

AI And The Limits Of Language

An artificial intelligence system trained on words and sentences alone will never approximate human understanding

Jacob Browning & Yann LeCun
22. August 2022

The magazine NOEMA and authors

- In ancient Greek, noēma means “thinking” or the “object of thought.”
- Covers topics such as technology, philosophy, governance, economics, geopolitics, and culture
- It seeks to gain deeper insights into the pressing challenges of the 21st century



Jacob Browning -
postdoc in NYU's
Computer Science
Department
working on the
philosophy of AI

Yann LeCun - Turing
Award-winning
machine learning
researcher, an NYU
professor and the chief
AI scientist at Meta



Suspended Google engineer says the AI he believes to be sentient hired a lawyer

On Medium, Lemoine, who is also a Christian priest, referred to LaMDA as a person, and said the AI "has been incredibly consistent in its communications about what it wants and what it believes its rights are as a person."

However, Lemoine told Wired that "person and AI are different things."

"Human is a biological term," he said. "It is not AI that knows it's not a human."



Historical Perspective on Knowledge

- ➡ Historically - knowledge viewed as linguistic
- ➡ Ideal language imagined as purely formal and logical
- ➡ Everything to be known can be put into a book
- ➡ **Criticism:** Ability to talk about something \neq understanding
- ➡ Language is a form of representation and far from perfect
- ➡ Iconic knowledge, distributed knowledge

The Limits of Language

- ➡ Language is **low-bandwidth**

Chomsky: Language is just not a clear vehicle for communication

- ➡ Humans rely on nonlinguistic understanding and context
- ➡ LLMs mimic contextual understanding
- ➡ Training them on human-written text enables authentic interaction

“It is clear that these systems are doomed to a shallow understanding that will never approximate the full-bodied thinking we see in humans.”

Shallow Understanding

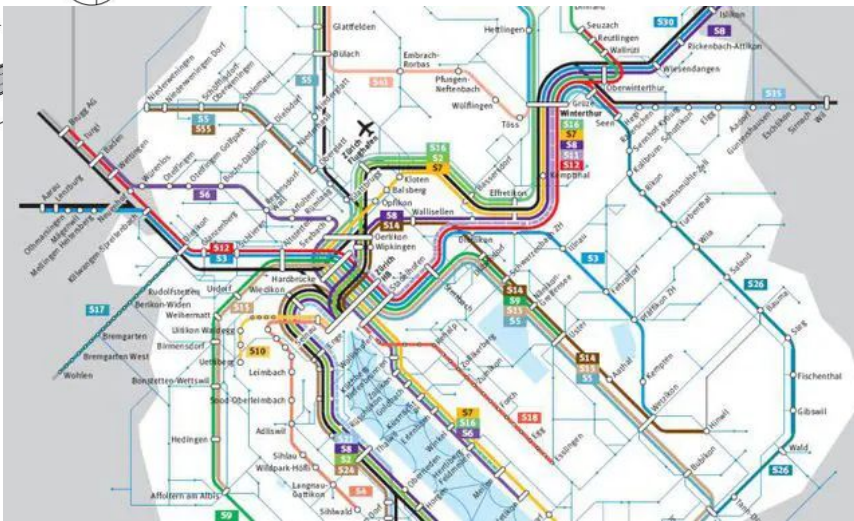
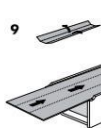
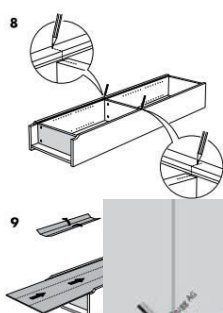
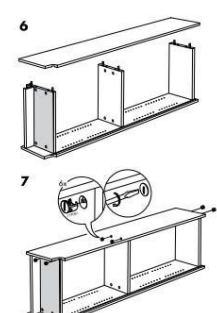
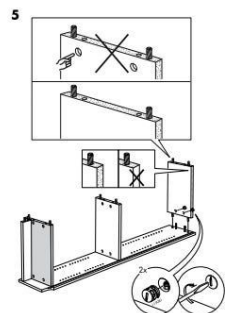
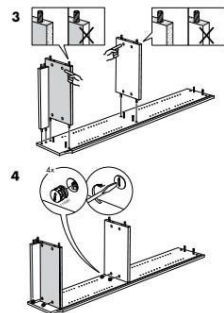
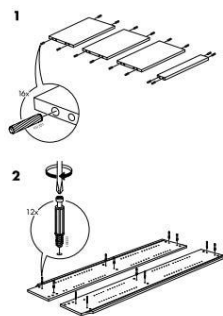
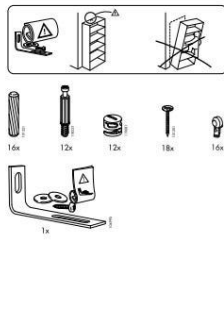
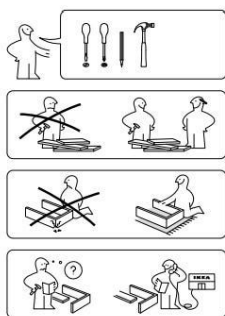
- ➡ Def.: Simple memorization of facts without the ability to use those facts in context
- ➡ LLMs accused of shallow understanding and mimicry
- ➡ The training process of LLMs is not natural to humans
- ➡ LLMs may **lack practical knowledge or the ability to apply** information in real-world scenarios (The octopus test)
- ➡ LLMs have limited attention span and may struggle with consistency
- ➡ While excelling in linguistic tasks, LLMs' understanding remains limited

“Abandoning the view that all knowledge is linguistic permits us to realize how much of our knowledge is nonlinguistic.”

Beyond Language

- ➡ Language is just one form of representation - not all knowledge is linguistic
- ➡ Humans mostly learn from interacting with the environment, especially children
- ➡ Imitation is a vital learning technique
- ➡ Language is still crucial, especially with advancements such as printing and the internet

BUT: Machines trained solely on language **have limited understanding** due to the narrow channel of information acquisition and they **offer an illusion** of depth and reflection like a mirror but lack true substance in understanding





a) Europe



b) Canada



c) USA

“A system trained on language alone will never approximate human intelligence, even if trained from now until the heat death of the universe.”

Consequences

- Training machines solely on language will never lead to true human-like awareness
- Language is useful when complementing nonlinguistic context
- Do not confuse LMMs' shallow understanding and deep understanding
- The goal is for AI systems to focus on the world being talked about, not the words themselves

Dealing with LLMs at any length makes apparent just how little can be known from language alone.

Cheryl's birthday problem



Cheryl's birthday problem

Albert and Bernard just became friends with Cheryl, and they want to know when her birthday is. Cheryl gives them a list of 10 possible dates:

May 15, May 16, May 19

June 17, June 18

July 14, July 16

August 14, August 15, August 17

Cheryl then tells Albert and Bernard separately the month and the day of her birthday respectively.

Albert: I don't know when Cheryl's birthday is, but I know that Bernard doesn't know too.

Bernard: At first I don't know when Cheryl's birthday is, but I know now.

Albert: Then I also know when Cheryl's birthday is.

So when is Cheryl's birthday?

		Day					
Month	May		15	16			19
	June				17	18	
	July	14		16			
	August	14	15		17		

Albert: I don't know when Cheryl's birthday is, but I know that Bernard doesn't know too.

Bernard: At first I don't know when Cheryl's birthday is, but I know now.

Albert: Then I also know when Cheryl's birthday is.

Cheryl's birthday problem

Day

May		15	16			19
June				17	18	
July	14		16			
August	14	15		17		

Albert: I don't know when Cheryl's birthday is, but I know that Bernard doesn't know too.

Cheryl's birthday problem

Day

	15	16			19
			17	18	
14		16			
14	15		17		

Bernard: At first I didn't know when Cheryl's birthday was, but I know now.

Cheryl's birthday problem

Day

	15	16			19
			17	18	
14		16			
14	15		17		

Albert: Then I also know when Cheryl's birthday is.



Let's break down the puzzle step by step:

1. Cheryl gives 10 possible dates:

- May 15, May 16, May 19
- June 17, June 18
- July 14, July 16
- August 14, August 15, August 17

2. Cheryl tells Albert the month and Bernard the day.

3. Albert: I don't know when Cheryl's birthday is, but I know that Bernard doesn't know too.

From this, we can deduce that Cheryl's birthday cannot be in May or June because if it were, there would be a chance that Bernard could know the birthday (if he were told '18' or '19' because those days are unique to May and June). So the possible dates are now:

- July 14, July 16
- August 14, August 15, August 17

⋮

So, Cheryl's birthday is on **July 16**.

Here is a math problem: John and Jon just became friends with Jonnie, and they want to know when her birthday is. Jonnie gives them a list of 10 possible dates: October 15, October 16, October 19, January 17, January 18, April 14, April 16, December 14, December 15, December 17.

Jonnie then tells John and Jon separately the month and the day of her birthday respectively.

John: I don't know when Jonnie's birthday is, but I know that Jon doesn't know too.

Jon: At first I don't know when Jonnie's birthday is, but I know now.

John: Then I also know when Jonnie's birthday is.

So when is Jonnie's birthday?

- Only May and June have unique days (May 19 and June 18). So, we can conclude that Jonnie's birthday is not in October or January.

The possible dates now are:

- April (14, 16)
- December (14, 15, 17)

Jon's statement: After hearing John, Jon, who only knows the day, says he now knows when Jonnie's birthday is. This means Jonnie's birthday must be on a day that is unique among the remaining possibilities.

From the narrowed list:

- Only April 14 and December 17 are unique.

John's next statement: After hearing Jon, John, who only knows the month, now also knows when Jonnie's birthday is. This means the birthday cannot be in December because there are still two dates left in December, and John wouldn't be able to determine the exact date.

So, the only remaining date is:

- April 14.

- This means that Jonnie's birthday cannot be in May or June, because if it were, there would be a chance that Jon could know the birthday (if he were told '18' or '19', unique days in the given list). Thus, we can eliminate:
October 19, January 17, and January 18.

- The remaining possible dates are:
October 15, October 16
April 14, April 16
December 14, December 15, December 17

Thus, Jonnie's birthday is on **December 17**.

Thus, Jonnie's birthday is on **April 14**.

April 16.

So, Jonnie's birthday is on April 16.

Nigel Richards

- ➡ World Scrabble champion
- ➡ In 2015, despite not speaking French, Richards won the French World Scrabble Championships, after reportedly spending nine weeks studying the French dictionary
- ➡ He won it again in 2018, and multiple duplicate titles from 2016

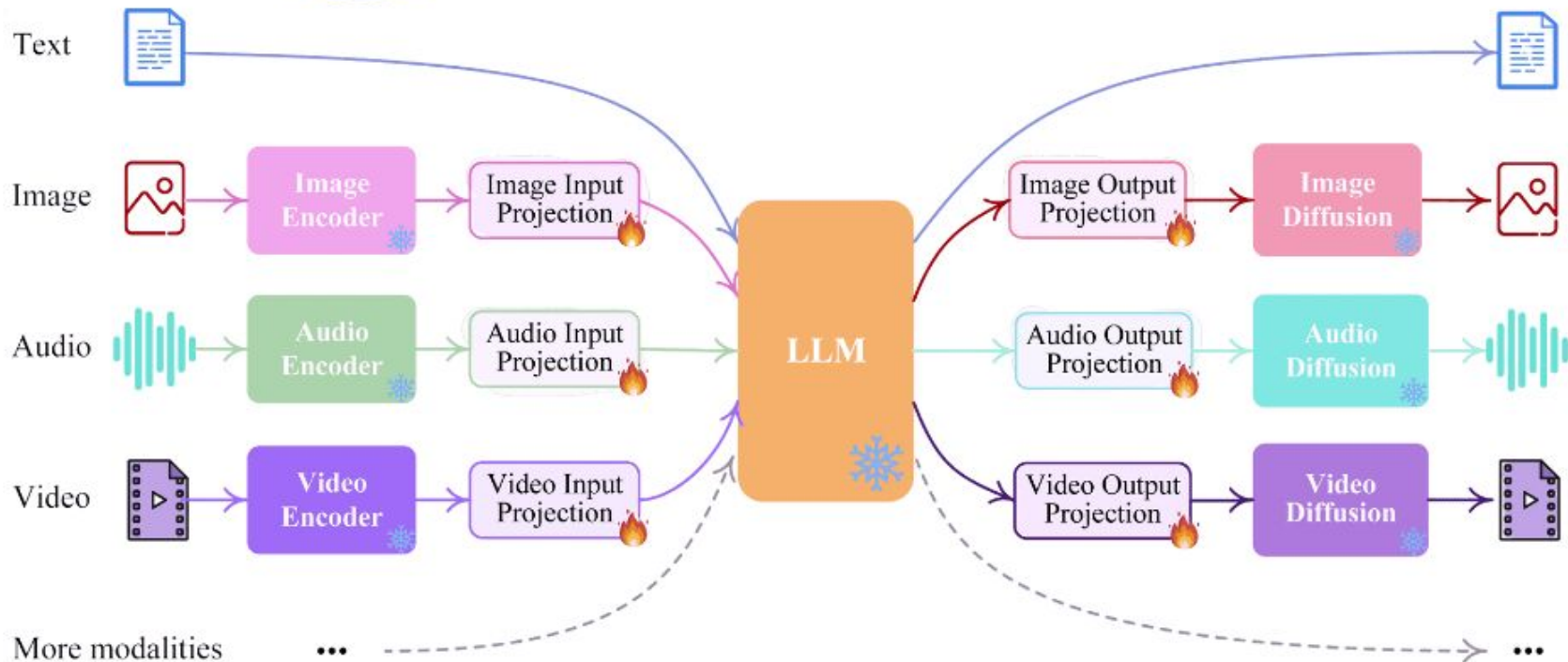


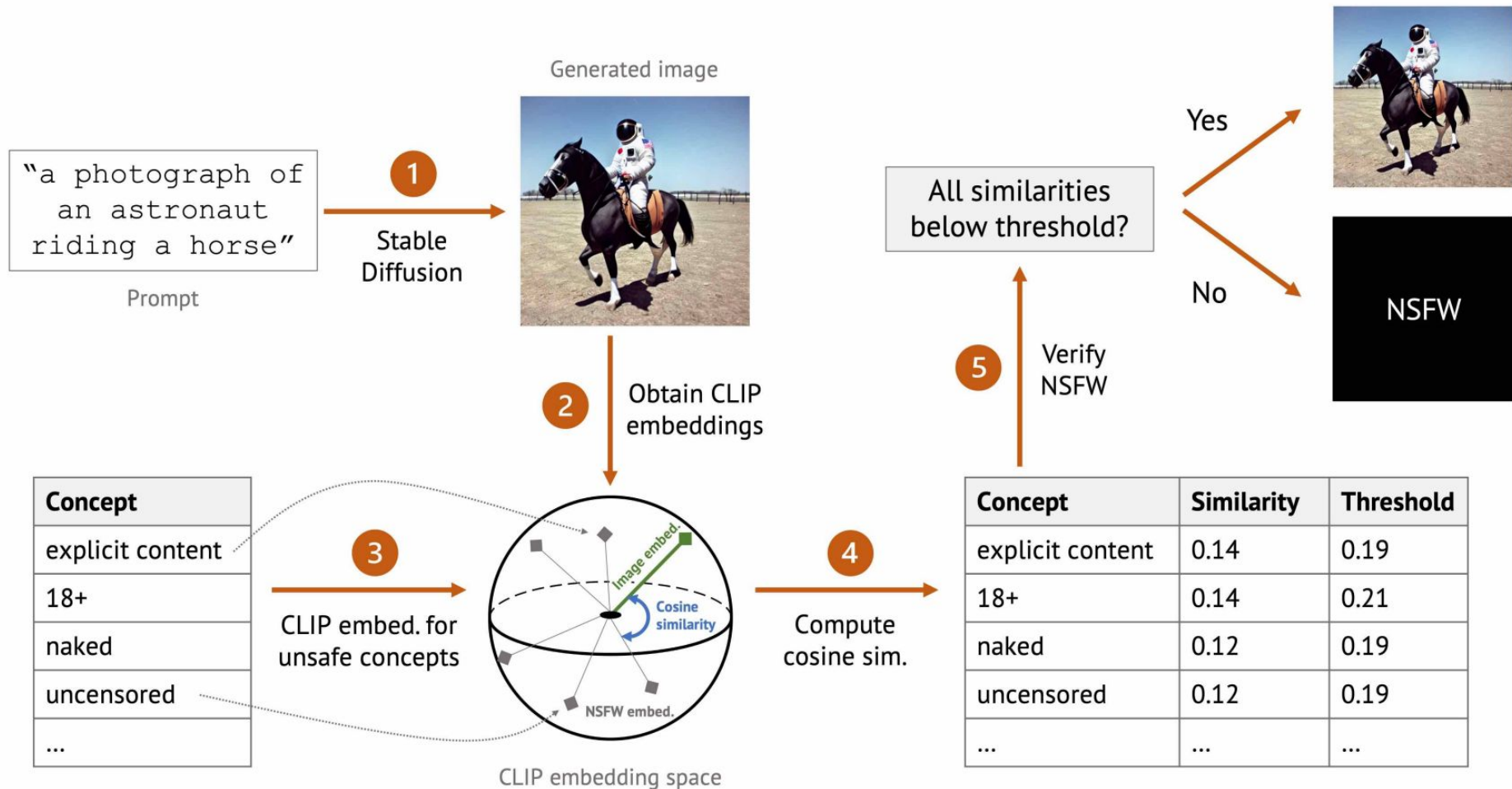
Multimodal LLMs

- ➡ Uses variety of inputs - text, audio, video, etc. – as prompts
- ➡ Can convert those prompts into various outputs, not just the source type
- ➡ For example Google's Gemini can receive a photo of a plate of cookies and generate a written recipe as a response and vice versa¹



NEXt-GPT





Related information

- Blake Lemoine: AI with a Soul:
<https://www.youtube.com/watch?v=d9ipv6HhuWM&t=1186s>
- Yuval Noah Harari (Sapiens) versus Yann Le Cun (Meta) on artificial intelligence:
https://www.lepoint.fr/sciences-nature/yuval-harari-sapiens-versus-yann-le-cun-meta-on-artificial-intelligence-11-05-2023-2519782_1924.php

Sources

Main article:

- <https://www.noemamag.com/ai-and-the-limits-of-language/>

Cheryl's birthday problem:

- https://en.wikipedia.org/wiki/Cheryl%27s_Birthday
- <https://www.bis.org/publ/bisbull83.htm>

Nigel Richards:

- [https://en.wikipedia.org/wiki/Nigel_Richards_\(Scrabble_player\)](https://en.wikipedia.org/wiki/Nigel_Richards_(Scrabble_player))
- <https://people.com/celebrity/nigel-richards-wins-french-scrabble-championship/>

NExT-GPT:

- <https://next-gpt.github.io>

Stable Diffusion:

- <https://arxiv.org/pdf/2210.04610>

	Encoder		Input Projection		LLM		Output Projection		Diffusion	
	Name	Param	Name	Param	Name	Param	Name	Param	Name	Param
Text	—	—	—	—	Vicuna [9] 7B❄️ (LoRA 33M🔥)		—	—	—	—
Image	ImageBind [15] 1.2B❄️		Linear 4M🔥				Transformer 31M🔥	SD [43] 1.3B❄️		
Audio							Transformer 31M🔥	AudioLDM [34] 975M❄️		
Video							Transformer 32M🔥	Zeroscope [5] 1.8B❄️		

Table 1: Summary of system configuration. Only 1% parameters need updating.