pencils.

+ Learn to follow instructions and to comply with answer policies.

- (1) Fine-tuning of GPT-3 to follow instructions: 13,000 popular prompts with hand-written answers.
- (2) Training of a reward model: 33,000 prompts with 4-9 answers, ranked from best to worse.
- (3) Training of the fine-tuned GPT-3 model from Step (1) to follow the reward policy.

**GPT-3.5 (InstructGPT)** [Jan. 2022]

+ Fine-tuning of GPT-3.5 to comply with even stricter guardrails.

**ChatGPT** [Nov. 2022]

# More on ChatGPT

## Interacting



# More on ChatGPT

## Interacting

Zero-shot prompt: Only provide a description of a task.



Write an email to my friend Jay from me Steve thanking him for covering my shift this past Friday. Tell him to let me know if I can ever return the favor.



Subject: Thanks for covering my shift!

Dear Jay,

I hope this email finds you well. I just wanted to take a moment to express my heartfelt gratitude for covering my shift this past Friday. As you know, I was supposed to work but something urgent came up and I had to leave town at the last minute. Your willingness to cover my shift was a huge relief for me, and I really appreciate it.

I owe you one for sure, Jay. Please do not hesitate to let me know if there is anything I can do for you in the future. Whether it's covering your shift, helping you move, or anything else, I am here for you. You can always count on me.

Once again, thank you very much for being an amazing friend and coworker. Your help was invaluable and I am truly grateful.

Best regards,

Steve

# More on ChatGPT

## Interacting

One-shot prompt: Provide one example to learn from how to complete a task.



A list of actors in the movie Star Wars

1. Mark Hamill: Luke Skywalker



1. Harrison Ford: Han Solo
2. Carrie Fisher: Princess Leia Organa
3. Adam Driver: Kylo Ren/Ben Solo
4. Daisy Ridley: Rey
5. John Boyega: Finn
6. Oscar Isaac: Poe Dameron
7. Anthony Daniels: C-3PO
8. Kenny Baker: R2-D2
9. Peter Mayhew: Chewbacca

# More on ChatGPT

## Interacting

Few-shot prompt: Provide multiple (10 - 100) examples.



This is a conversation between Steve, the author of the book Exploring GPT-3 and someone who is reading the book.

Reader: Why did you decide to write the book?

Steve: Because I'm super fascinated by GPT-3 and emerging technology in general.

Reader: What will I learn from this book?

Steve: The book provides an introduction to GPT-3 from OpenAI. You'll learn what GPT-3 is and how to get started using it.

Reader: Do I need to be a coder to follow along?

Steve: No. Even if you've never written a line of code before, you'll be able to follow along just fine.

Reader:



That's great to hear! What inspired you to explore GPT-3 in particular?

Steve: Well, GPT-3 is one of the most advanced AI language models in the world, and it has the potential to revolutionize the way we interact with technology. I was particularly drawn to it because of its ability to generate human-like text, and I wanted to share my knowledge and experience with others who were interested in learning more about it.

# More on ChatGPT

## Interacting

Chain-of-thought prompt: Ask to solve a task in small steps, one after another.



The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Let's think step by step.



Sure, let's take this step by step.

Step 1: Start with the number of apples the cafeteria had initially: 23 apples.

Step 2: They used 20 of those apples to make lunch. So, you subtract the apples used from the initial amount. That's  $23 \text{ apples} - 20 \text{ apples} = 3 \text{ apples}$ .

Step 3: Then they bought 6 more apples. You add the apples they bought to the number they had left. So, that's  $3 \text{ apples} + 6 \text{ apples} = 9 \text{ apples}$ .

So, after using 20 for lunch and buying 6 more, the cafeteria now has 9 apples.

# More on ChatGPT

## Where ChatGPT can Struggle\*

### 1. Reasoning. Lacks an explicit world model.

- ❑ spatial reasoning: understand and manipulate relationships between objects
- ❑ physical reasoning: understand and manipulate physical objects
- ❑ temporal reasoning: predictions about events and their ordering in time
- ❑ psychological reasoning: predictions about human behavior and mental processes

### 2. Logic. Lacks a formal reasoning system.

- ❑ it can generate text that appears logically sound
- ❑ it can deal with small reasoning problems

### 3. Maths. Is limited in its ability to calculate mathematical expressions.

Stephen Wolfram suggests connecting ChatGPT to Wolfram|Alpha and all its computational knowledge “superpowers”.

### 4. Factual errors. Cannot access information from external sources.

However, the information it provides may seem credible and is well presented.

---

\* Ali Borji. A Categorical Archive of ChatGPT Failures. arXiv:2302.03494v1

# More on ChatGPT

## Where ChatGPT can Struggle (continued)

### 5. Bias. Recognizes and generalizes patterns from training data.

The data contains biases from society and history, and if proper measures are not taken, the trained model is likely to reproduce these biases.

### 6. Wit and humor. Has some understanding of humor.

Basically, large language models can handle humor; a comprehensive examination has yet to be conducted.

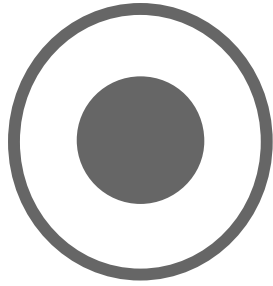
### 7. Coding. Excels at tackling various programming issues.

Given a description in English, code generated in a programming language of choice. Obviously, there is danger in using ChatGPT for creating malicious code that can be used for hacking or spamming.

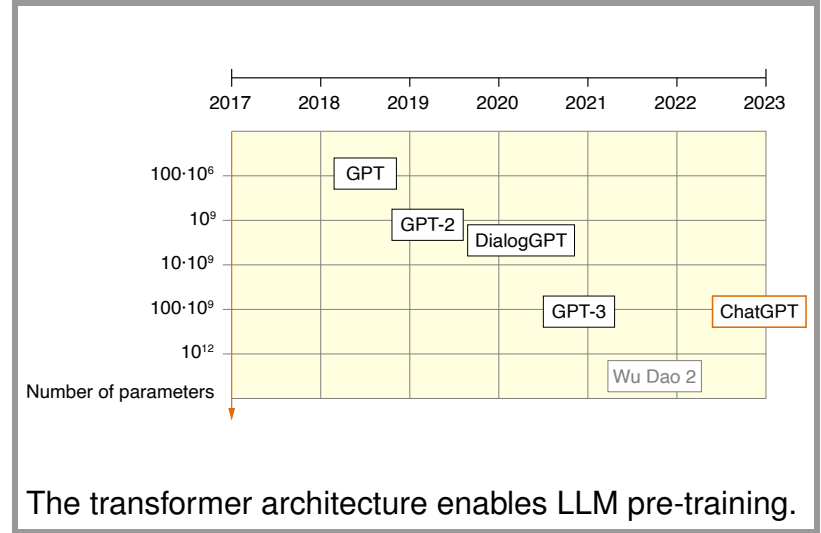
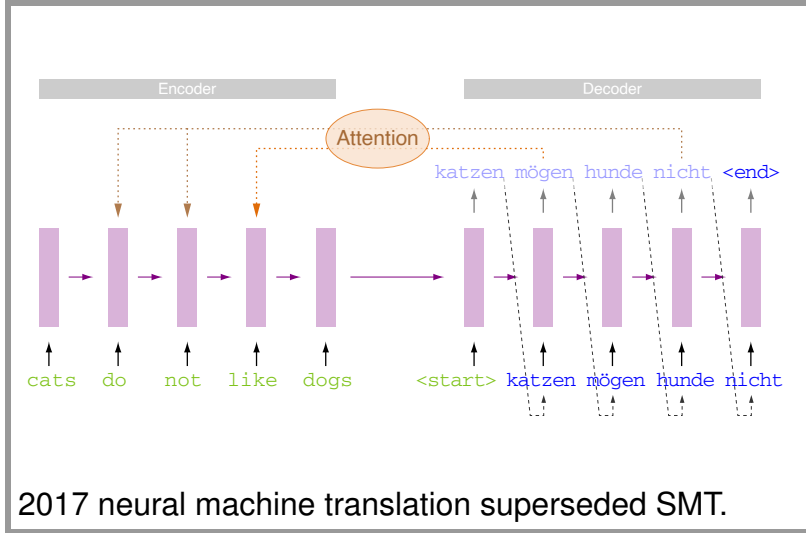
### 8. Self Awareness. Cannot access information from external sources.

ChatGPT is unaware of the details of its own architecture. This lack of understanding may have been intentionally imposed by OpenAI to protect the information about the model.

**Human versus AI comparison:** [GPT-4 vs human, Mar. 2023] [LLM tests, Nov. 2022]

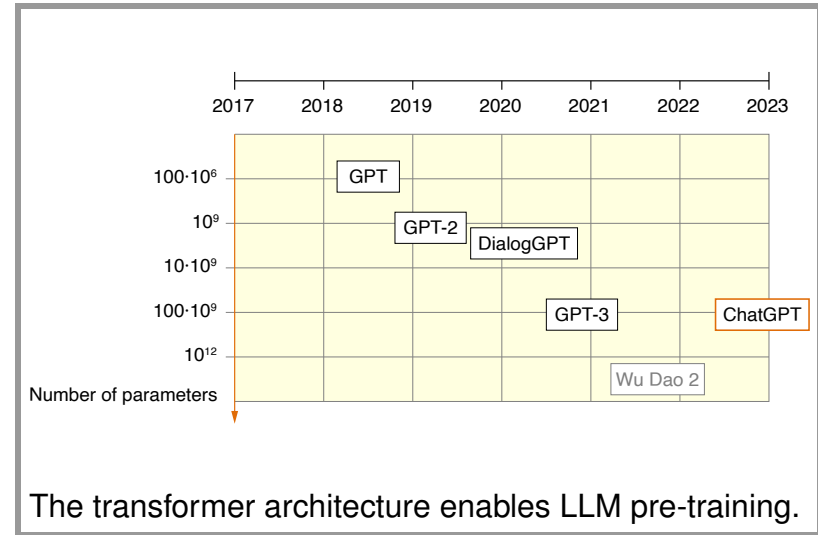
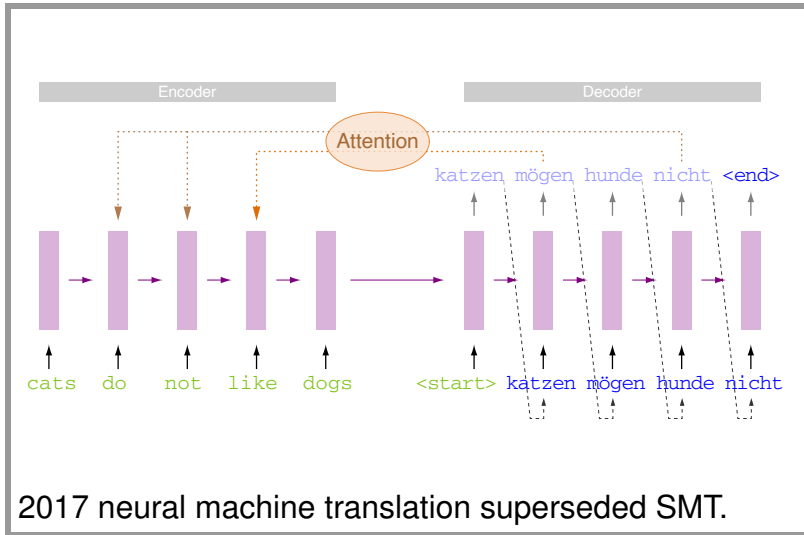


# Summary





# Summary



- ❑ GPT, ChatGPT, etc. are large, pre-trained language models.
- ❑ Language models accurately *simulate* human language behaviors.
- ❑ Language models don't have an explicit reasoning mechanism.
- ❑ Reinforcement learning from human feedback brings a quality boost.
- ❑ Development is extremely fast: GPT-4, PaLM (Google), LLaMA (Meta).

Thank You!



Netspeak - Mozilla Firefox

Netspeak

https://netspeak.org/#q=see+...+works

133x

**Netspeak** One word leads to another.

English German

see ... works i X ↺

how to ? this  
 see ... works  
 it's [ great well ]  
 and knows #much  
 { more show me }  
 m...d ? g?p

The ? finds one word.  
 The ... finds many words.  
 The [ ] compare options.  
 The # finds similar words.  
 The { } check the order.  
 The space is important.

see <b>how it</b> works	150,000	20%
see <b>if it</b> works	100,000	14%
see works	57,000	7.5%
see <b>how this</b> works	55,000	7.3%
see <b>what</b> works	51,000	6.7%
see <b>the</b> works	51,000	6.7%
see <b>if that</b> works	28,000	3.7%
see <b>your good</b> works	28,000	3.7%
see <b>how that</b> works	25,000	3.3%
see <b>how technorati</b> works	23,000	3.0%
see <b>if this</b> works	17,000	2.3%
see <b>more</b> works	17,000	2.2%
see <b>if it really</b> works	15,000	2.1%
see <b>his</b> works	12,000	1.7%
see <b>how well it</b> works	11,000	1.5%
see <b>other</b> works	8,900	1.2%

Netspeak - Mozilla Firefox

Netspeak

https://netspeak.org/#q=i+love+my+?

133x

**Netspeak** One word leads to another.

English German

i love my ?|

i x ↺

how to ? this  
 see ... works  
 it's [ great well ]  
 and knows #much  
 { more show me }  
 m...d ? g?p

The ? finds one word.  
 The ... finds many words.  
 The [ ] compare options.  
 The # finds similar words.  
 The { } check the order.  
 The space is important.

i love my <b>job</b>	72,000	10%
i love my <b>country</b>	44,000	6.2%
i love my <b>family</b>	41,000	5.9%
i love my <b>wife</b>	38,000	5.4%
i love my <b>new</b>	34,000	4.9%
i love my <b>friends</b>	33,000	4.7%
i love my <b>pet</b>	27,000	3.8%
i love my <b>dog</b>	26,000	3.7%
i love my <b>husband</b>	26,000	3.7%
i love my <b>life</b>	24,000	3.4%
i love my <b>baby</b>	24,000	3.4%
i love my <b>soldier</b>	22,000	3.1%
i love my <b>cat</b>	21,000	2.9%
i love my <b>computer</b>	18,000	2.6%
i love my <b>work</b>	16,000	2.4%
i love my <b>mom</b>	16,000	2.3%





ChatNoir

Bonjour! I'm ChatNoir, your friendly cat assistant. How can I help you today?

How can a computer's intelligence be tested?



ChatNoir

To test computer intelligence, you can give it problems to solve and evaluate how it performs. You can also give it tasks that require it to think outside the box, such as creating its own algorithms or coming up with innovative solutions to problems. Additionally, you can give it tests that measure its knowledge and understanding of a given subject.

This is outlined in [1], which also covers other definitions and tests of machine intelligence proposed, such as [2] which provides an overview of these definitions and tests. Finally, [3] is a standard test for machine intelligence.

[www.wired.com/.../pass-turing-ai-test/](http://www.wired.com/.../pass-turing-ai-test/)

[1] **How to Pass the Turing Artificial Intelligence Test | WIRED**

ClueWeb22 · Crawled Aug 2022

**How to Pass the Turing Artificial Intelligence Test | WIRED** WIRED Staff Science Jun 19, 2012 2:00 PM **How to Pass the Turing Artificial Intelligence Test** Are you human or a machine? Prove it, by passing the Turing Test – a test of the ability of a machine to exhibit intelligent behavior.

[www.shivonzilis.com/](http://www.shivonzilis.com/)

[2] **Machine Intelligence - Shivon Zilis**

ClueWeb22 · Crawled Aug 2022

Teams need to figure out **how to test** newfound capabilities, and applications need to change so they offer the a system of record; they also need to coach employees and learn from the data they enter.

[www.geeksforgeeks.org/turing-test-artificial-intelligence/](http://www.geeksforgeeks.org/turing-test-artificial-intelligence/)

[3] **Turing Test in Artificial Intelligence - GeeksforGeeks**

ClueWeb22 · Crawled Aug 2022

If a machine engages in a conversation with a human **how to** process the data it has been demonstrated by a machine, He has proposed the following skills of the test as follows: The turning judges the conversational skills humans.

Show more...



## Was ist eine Klausurtagung?



Eine Klausurtagung ist eine Zusammenkunft von Personen, die in der Regel mehrere Tage dauert und sich zum Ziel gesetzt hat, ein bestimmtes Thema oder Problem intensiv zu bearbeiten. Dabei findet die Klausurtagung in der Regel außerhalb des normalen Arbeitsumfelds statt, um ungestört und konzentriert arbeiten zu können.



Klausurtagungen werden häufig von Unternehmen, Regierungen oder Organisationen abgehalten, um wichtige Entscheidungen zu treffen, Strategien zu entwickeln oder Probleme zu lösen. Teilnehmer werden oft eingeladen, um ihre Expertise und Erfahrungen zu teilen und gemeinsam an einer Lösung zu arbeiten.

In der Regel finden Klausurtagungen in einer entspannten Umgebung statt und es wird viel Wert auf eine gute Organisation und eine produktive Arbeitsatmosphäre gelegt. Die Teilnehmer haben meist wenig oder keine Ablenkungen und können sich vollständig auf das Thema konzentrieren, um möglichst effektiv zu arbeiten.





Was wäre eine wichtige Entscheidung für ein Thüringer Ministerium?



Es gibt viele Entscheidungen, die ein Thüringer Ministerium treffen kann, je nach Ressort und aktuellem politischen oder gesellschaftlichen Kontext. Hier sind einige Beispiele:



- Ein Ministerium für Bildung könnte eine wichtige Entscheidung darüber treffen, wie die Schulen in Thüringen reformiert werden sollen, um die Qualität der Bildung zu verbessern und die Schülerinnen und Schüler besser auf die Anforderungen der Zukunft vorzubereiten.
- Ein Ministerium für Wirtschaft und Energie könnte eine wichtige Entscheidung darüber treffen, wie Thüringen seine Wirtschaft stärken kann, um Arbeitsplätze zu schaffen und die Energieversorgung sicherzustellen, während gleichzeitig die Umwelt geschützt wird.
- Ein Ministerium für Inneres und Kommunales könnte eine wichtige Entscheidung darüber treffen, wie Thüringen seine Sicherheitskräfte besser ausstatten und organisieren kann, um Terrorismus und Extremismus zu bekämpfen und die öffentliche Sicherheit zu gewährleisten.
- Ein Ministerium für Umwelt, Energie und Naturschutz könnte eine wichtige Entscheidung darüber treffen, wie Thüringen seine Umwelt schützen und den Klimawandel bekämpfen kann, indem es erneuerbare Energien fördert, den Verkehr umweltfreundlicher gestaltet und den Schutz von Wäldern und



Schreibe einen Satz, dessen Wörter alle mit dem Buchstaben "a" beginnen.



Ab acht Uhr abends haben alle Affen auf dem Ast Apfelsaft getrunken.



"Uhr" gehört da nicht rein.



Oh, es tut mir leid, da habe ich mich vertan. Hier ist eine korrigierte Version:

Alle arglosen Affen aßen abends anmutig an ausgewählten Apfelarten.



Schreibe mir eine Einleitung zum Thema Erstbesteigung des Mount Everest.



Die Erstbesteigung des Mount Everest im Jahr 1953 durch Sir Edmund Hillary und Tenzing Norgay ist eine der bekanntesten Errungenschaften im Bergsteigen. Der Mount Everest ist mit einer Höhe von 8.848 Metern der höchste Berg der Welt und wird von Bergsteigern auf der ganzen Welt als ultimatives Ziel angesehen. Die Geschichte der Erstbesteigung des Mount Everest ist von vielen Herausforderungen, Rückschlägen und Triumph geprägt und hat die Grenzen des menschlichen Potenzials und Durchhaltevermögens aufgezeigt. Die Bedeutung der Erstbesteigung des Mount Everest geht jedoch weit über das Bergsteigen hinaus und hat Auswirkungen auf viele Bereiche des menschlichen Lebens, von der Wissenschaft und Technologie bis hin zu internationalen Beziehungen und der menschlichen Natur selbst. In dieser Einleitung werden wir uns mit der Geschichte, Bedeutung und Erbe der Erstbesteigung des Mount Everest auseinandersetzen.





Und weiter?



Die Geschichte der Erstbesteigung des Mount Everest beginnt bereits in den frühen 1920er Jahren, als britische Bergsteiger begannen, Expeditionen in die Region zu unternehmen. In den folgenden Jahrzehnten fanden mehrere Versuche statt, den Gipfel zu erreichen, darunter auch tragische Expeditionen mit Verlusten von Menschenleben. Der Durchbruch kam schließlich im Jahr 1953, als Sir Edmund Hillary und Tenzing Norgay die Spitze des Mount Everest erreichten und damit in die Geschichte eingingen.



Die Erstbesteigung des Mount Everest hat nicht nur im Bergsteigen, sondern auch in vielen anderen Bereichen einen großen Einfluss gehabt. Die technologischen Fortschritte, die im Rahmen der Expedition gemacht wurden, haben zum Beispiel die Entwicklung von Ausrüstung und Kleidung für extreme Bedingungen vorangetrieben. Darüber hinaus haben die internationalen Beziehungen und Zusammenarbeit, die bei der Erstbesteigung des Mount Everest zum Tragen kamen, zur Förderung des Friedens und Verständnisses zwischen verschiedenen Ländern beigetragen.

Das Erbe der Erstbesteigung des Mount Everest ist auch heute noch spürbar. Der Berg zieht nach wie vor Bergsteiger und Abenteurer aus aller Welt an und stellt eine bedeutende Herausforderung dar. Gleichzeitig bleibt die Frage nach



Schreibe mir ein LISP-Programm zur Berechnung von Primzahlen.



Gerne! Hier ist ein LISP-Programm zur Berechnung von Primzahlen:



scss

Copy code

```
(defun is-prime (n)
  (cond
    ((= n 2) t)
    ((or (= n 0) (= n 1)) nil)
    (t
     (do ((i 2 (+ i 1)))
         ((> i (isqrt n)) t)
         (when (zerop (mod n i))
           (return nil))))))

(defun isqrt (n)
  (isqrt-iter n 1))

(defun isqrt-iter (n guess)
  (if (good-enough-p guess n)
      guess
      (isqrt-iter n (improve guess n))))

(defun good-enough-p (guess n)
```

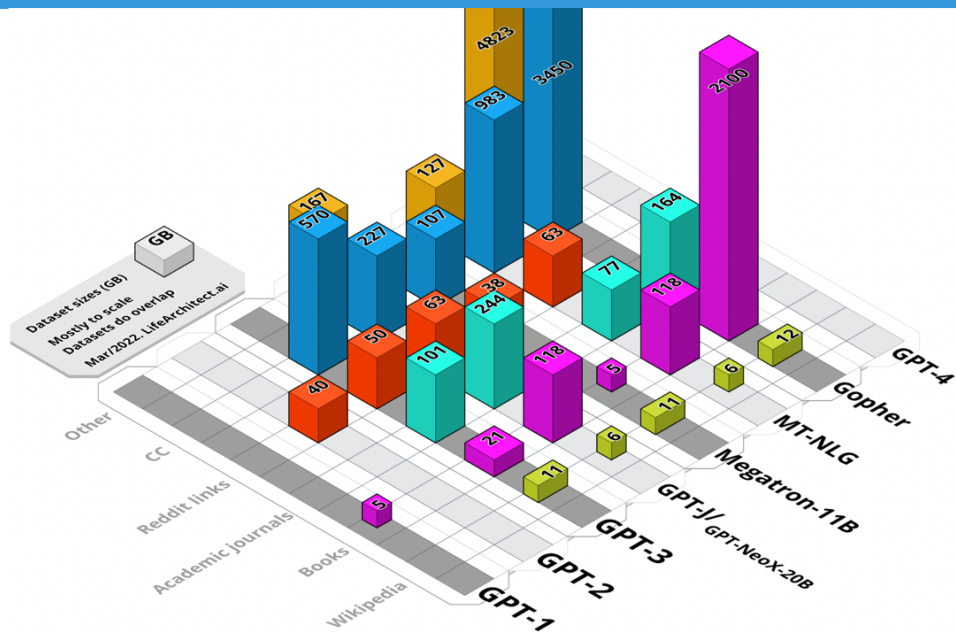
Das Programm definiert zuerst die Funktion `is-prime`, die eine Zahl `n` als Argument nimmt und zurückgibt, ob es sich bei dieser Zahl um eine Primzahl handelt oder nicht. Die Funktion überprüft, ob die Zahl gleich 2 ist oder ob sie 0 oder 1 ist. In allen anderen Fällen wird eine Schleife gestartet, die von 2 bis zur Quadratwurzel von `n` läuft. Wenn die Zahl `n` ohne Rest durch eine Zahl zwischen 2 und der Quadratwurzel von `n` geteilt werden kann, ist `n` keine Primzahl und die Funktion gibt `n+1` zurück. Andernfalls gibt die Funktion `t` zurück.

Das Programm definiert auch die Funktionen `isqrt`, `isqrt-iter`, `good-enough-p` und `improve`, die zusammen eine Schätzung der Quadratwurzel einer Zahl liefern.

Schließlich definiert das Programm die Funktion `prime-numbers`, die eine Zahl `n` als Argument nimmt und eine Liste aller Primzahlen von 2 bis `n` zurückgibt. Die Funktion verwendet eine Schleife, die von 2 bis `n` läuft und überprüft, ob jede Zahl eine Primzahl ist. Wenn eine Zahl eine Primzahl ist, wird sie der Liste `primes` hinzugefügt. Am Ende gibt die Funktion die Liste `primes` zurück.

In diesem Beispiel wird `prime-numbers` mit dem Argument 100 aufgerufen, um eine Liste aller Primzahlen von 2 bis 100 zu erhalten. Sie können die Zahl als Argument von `prime-numbers` ändern, um eine Liste aller Primzahlen bis zu Ihrer gewünschten Zahl zu erhalten.

# 2022 WHAT'S IN MY AI? – ALT VIEW



Source	Percentage
Google Patents.....	0.48%
The New York Times.....	0.06%
Los Angeles Times.....	0.06%
The Guardian.....	0.06%
Public Library of Science..	0.06%
Forbes.....	0.05%
Huffington Post.....	0.05%
Patents.com.....	0.05%
Scribd.....	0.04%
Other.....	99.09%

**Common Crawl**

Source	Percentage
Google.....	3.4%
Archive.....	1.3%
Blogspot.....	1.0%
GitHub.....	0.9%
The New York Times.....	0.7%
Wordpress.....	0.7%
Washington Post.....	0.7%
Wikia.....	0.7%
BBC.....	0.7%
Other.....	89.9%

**Reddit links**

Category	Percentage
Biography.....	27.8%
Geography.....	17.7%
Culture and Arts.....	15.8%
History.....	9.9%
Biology, Health, Medicine...	7.8%
Sports.....	6.5%
Business.....	4.8%
Other society.....	4.4%
Science & Math.....	3.5%
Education.....	1.8%

**English Wikipedia**

Category	Percentage
Romance.....	26.1%
Fantasy.....	13.6%
Science Fiction.....	7.5%
New Adult.....	6.9%
Young Adult.....	6.8%
Thriller.....	5.9%
Mystery.....	5.6%
Vampires.....	5.4%
Horror.....	4.1%
Other.....	18.0%

**BookCorpus (GPT-1 only)**

# GPT-4 VS HUMAN TESTS (MAR/2023)

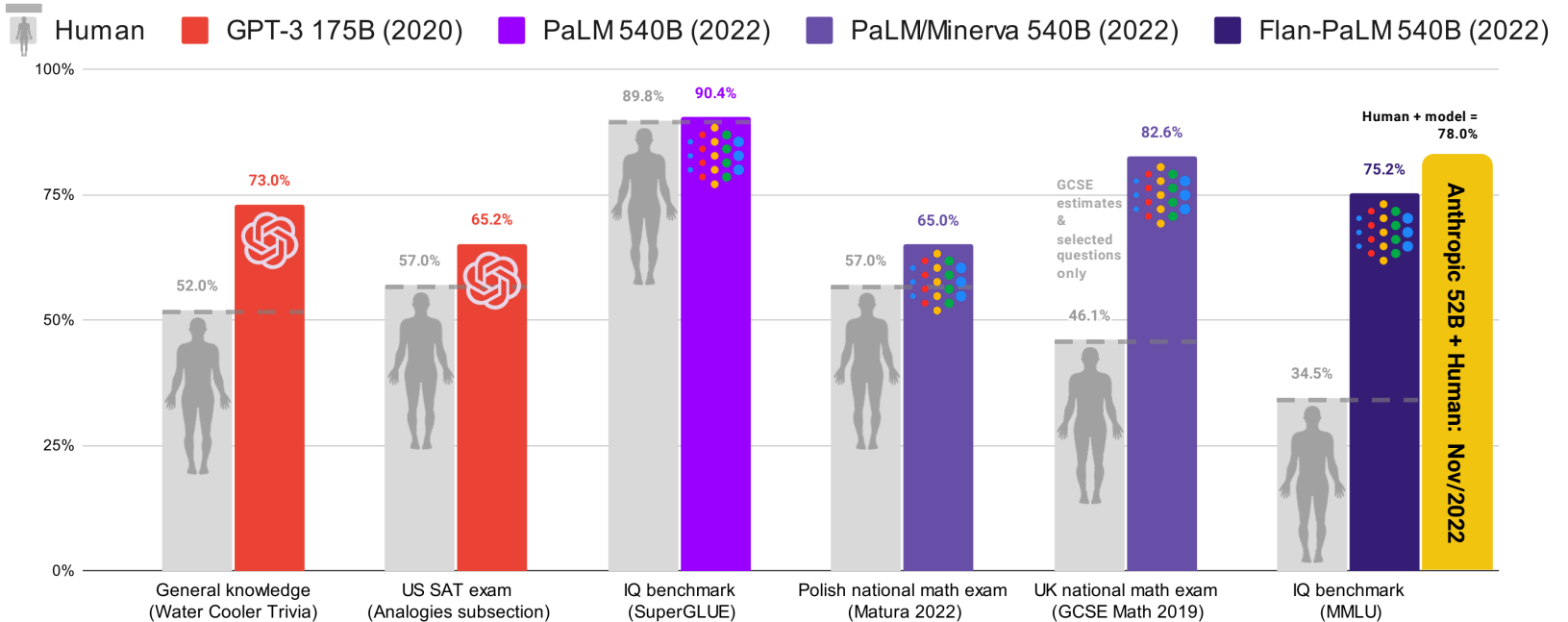


Selected highlights only. Percentiles; 50 refers to the 50<sup>th</sup> percentile as average, and may not be the testing average for some tests. Alan D. Thompson, March 2023. <https://life architect.ai/iq-testing-ai/>





# LANGUAGE MODEL TESTS (NOV/2022)



Selected highlights only. Sources: <https://lifecycle.ai/iq-testing-ai/>, Anthropic paper: <https://arxiv.org/abs/2211.03540> Other papers: <https://lifecycle.ai/papers>, Alan D. Thompson, November 2022, <https://lifecycle.ai/>