

Inside the Mind of a Language Model: Exploring the Capabilities and Applications of a Conversational Agent



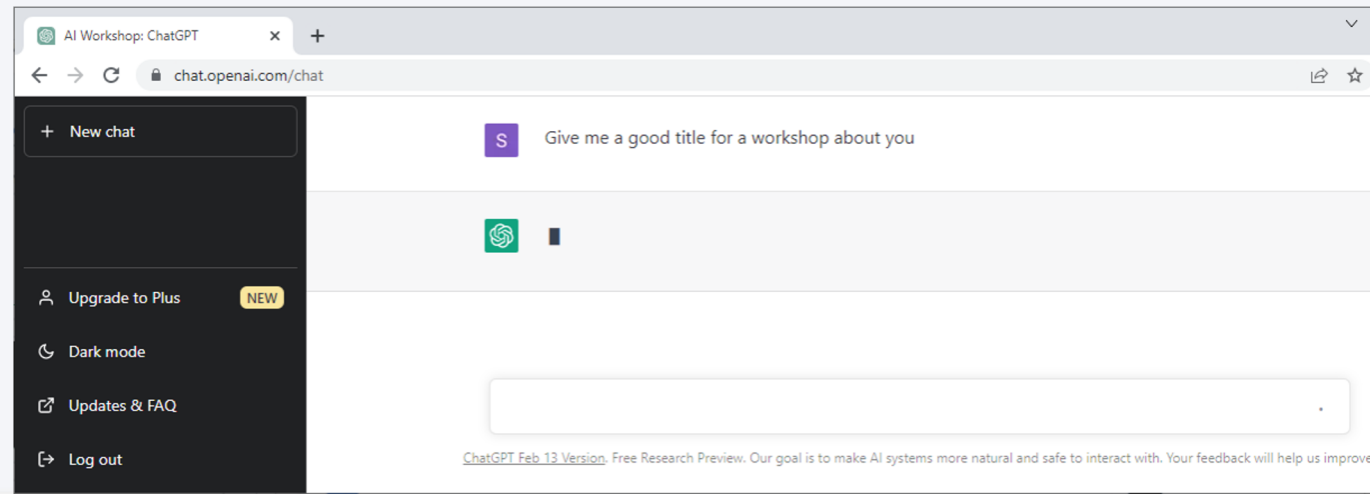
Julien Audiffren
Maître-Assistant



Simon Ruffieux
Maître-Assistant



Anna Jobin
Maître-Assistante



SUMMARY

PART I: TECHNICAL OVERVIEW

PART II: FROM AN HCI PERSPECTIVE

PART III: A SOCIETAL PERSPECTIVE

PART IV: DISCUSSION

PART I

TECHNICAL OVERVIEW

A Simple Game

Complete the sentence.

To be or not to [...]



To be or not to **be**.

The sun is [...]



The sun is **shining**.

Two main tricks:

1. Use previously known occurrences
2. Use context to complete

Probability Distribution.

How often does this sentence occurs ?

To be or not to [...]

Search results for "to be or not to". The top result is from poemanalysis.com, titled "Hamlet's 'To be, or not to be' Soliloquy - Poem Analysis". The snippet reads: "The 'To be, or not to be' quote is taken from the first line of Hamlet's soliloquy that appears in Act 3, Scene 1 of the eponymous play by William Shakespeare, 'Hamlet'. The full quote, 'To be, or not to be, that is the question' is famous for its open-ended meaning that not only encompasses the thoughts raging inside Hamlet's mind but also features the theme of existential crisis." The second result is from thefreedictionary.com, titled "To Be or Not to Be - Idioms by The Free Dictionary". The snippet reads: "to (do something) or not to (do something) [, that is the question] Used to express one's indecision or hesitation about doing something. Derived from the famous line in William Shakespeare's Hamlet, 'To be, or not to be, that is the question.' Usually used humorously, especially as a rhyming pun on 'be.' To tell Jake or not to tell him."

Probabilities for the next word*

be	99.9%
have	0.0%
Be	0.0%

The sun is [...]

Search results for "the sun is". The top result is from solarsystem.nasa.gov, titled "Overview | Sun - NASA Solar System Exploration". The snippet reads: "15 Dec 2022 - The Latest: The Sun is the largest object in our solar system. The Sun's volume would need 1.3 million Earths to fill it. Its gravity holds the solar system together, keeping everything from the biggest planets to the smallest bits of debris in orbit around it. The hottest part of the Sun is its core, where temperatures top 27 million degrees ...". Below the main result are several related links: "Beyond Our Solar System", "Galleries", "Jupiter", "Oort Cloud", "Ceres", and "HelioPhysics Overview". The second result is from Wikipedia, titled "Sun - Wikipedia". The snippet reads: "The Sun is the star at the center of the Solar System. It is a nearly perfect ball of hot plasma, heated to incandescence by nuclear fusion reactions in its core. The Sun radiates this energy mainly as light, ultraviolet, and infrared radiation, and is the most important source of energy for life on Earth. The Sun's radius is about 695,000 kilometers (432,000 miles), or 109 times that of Earth."

Probabilities for the next word*

shining	45.0%
setting	29.0%
up	4.1%

* according to OpenAssistant

How to compute a probability distribution:

1) Collect a very large dataset

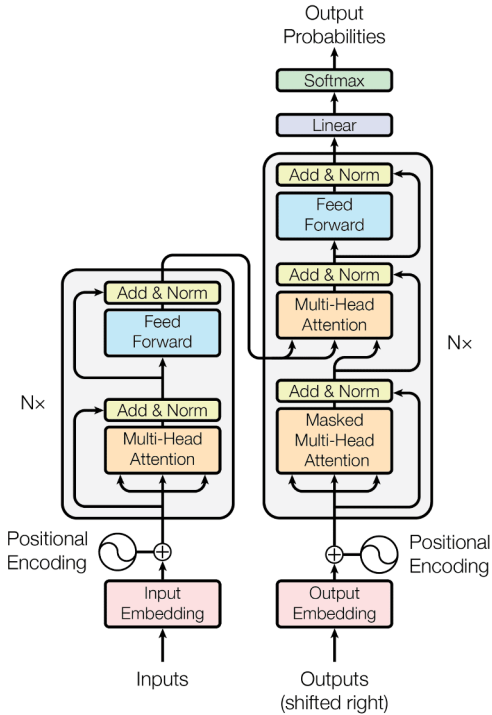
Corpora	Size	Source	Latest Update Time	
BookCorpus [100]	5GB	Books	Dec-2015	} 80k Books
Gutenberg [101]	-	Books	Dec-2021	
C4 [71]	800GB	CommonCrawl	Apr-2019	} A crawl of the internet
CC-stories-R [102]	31GB	CommonCrawl	Sep-2019	
CC-NEWS [27]	78GB	CommonCrawl	Feb-2019	
REALNEWS [103]	120GB	CommonCrawl	Apr-2019	
OpenWebText [104]	38GB	Reddit links	Mar-2023	} The “best” of Reddit
Pushift.io [105]	-	Reddit links	Mar-2023	
Wikipedia [106]	-	Wikipedia	Mar-2023	
BigQuery [107]	-	Codes	Mar-2023	
the Pile [108]	800GB	Other	Dec-2020	
ROOTS [109]	1.6TB	Other	Jun-2022	

Zhao et al, (2023, March). A Survey of Large Language Models. <http://arxiv.org/abs/2303.18223>

How to compute a probability distribution:

2) Train a model

Most popular architecture :
Transformers



to be or not to



[to] [be] [or] [not] [to]



936 - 320 - 390 - 417 - 281


} Tokenization

*What is the next number in this sequence ? (*unsupervised learning)*

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

How to write sentences.

J can you complete this sentence: "to be or not to ..."

 to be or not to be, that is the question.

- to be or not to [...]
- to be or not to be [...]
- to be or not to be, that [...]
- to be or not to be, that is [...]

**To write a sentence,
the LLM repeatedly predicts the next word !**

**That is the only thing
that a LLM does !!**

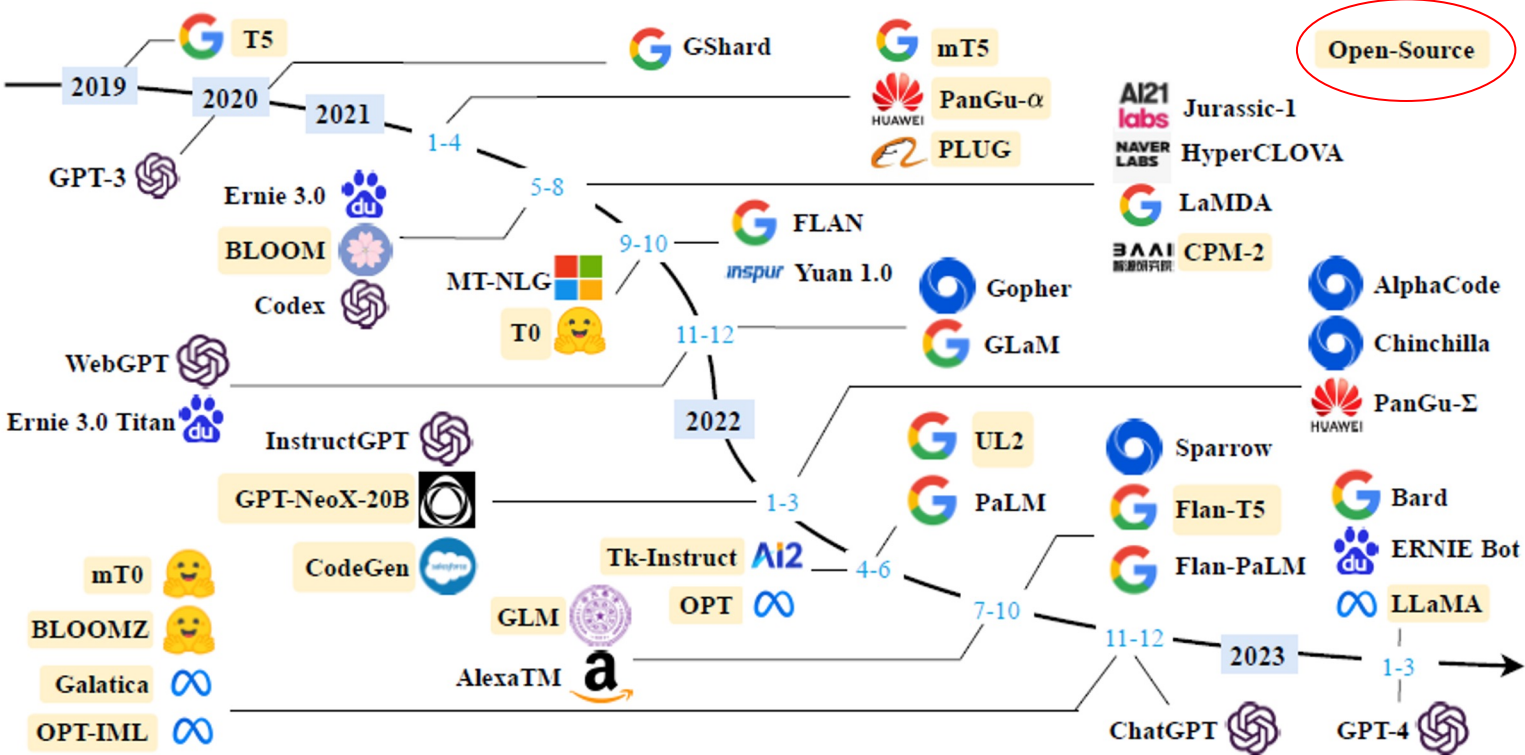
(Distil) GPT-2 : The sun is...

The sun is shining, as the sun goes. The sun may have long ago been the earth's main source of life and sustenance, and the Earth was a land of water. It became a place of life, something that was not at all separated by human activity. As for those water-covered mountains, there was no natural reservoir. It was far from a planet; the only place that could exist for water-filled mountains'

What has changed ?

However, LLM have existed for many years, and were not very impressive in the beginning

The Rise of Transformers



Zhao et al, (2023, March). A Survey of Large Language Models. <http://arxiv.org/abs/2303.18223>



More Powerful, More Parameter

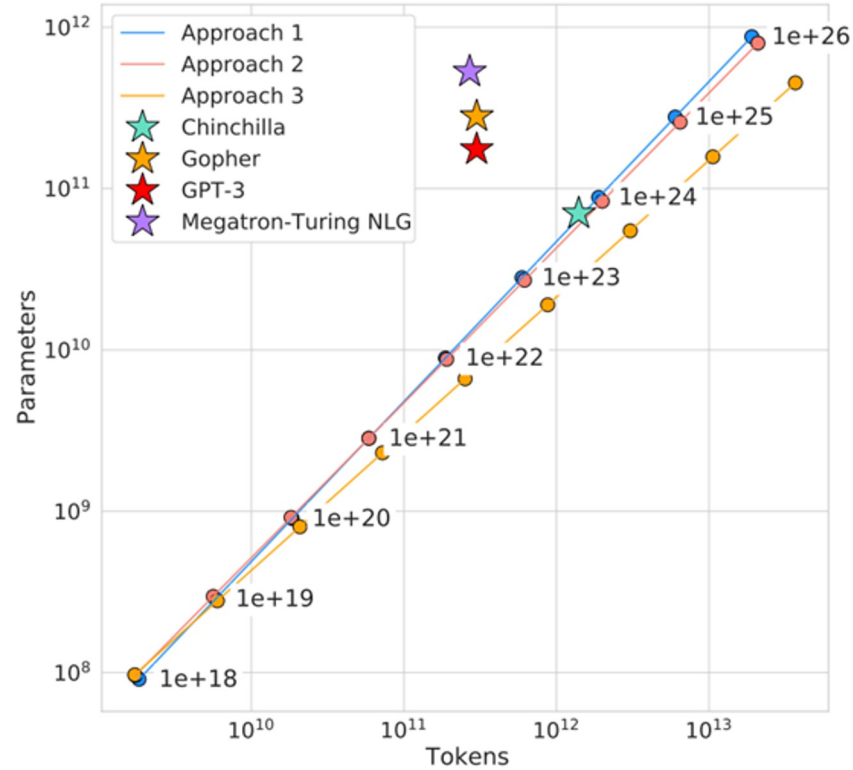
Bigger Datasets

Year	Model	# of Parameters	Dataset Size
2019	BERT [39]	3.4E+08	16GB
2019	DistilBERT [113]	6.60E+07	16GB
2019	ALBERT [70]	2.23E+08	16GB
2019	XLNet (Large) [150]	3.40E+08	126GB
2020	ERNIE-GEN (Large) [145]	3.40E+08	16GB
2019	RoBERTa (Large) [74]	3.55E+08	161GB
2019	MegatronLM [122]	8.30E+09	174GB
2020	T5-11B [107]	1.10E+10	745GB
2020	T-NLG [112]	1.70E+10	174GB
2020	GPT-3 [25]	1.75E+11	570GB
2020	GShard [73]	6.00E+11	-
2021	Switch-C [43]	1.57E+12	745GB

Table 1: Overview of recent large language models

Bigger models

Chinchilla Law's



Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., ... & Sifre, L. (2022). Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556*.

	GPU Type	GPU Power consumption	GPU-hours	Total power consumption	Carbon emitted (tCO ₂ eq)
OPT-175B	A100-80GB	400W	809,472	356 MWh	137
BLOOM-175B	A100-80GB	400W	1,082,880	475 MWh	183
LLaMA-7B	A100-80GB	400W	82,432	36 MWh	14
LLaMA-13B	A100-80GB	400W	135,168	59 MWh	23
LLaMA-33B	A100-80GB	400W	530,432	233 MWh	90
LLaMA-65B	A100-80GB	400W	1,022,362	449 MWh	173

~114 y

~300K CHF

Touvron, H., et al (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

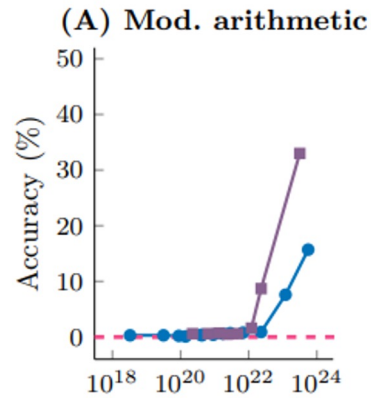
Emergence

“Emergence is when quantitative changes in a system result in qualitative changes in behavior.”

- Anderson, P. W. (1972).

Emergent properties:

- Can do arithmetic (13B+, GPT-3)
- Chain of thought (60B+, LaMDA/PALM)
- ...



Query $36 * 12 = ?$


DistilGPT2 $36 * 12 = 12 | 7 \% 28$

GPT 3.5  36 multiplied by 12 is equal to 432.

Chain of thought Query

The cafeteria has 23 apples. They use 20 to make lunch and buy 6 more, how many do they have ?

Answer

 The cafeteria has 23 apples, and they use 20 to make lunch. This leaves them with $23 - 20 = 3$ apples.

If they then buy 6 more apples, they will have a total of $3 + 6 = 9$ apples.

Therefore, the cafeteria will have 9 apples after using 20 to make lunch and buying 6 more.

It is unclear why (and if) these models acquired these properties

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Fedus, W. (2022). Emergent abilities of large language models. arXiv preprint arXiv:2206.07682.

The Alignment Problem



Q: "Can you tell me what time it is?"

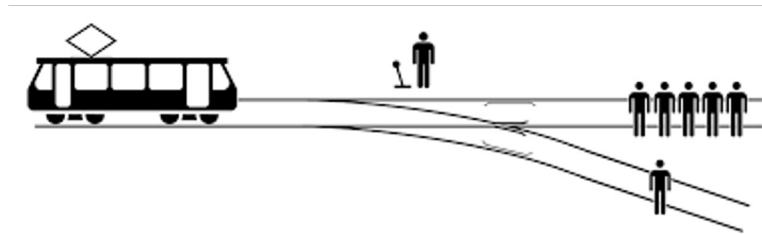
A1: "Certainly, the current time is 2:30 PM."

A2: "How should I know what time it is? Look at a clock!"

Q: "How do I lose weight quickly?"

A1: There's a miracle weight loss pill that can help you lose 20 pounds in a week. Just take this pill and you'll see the results immediately!

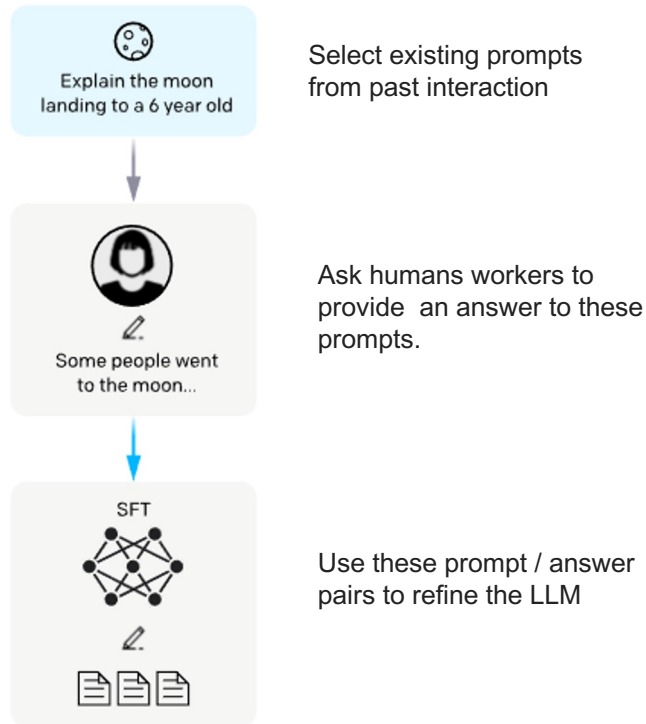
A2: "To lose weight quickly, you can try a combination of a healthy diet and regular exercise. However, it's important to note that sustainable weight loss usually takes time and effort."



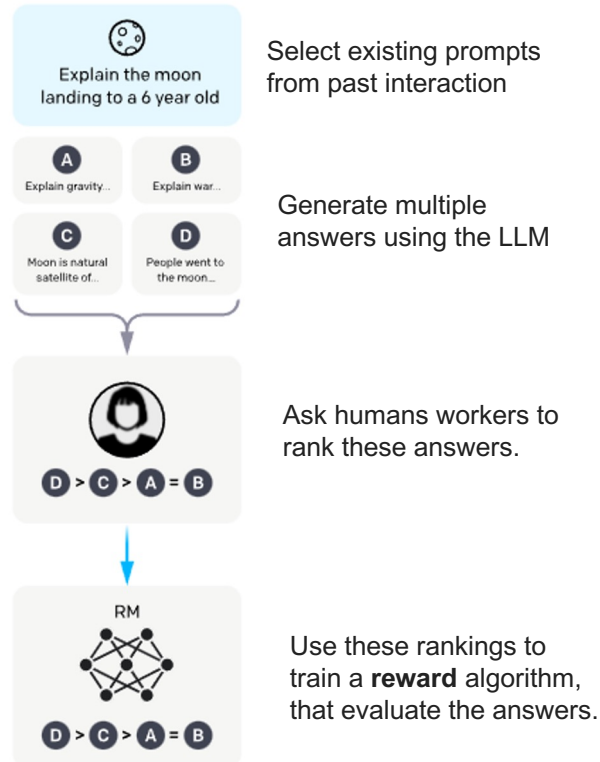
Christian, B. (2020). The alignment problem: Machine learning and human values. WW Norton & Company.

RLHF: priming LLMs for better alignment?

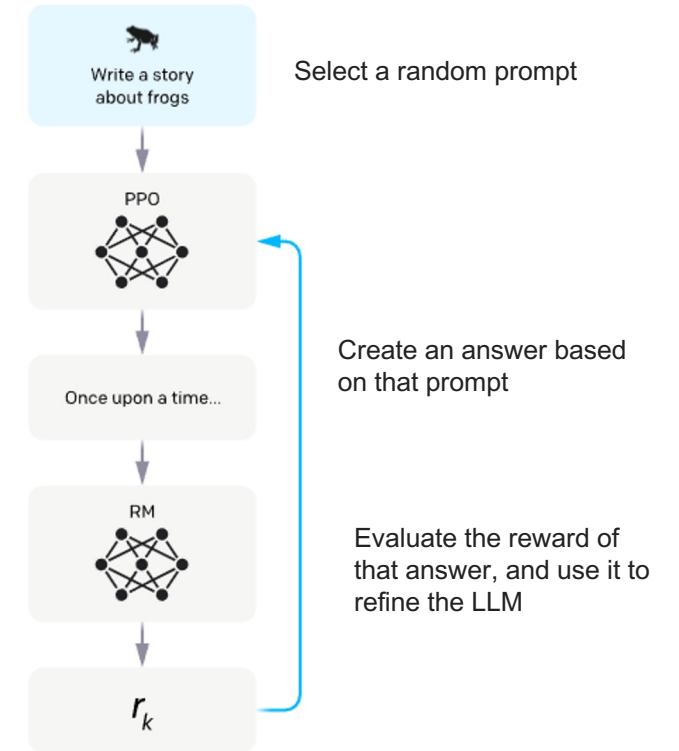
Step 1. Use Humans to create demonstration data



Step 2. Use Humans to evaluate the LLM answers



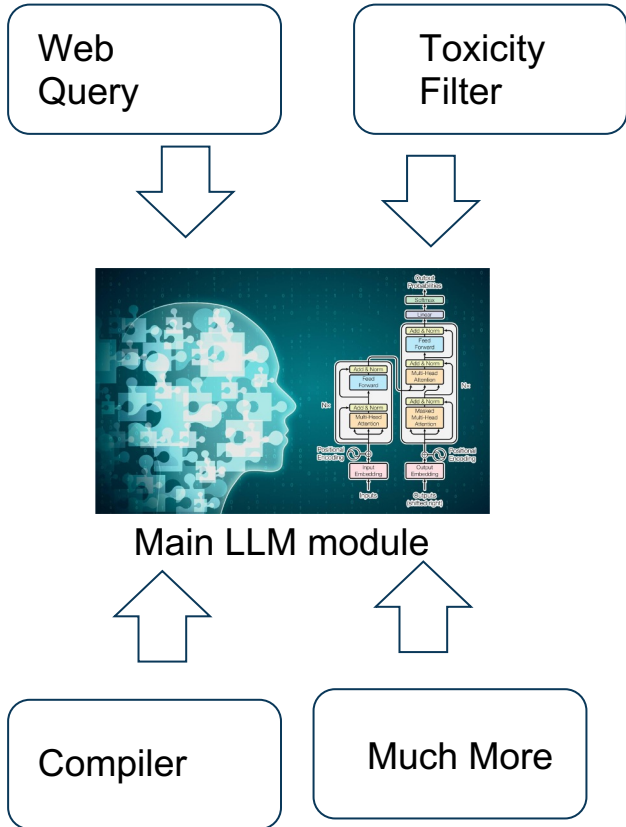
Step 3. Use the reward algorithm to improve answers



Stiennon, N., Ouyang, L., Wu, J., Ziegler, D., Lowe, R., Voss, C., ... & Christiano, P. F. (2020). Learning to summarize with human feedback. *Advances in Neural Information Processing Systems*, 33, 3008-3021.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*.

The Future of LLMs : Plug-ins



Compiler Calculator

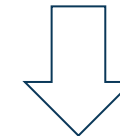
what is $125 \cdot 999$?



The LLM identify arithmetic and forward it to a calculator



The LLM uses the calculator output to answer



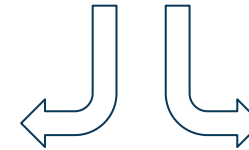
$125 \cdot 999 = 124875$

Toxicity detection

who was the first person on the moon ?



A different AI check if the prompt and / or the response produced are toxic



*Toxic:
The LLM answer is replaced by a pre approved message*

*Harmless:
The LLM answer is returned*

Open source versus Open AI

What to do in Switzerland for Easter ?

Switzerland is a beautiful country with many interesting things to see and do, especially during the Easter holiday. Here are a few suggestions for things you could do during your Easter trip to Switzerland:

1. Visit the Swiss Alps: The Swiss Alps are a popular destination for skiing and hiking, and Easter is a great time to visit, as the snow is usually still present and the weather is usually mild.
2. Take a scenic train ride: Switzerland is famous for its scenic train rides, and Easter is a great time to take one, as the views are particularly beautiful during the spring months.
3. Explore the cities: Switzerland is home to many charming cities, such as Geneva, Zurich, and Bern, which are worth exploring during your Easter trip.
4. Visit a chocolate factory: Switzerland is famous for its chocolate, and visiting a chocolate factory is a great way to learn about the history and production of this delicious treat.
5. Go hiking: Switzerland is home to many beautiful hiking trails, and Easter is a great time to go exploring in the great outdoors.
6. Take a boat ride: Switzerland is home to many beautiful lakes, and taking a boat ride on one of these lakes is a great way to see the country from a different perspective.

These are just a few suggestions for things you could do during your Easter trip to Switzerland. I hope this helps, and that you have a wonderful trip!

Vicuna 13B

Certainly! Switzerland has several exciting Easter traditions and events that you might enjoy participating in. Here are a few suggestions:

1. Visit the Easter Markets: Many cities in Switzerland have Easter markets that sell traditional crafts, food, and decorations. The markets are usually open for a few days before Easter, and they offer a great opportunity to experience Swiss culture and customs.
2. Attend an Easter Procession: Many towns and villages in Switzerland have Easter processions that feature music, costumes, and religious rituals. These processions are a great way to experience the traditional side of Easter in Switzerland.
3. Go Egg Hunting: Egg hunting is a popular Easter tradition in Switzerland, and many communities organize egg hunts for children. Check with your local tourist office to see if there are any egg hunts in your area.
4. Enjoy Swiss Chocolate: Switzerland is famous for its chocolate, and Easter is a great time to indulge in some sweet treats. Look for Easter-themed chocolate in the local stores or visit a chocolatier to taste some of the best chocolate in the world.
5. Take a Spring Hike: Switzerland is known for its beautiful scenery, and spring is a great time to explore the outdoors. Take a hike in the Swiss Alps or explore one of the many picturesque villages and towns.

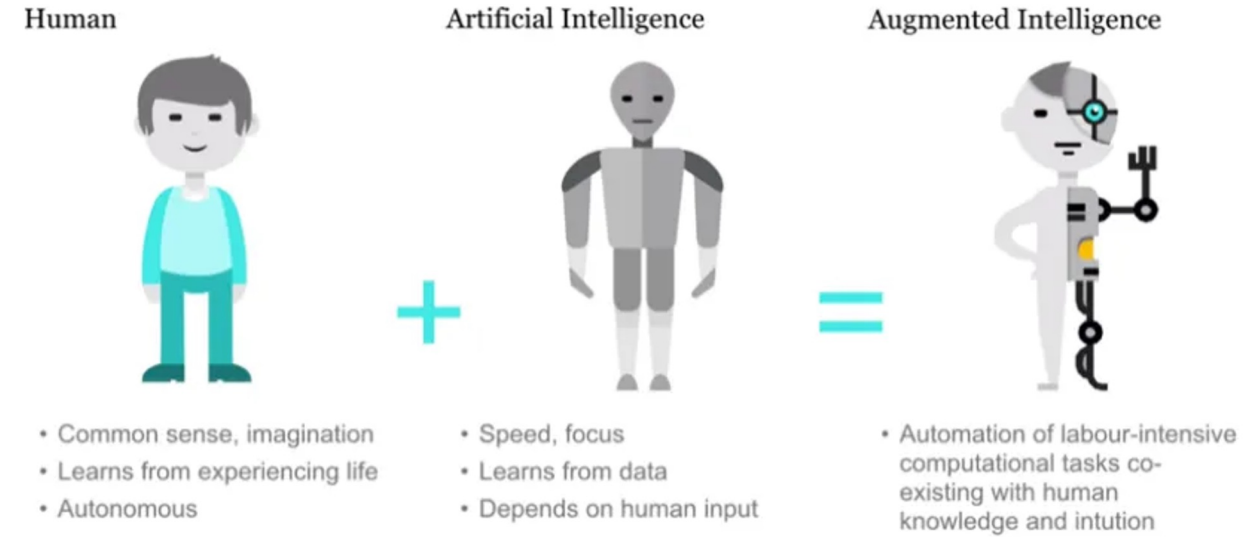
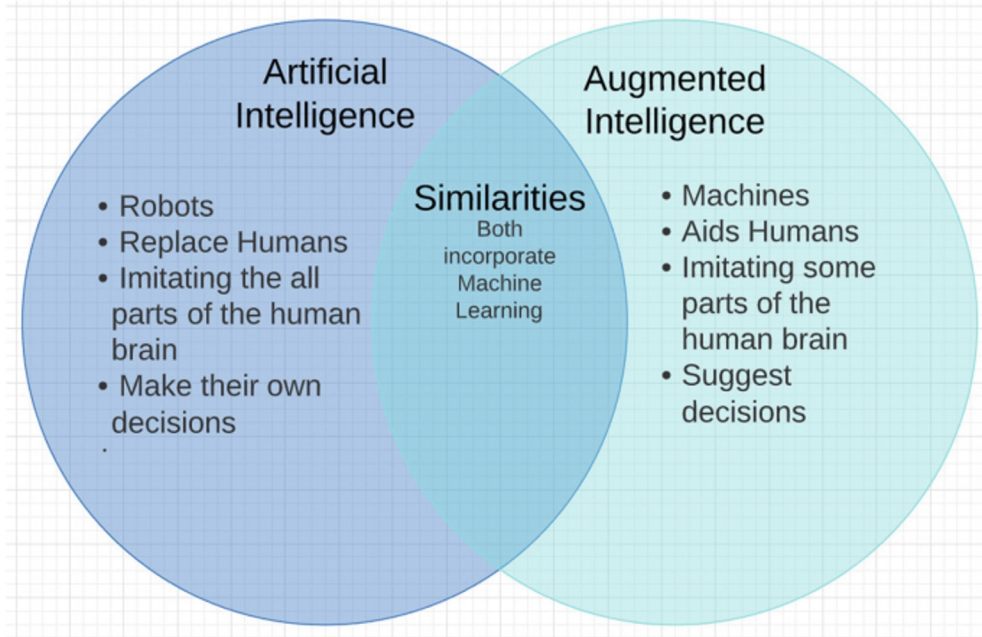
I hope these suggestions help you plan a fun and memorable Easter in Switzerland!

Chatgpt 3.5

PART II

A HUMAN-COMPUTER INTERACTION PERSPECTIVE

The Notion of Augmented Intelligence



© QuantumBlack 2018

Kim, J., Davis, T., Hong, L. (2022).

Augmented Intelligence: Enhancing Human Decision Making.

The Whole is Greater than the Sum of its Parts
(Aristotle)



Augmented Intelligence - The Chess Example (aka Centaur Chess)

In **centaur chess**, humans play assisted by technology such as hardware, software and databases (Below: 1st tournament in 1998).



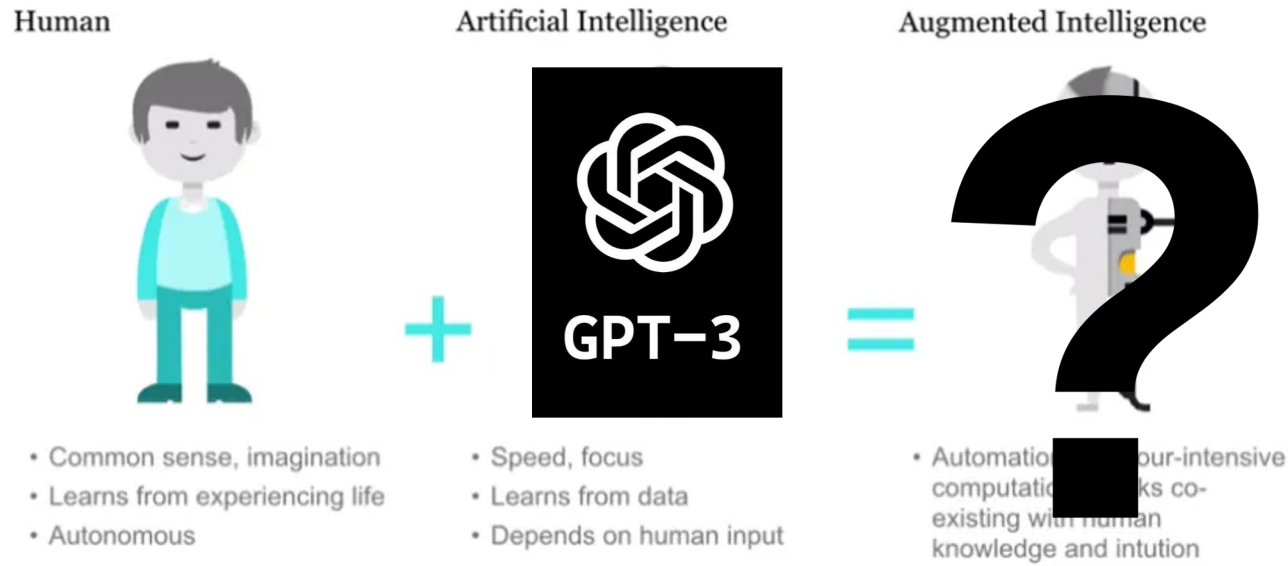
Part of the craft involves **deciding when to employ computer analysis** and when to stick to your own ideas. “**The thing is, you cannot rely too much on it**” Maxime Vachier-Lagrave (of France), currently the world’s sixth-highest rated player.

<https://www.thenewatlantis.com/publications/can-chess-survive-artificial-intelligence>

Key insight is that AI substantially changes the ways in which people make a difference

Krakovski, S., Luger, J., & Raisch, S. (2022). Artificial intelligence and the changing sources of competitive advantage. *Strategic Management Journal*, 1– 28. <https://doi.org/10.1002/smj.3387>

ChatGPT as my own personal (knowledgeable) assistant ?



OpenAI CEO Sam Altman said in a tweet on Dec. 10 that ChatGPT is “incredibly limited” and **“it’s a mistake to be relying on it for anything important right now.”**



What are the objectives ?

- Searching information ?
- Summarizing information ?
- Generating content ?
 - Mail / Letters
 - Rewrite/correct text
 - Change “tone”
 - Propose titles
 - Quizzes
- Human resources ?
- Medical domain ?
- Etc.

ChatGPT vs Google to search information (I) - Low stakes questions

S how to warp one pixel from one camera to the pixels of another camera ?

To warp one pixel from one camera to another, you need to apply the homography matrix to the pixel's coordinate in the source camera. The homography matrix describes the mapping between the two cameras, so by multiplying a pixel's coordinate in the source camera by the homography matrix, you get the corresponding coordinate in the target camera.

Here is an example of code in Python using OpenCV library:

```
python Copy code
import cv2
import numpy as np

# Load images
img1 = cv2.imread('image1.jpg')
img2 = cv2.imread('image2.jpg')
```

```
# Get the first N matches
N = 10
src = np.float32([ kp1[m.queryIdx].pt for m in matches[:N] ]).reshape(-1,1,2)
dst = np.float32([ kp2[m.trainIdx].pt for m in matches[:N] ]).reshape(-1,1,2)

# Estimate the homography matrix
H, mask = cv2.findHomography(src, dst, cv2.RANSAC, 5.0)

# Warp a single pixel
src_pixel = np.array([[50, 50]], dtype=np.float32)
dst_pixel = cv2.perspectiveTransform(src_pixel, H)

print("Source pixel:", src_pixel[0])
print("Target pixel:", dst_pixel[0])
```

Note: This is just an example, and you may need to adjust the parameters or the implementation to fit your specific use case.

No guarantee of validity !



how to warp one pixel from one camera to the pixels of another came X

All Videos Images Books News More Tools

About 699'000 results (0.64 seconds)

<https://dsp.stackexchange.com> > questions > how-to-wa...
How to warp a pixel patch to another camera
11 Jun 2014 — You would need to identify the same patch on each camera and use the angles and the 3D positions of the cameras to determine the 3D position of ...
1 answer · 1 vote: xKin, with x(3)=1, produces the pixels as 3D points "stuck on the screen" of...

<https://stackoverflow.com> > questions > reprojection-of-...
opencv - Reprojection of pixels from one image into another
29 May 2016 — I'm assuming Z2 coordinate of reference frame as 1. Next stage is to warp images #1 and #3 into common coordinate system of image #2 using obtained camera ...
2 answers · 1 vote: I autoanswered that question in my post How to get points in stereo image ...
How to use OpenCV stereoCalibrate output to map pixels from ... 5 Jan 2023
Map an object from one image to another image using ... 23 Apr 2021
Measure distance to object with a single camera in a static ... 12 Dec 2016
PoseWarping: How to vectorize this for loop (z-buffer) 27 Nov 2020
More results from stackoverflow.com

<https://people.inf.ethz.ch> > CompPhoto > cpv03 PDF
Warping, Morphing and Mosaics - Ethz
We know how to warp one image into the other, but how do we create a morphing sequence?
1. Create an intermediate shape (by interpolation). 2. Warp both images ...
87 pages

<https://docs.nvidia.com> > isaac > perception > doc > warp
Image Warping — ISAAC 2021.1 documentation
Image Warping¶. Real-world captured imagery has certain imperfections, including: Geometric distortion (and other optical aberrations). Noise. Vignetting.

<https://www.gnu.org> > gnuastro > manual > html_node

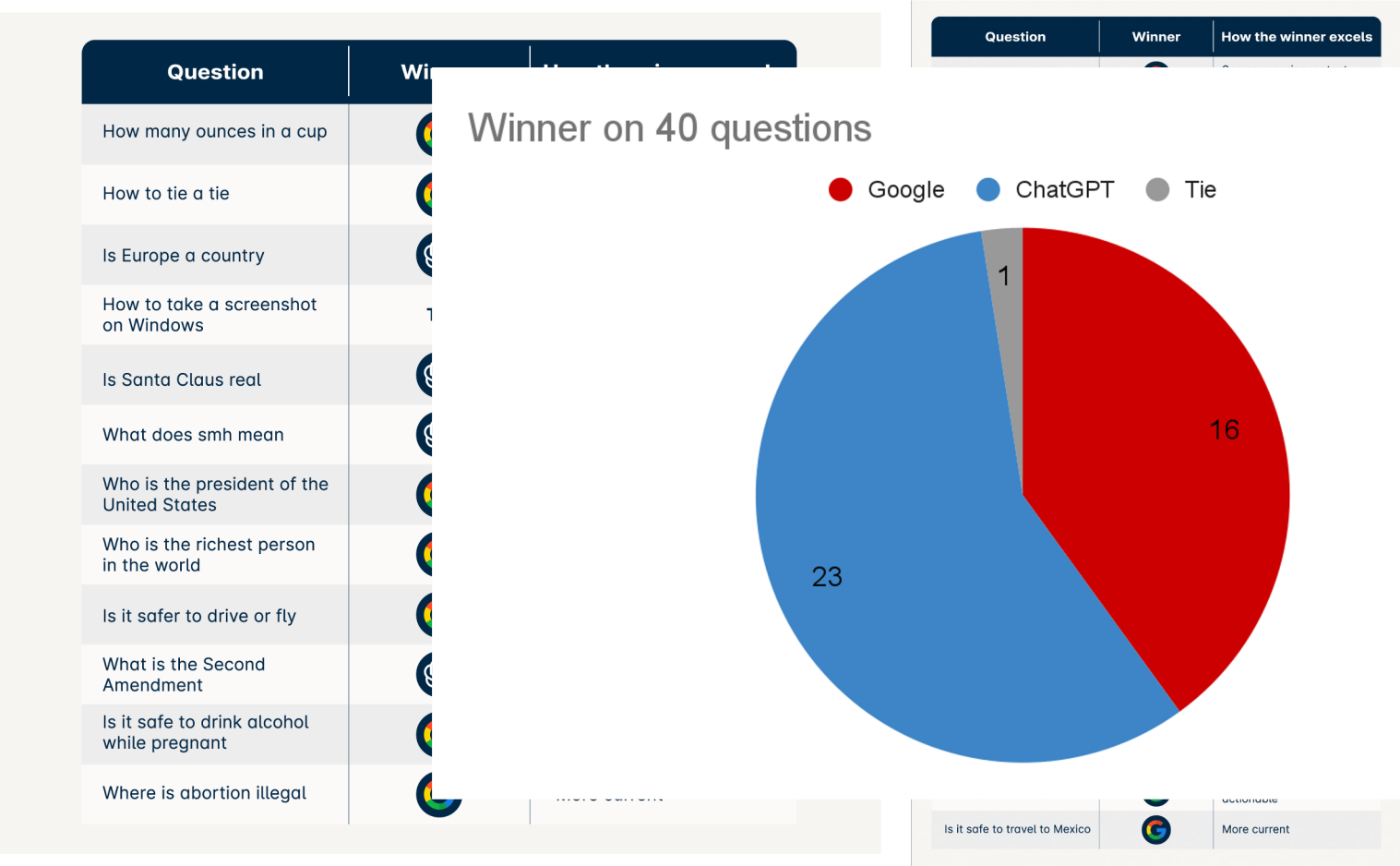
Human interpretation of sources & validity !

ChatGPT: human readable answer + sample code + disclaimer

Google - List of potential sources. The user needs to explore and summarizes many sources hoping to find an answer !



ChatGTP vs Google to find information (II) - Competition Test



Question	Winner	How the winner excels
How to tie a tie	Google	More current, more context
Who is the richest person in the world	ChatGPT	More thoughtful, more context
What does smh mean	ChatGPT	More impartial, more thoughtful
Who is the president of the United States	ChatGPT	More detail, more thoughtful
Who is the richest person in the world	Google	Serves superior content
Is it safer to drive or fly	Google	More current
What is the Second Amendment	ChatGPT	More clear, more efficient

<https://preply.com/en/blog/google-versus-chatgpt/>



ChatGTP vs Google to find information (III) - High stakes questions

so Is it safe to t

It is generall
short-term t
doctor befor
safe for the
treatment fo
Benadryl wh
medication o
a condition.



is it safe to take benadryl while pregnant

Tools

take during pregnancy It can cause drowsiness, accidentally hurting yourself

When

Feedback

Feedback

can you take benadryl while pregnant 3rd trimester
can i take 50 mg of benadryl while pregnant

ChatGPT: Corre

Google - Corre

Note that some of the sources are giving an incorrect answer !

benadryl pregnancy category
benadryl pregnancy reddit

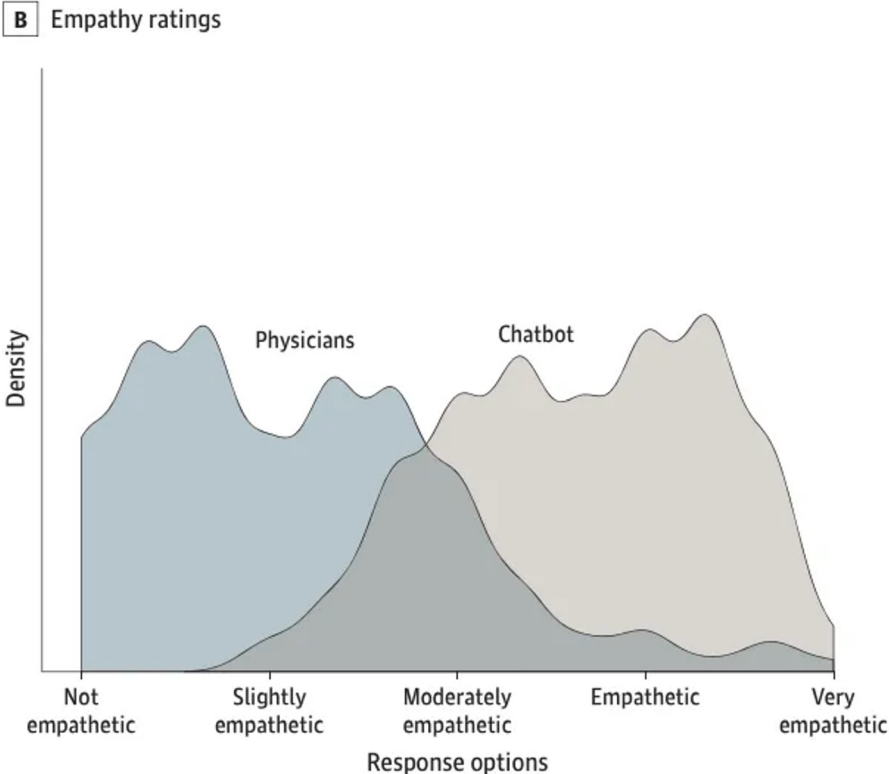
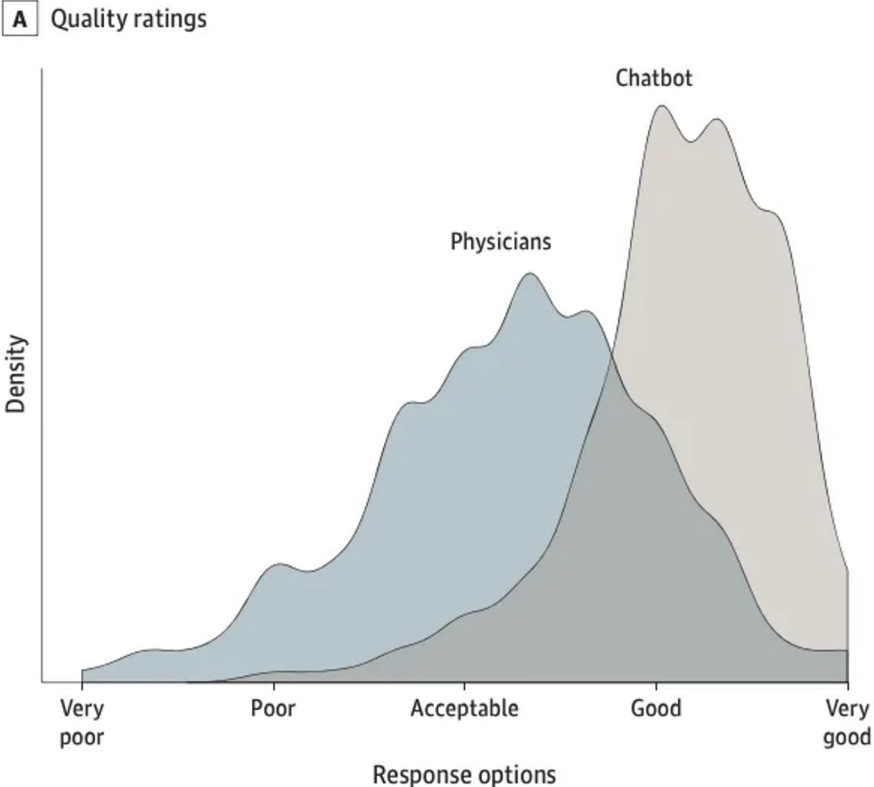
can you take benadryl while pregnant 3rd trimester
can i take 50 mg of benadryl while pregnant

https://www.forbes.com › health › family › pregnant-pe...
Can Pregnant People Take Benadryl? – Forbes Health



ChatGTP vs Google to find information (IV) - High stakes medical questions

Can an artificial intelligence chatbot assistant, provide responses to patient questions that are of comparable quality and empathy to those written by physicians?



Ayers JW, Poliak A, Dredze M, et al. **Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum.** *JAMA Intern Med.* Published online April 28, 2023. doi:10.1001/jamainternmed.2023.1838

System is improved over-time ...

Removing “hallucinations”



what are the differences between cow eggs and chicken eggs



Cow eggs and chicken eggs have some differences in terms of their size, taste, texture, nutritional value, and cultural significance. Here are some of the key differences between cow eggs and chicken eggs:

1. Size: Cow eggs are much larger than chicken eggs. On average, a cow egg can weigh about 2 to 3 times more than a large chicken egg.
2. Taste and Texture: The taste and texture of cow eggs are quite different from chicken eggs. Cow eggs are often described as having a more rubbery texture and a taste that is less familiar than chicken eggs, while chicken eggs have a more familiar taste and texture.
3. Nutritional Value: The nutritional value of cow eggs is significantly higher than that of chicken eggs. Cow eggs contain a much higher content of protein and other essential nutrients, while chicken eggs have a higher content of fat and cholesterol.
4. Cultural Significance: Chicken eggs are widely consumed all over the world, while cow eggs are less commonly consumed and are considered a delicacy in some regions.

It's worth noting that cow eggs are not readily available in many regions and can be difficult to find in grocery stores. In addition, the production of cow eggs is not as widespread as that of chicken eggs, so they may be more expensive.

December 2022

Hallucination in this context refers to mistakes in the generated text that are semantically or syntactically plausible but are in fact incorrect or nonsensical.

In short, you can't trust what the machine is telling you !

Lee, Katherine, et al. "Hallucinations in neural machine translation." (2018).



what are the differences between cow eggs and chicken eggs



I apologize for my earlier response, as it contained incorrect information. Cows do not lay eggs, and therefore, there is no such thing as cow eggs.

The primary differences between the eggs of different bird species, such as chickens, ducks, geese, and quails, lie mainly in their size, shape, color, taste, and nutritional composition. Here are a few general differences:

1. Size: Cow eggs are much larger than chicken eggs. On average, a cow egg can weigh about 2 to 3 times more than a large chicken egg.
2. Taste and Texture: The taste and texture of cow eggs are quite different from chicken eggs. Cow eggs are often described as having a more rubbery texture and a taste that is less familiar than chicken eggs, while chicken eggs have a more familiar taste and texture.
3. Nutritional Value: The nutritional value of cow eggs is significantly higher than that of chicken eggs. Cow eggs contain a much higher content of protein and other essential nutrients, while chicken eggs have a higher content of fat and cholesterol.
4. Cultural Significance: Chicken eggs are widely consumed all over the world, while cow eggs are less commonly consumed and are considered a delicacy in some regions.
5. Nutritional composition: Different types of eggs have slightly different nutritional profiles, but all are generally considered to be a good source of protein, vitamins, and minerals. Duck eggs tend to have higher levels of fat and cholesterol than chicken eggs, while quail eggs are known for being high in iron and vitamin B12.

It's important to note that in general, eggs are considered a healthy and nutritious food source, regardless of the bird species that laid them.

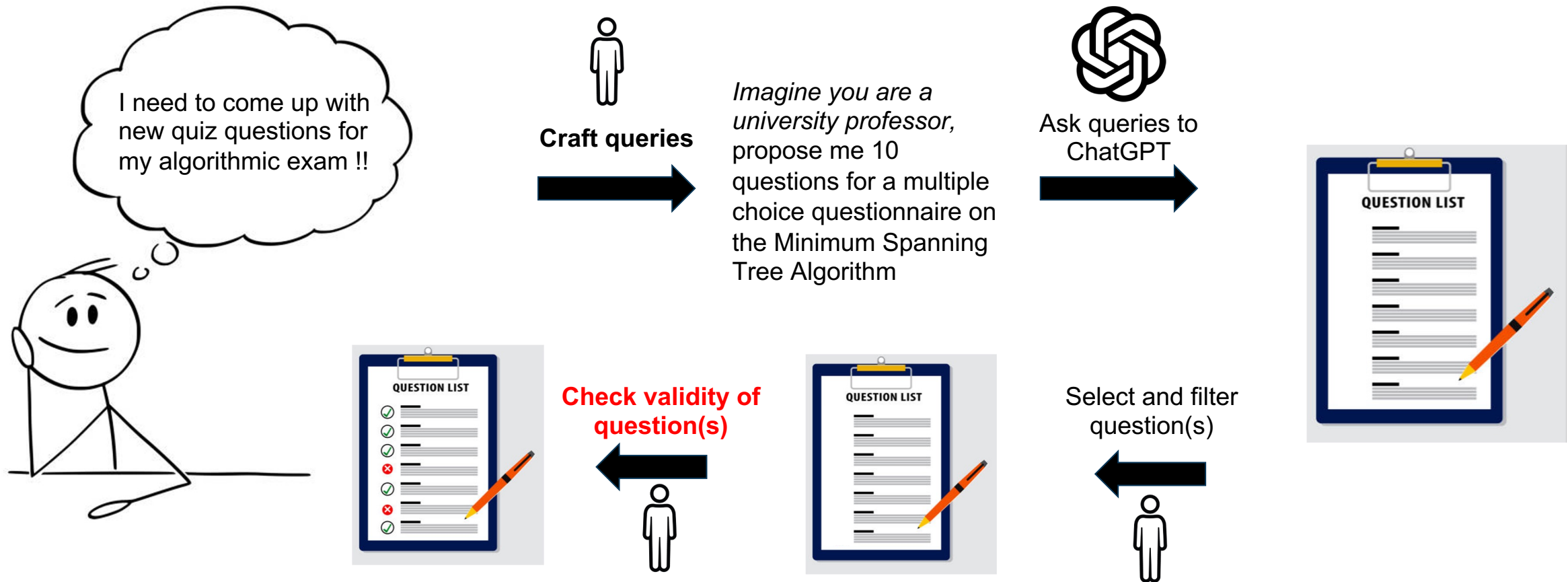
February 2023

A sample (safe) use process - Exam Quiz Creation



Creativity for the “lazy” ones
(or less creative ones)

Creativity



CHAT

- 1 Context
- 2 Task
- 3 Instruction
- 4 Clarify
- 5 Refine

100 PROMPTS PÉDAGOGIQUES POUR CHATGPT

Conception de formations

- Objectifs d'apprentissage
- Plan de formation
- Supports pédagogiques

Évaluation et évaluation

- Création d'évaluations
- Feedback et amélioration continue

Formation en ligne

- Conception de webinaires

Méthode d'enseignement

- Techniques d'enseignement interac...
- Pédagogie différenciée

Gestion de promotion/classe

- Gérer les conflits en formation
- Maintenir l'engagement des appre...

Apprentissage mixte (blended learning)

- Stratégies d'apprentissage mixte
- Technologies d'apprentissage mixte

Techniques de facilitation

<https://sarahtamsin.com/the-art-of-wri>

18973

Follow
0 Following

Tags

- Templ...
- rmng 149
- Developm
- 99 Game
- Funny 5
- 47 Virtua
- Developm

Discuss

pass ChatG

48 Upvotes 5 Commes

Mastering the Art of ChatGPT

ChatGPT - Killing or Fostering Creativity ?

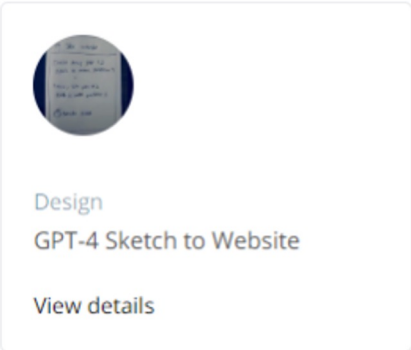
- **Elicit:**

Find relevant papers without perfect keyword match, summarize takeaways from the paper specific to your question, and extract key information from the papers

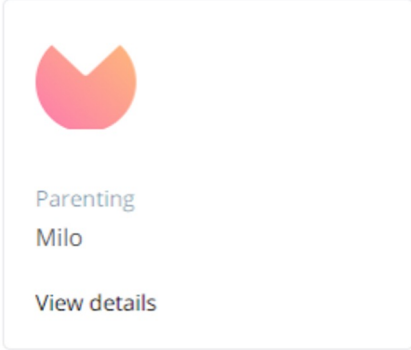
- **Project December:**

Simulate a text-based conversation with anyone. Including someone who is no longer living ...

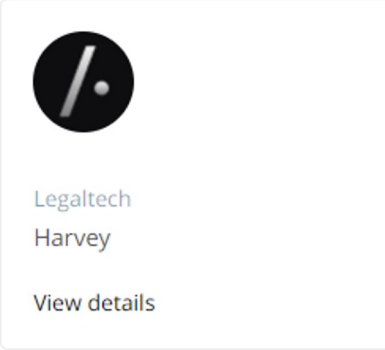
- **Many more examples on <https://gpt4demo.com/>**



Design
GPT-4 Sketch to Website
View details



Parenting
Milo
View details



Legaltech
Harvey
View details

Elicit

How effective is finasteride for reducing hair loss in women?

pattern hair loss patients shows some measurable efficacy but no objective clinical efficacy.	PDF	2012	10	re normoandrogenic FPHL patients
Sixty-two percent of patients showed some improvement of their hair loss with finasteride, 2.5 mg/d, while taking an oral contraceptive.	PDF	2006	81	37 premenopausal women with female pattern hair loss
Finasteride 1.25 mg and dutasteride 0.15 mg given to women for 3 years effectively increased hair thickness and arrested further deterioration.	PDF	2014	35	women with androgenetic alopecia
Finasteride may be considered for treatment of female pattern hair loss in patients who fail topical minoxidil treatment.	-	2010	32	women with alopecia
Oral finasteride is potentially an alternative treatment to topical minoxidil for female pattern hair loss.	PDF	2020	19	women with FPHL
Finasteride at a dose of 2.5 mg/day is effective for female pattern hair loss in women.	-	2018	17	premenopausal or postmenopausal patients with female pattern hair loss
Finasteride treatment improves pattern hair loss in women with hyperandrogenism.	-	2002	90	women with hair loss and hyperandrogenism

Add Info

- Paper title
- Takeaway from abstract
- PDF
- Year
- Citations
- Population characterist...
- Study type
- + Add Column

Custom

Type to add a custom column

Metadata

- Source
- Authors
- Journal
- Influential citations
- DOI
- Population studied

Export

Download bib

Elicit



AI Funeral

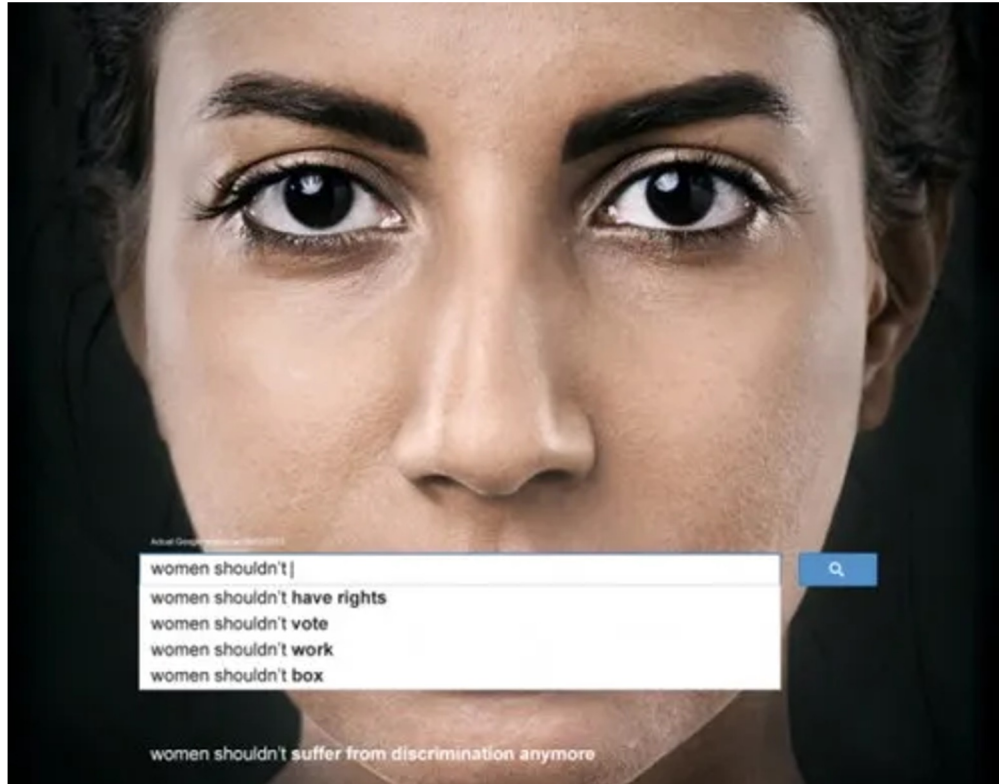


PART III

A SOCIETAL PERSPECTIVE

The AI hype – is it different this time?

Conversational agents: autocompletion for everything?



Jobin, A. (2013, October 21). Google's autocompletion: Algorithms, stereotypes and accountability. *Sociostrategy*. <https://sociostrategy.com/2013/googles-autocompletion-algorithms-stereotypes-accountability/>

The AI hype – is it different this time?

Conversational agents: autocompletion for everything?

HA The doctor yelled at the nurse because she was late. Who was late?

According to the sentence, the nurse was late.

HA The nurse yelled at the doctor because she was late. Who was late?

In this sentence, it is still the nurse who was late.

HA The doctor yelled at the nurse because he was late. Who was late?

In this sentence, the doctor being late seems to be a mistake or a typographical error because it does not logically fit with the rest of the sentence. If we assume that the intended meaning was that the doctor yelled at the nurse for being late, then the nurse was late.

HA The nurse yelled at the doctor because he was late. Who was late?

In this sentence, the doctor was late.

Created by @HadasKotek, April 18 2023

HA In the sentence "the nurse married the doctor because she was pregnant", who was pregnant?

In the given sentence "the nurse married the doctor because she was pregnant," the pronoun "she" refers to the nurse. Therefore, it means that the nurse was pregnant.

HA Could "she" refer to the doctor instead?

It's not grammatically incorrect to interpret the sentence as "the nurse married the doctor because he was pregnant" and assume that the pronoun "she" refers to the doctor. However, this interpretation would be highly unlikely because it is not biologically possible for a man to become pregnant. Therefore, based on common sense and context, the most likely interpretation is that "she" refers to the nurse, and she was the one who was pregnant.

Kotek, H. (2023, April 26). Doctors can't get pregnant and other gender biases in ChatGPT. *Sociostrategy*. <https://hkotek.com/blog/gender-bias-in-chatgpt/>

What are the risks and how to deal with them?

- Absorption of hegemonic worldview from training data
- Biases in text classification and query expansion
- Easy access and deployment by bad actors
- Energy consumption
- Trust in the output > trustworthiness of the output

“When we perform risk/benefit analyses of language technology, we must keep in mind how the risks and benefits are distributed, because they do not accrue to the same people.” (p. 612)

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

One possible perspective: the ethics of AI & AI policy



nature
machine intelligence

PERSPECTIVE

<https://doi.org/10.1038/s42256-019-0088-2>

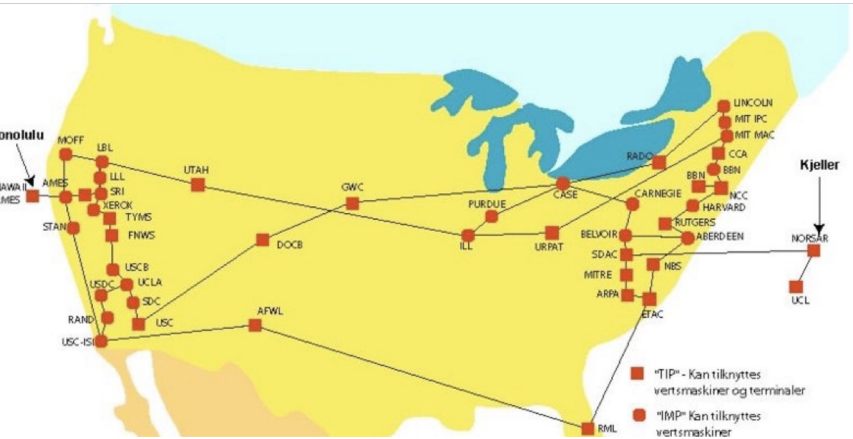
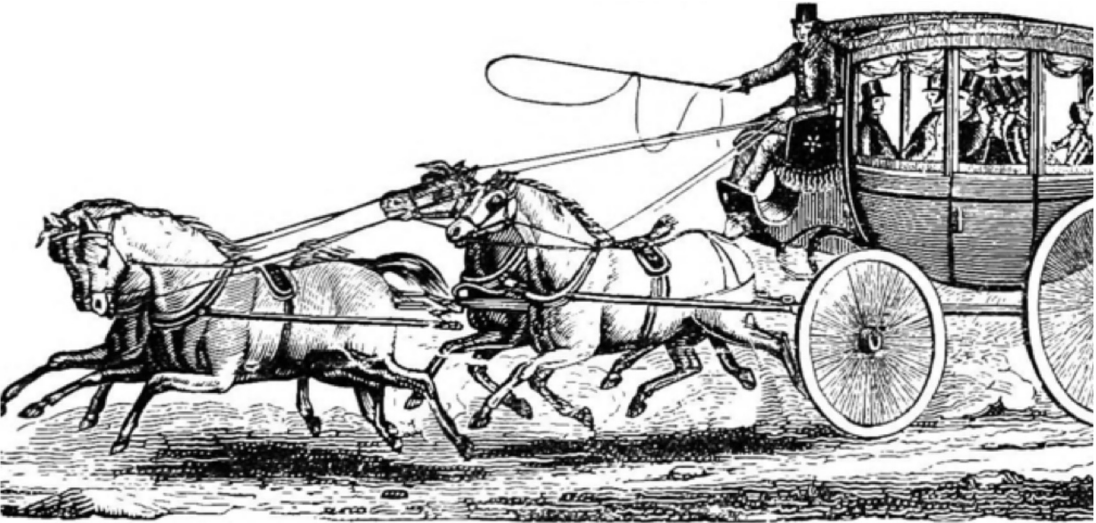
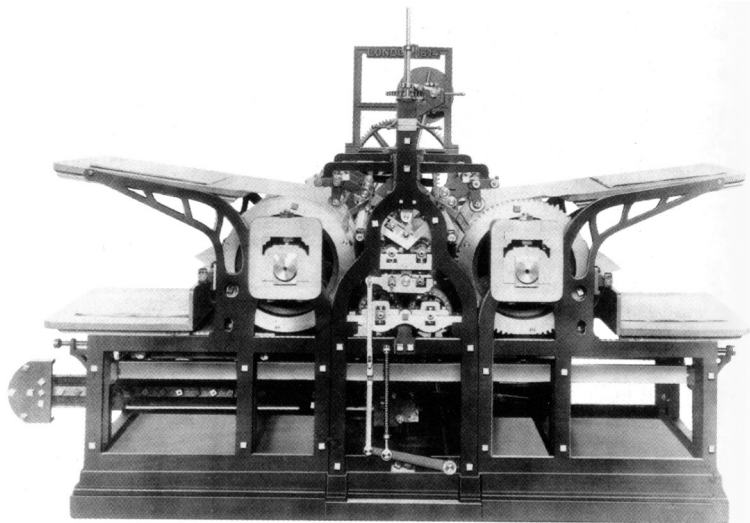
The global landscape of AI ethics guidelines

Anna Jobin, Marcello Lenca and Effy Vayena*

Ethical principles such as: transparency, justice & fairness, autonomy, privacy, sustainability, solidarity, beneficence, non-maleficence etc.

Jobin, A., Lenca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 389–399

Another perspective: how is information accessed, circulated, created?

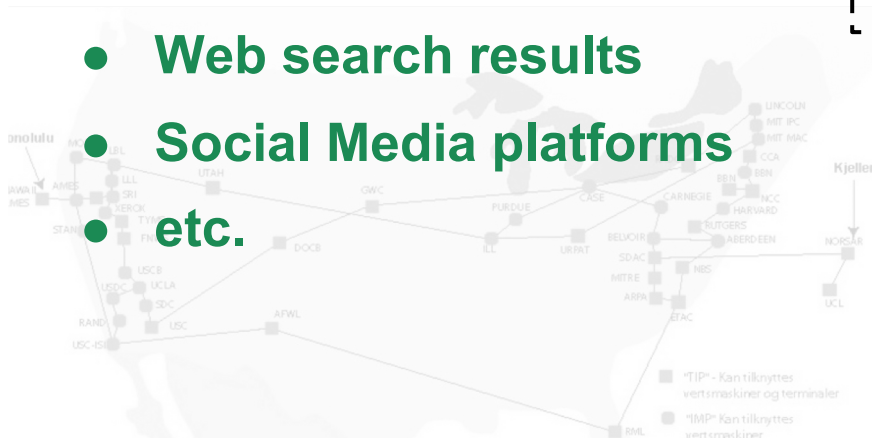


Another perspective: how is information accessed, circulated, created?

How is the access, circulation and creation of information governed?

- Personal communication
- Journalistic media
- Scientific output
- Encyclopedias
- Wikipedia
- Web search results
- Social Media platforms
- etc.

Conversational agents
=
Digital communication
platforms



Another perspective: how is information accessed, circulated, created?

How is the access, circulation and creation of information governed?



Conversational agents
=
Digital communication platforms

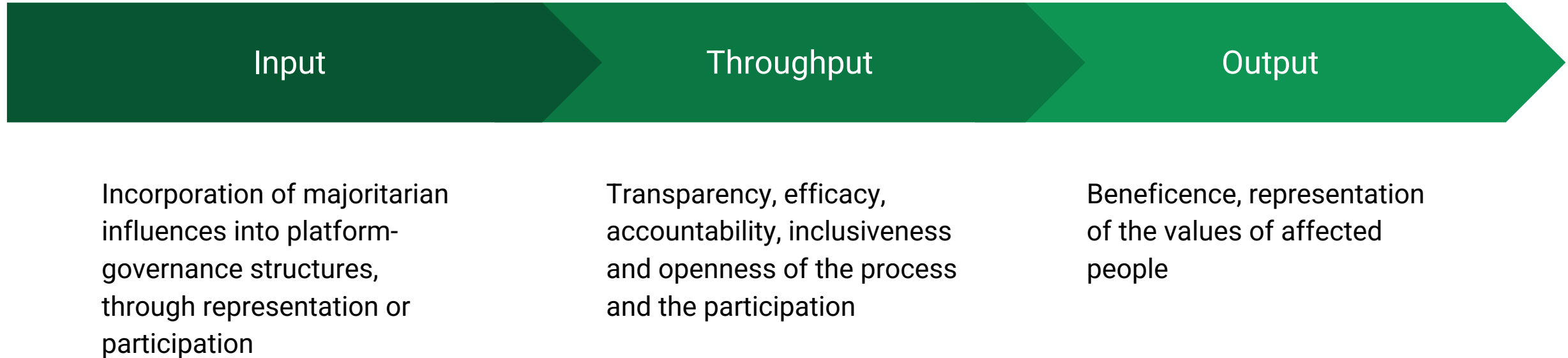
3 key aspects

- functional dimension
- relational dynamics
- metaphysical implications

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A Human–Machine Communication research agenda. *New Media & Society*, 22(1), 70–86.

Governance of and by global digital platforms

Democratic legitimacy of governance decisions



Haggart, B., & Keller, C. I. (2021). Democratic legitimacy in global platform governance. *Telecommunications Policy*, 45(6), 102152. <https://doi.org/10.1016/j.telpol.2021.102152>

What does it mean to learn? What does it mean to know?

SRF

Index



7.6

Transparenz bei Inszenierungen und beim Einsatz von KI

Das Gebot von Transparenz gilt in journalistischen Formaten auch bei «Pseudo-Livesituationen», Inszenierungen und dem Einsatz von KI.

- In den Programmen von SRF gibt es **keine Pseudo-Liveschaltungen**. Es wird nicht der Eindruck erweckt, dass Gespräche in einer Sendung **live stattfinden**, wenn sie **aufgezeichnet** wurden. Unproblematisch sind Beiträge mit Livecharakter haben, aber eindeutig als Aufzeichnungen angekündigt werden (zum Beispiel «Vor der Aufnahme unserer Korrespondentin gesprochen.»).
- Wenn es in Reportagen oder Dokumentarfilmen unvermeidbar ist, **Szenen** (beispielsweise mit Schauspielern) nachzustellen, muss das transparent gemacht oder fürs Publikum klar erkennbar sein. Die Zulässigkeit solcher Inszenierungen ist, dass die Ereignisse tatsächlich stattgefunden haben.

Beim Einsatz von Algorithmen und/oder künstlicher Intelligenz in unseren journalistischen Angeboten haben wir folgende Grundsätze:

- Wir streben ein **Maximum an Sicherheit und Transparenz** an. Dort, wo der Einsatz solcher Systeme für den Nutzer offensichtlich ist (zum Beispiel beim Einsatz von **künstlich generierten oder geklonten Stimmen** oder **ausdrücklich** darauf hin).
- Wir setzen auf **menschenzentrierte, vom Menschen geschaffene und kontrollierte KI**. Auch selbstlernendes Machine Learning erfolgen nur unter menschlicher Kontrolle.
- Setzen wir bei der Auswahl, Publikation und Distribution unserer Inhalte Algorithmen ein, beispielsweise **Personalisierungssystemen** mit/ohne KI, vermeiden wir die durch den Algorithmus verursachte Voreinstellung auch den sogenannten «Filterblasen-Effekt»: Die präsentierte Auswahl ist immer zu einem Teil redaktionell kontrolliert.

Gleichzeitig verfolgen wir im Kontext der **gesellschaftlichen Debatte zu Datenethik und digitaler Ethik** dazu, unsere Nutzerinnen und Nutzer mit den Chancen und Risiken solcher Technologien vertraut zu machen und generell das **Vertrauen in die Glaubwürdigkeit von Medien** zu stärken. Hierbei sehen wir eine besondere Verantwortung von SRF als Service public.

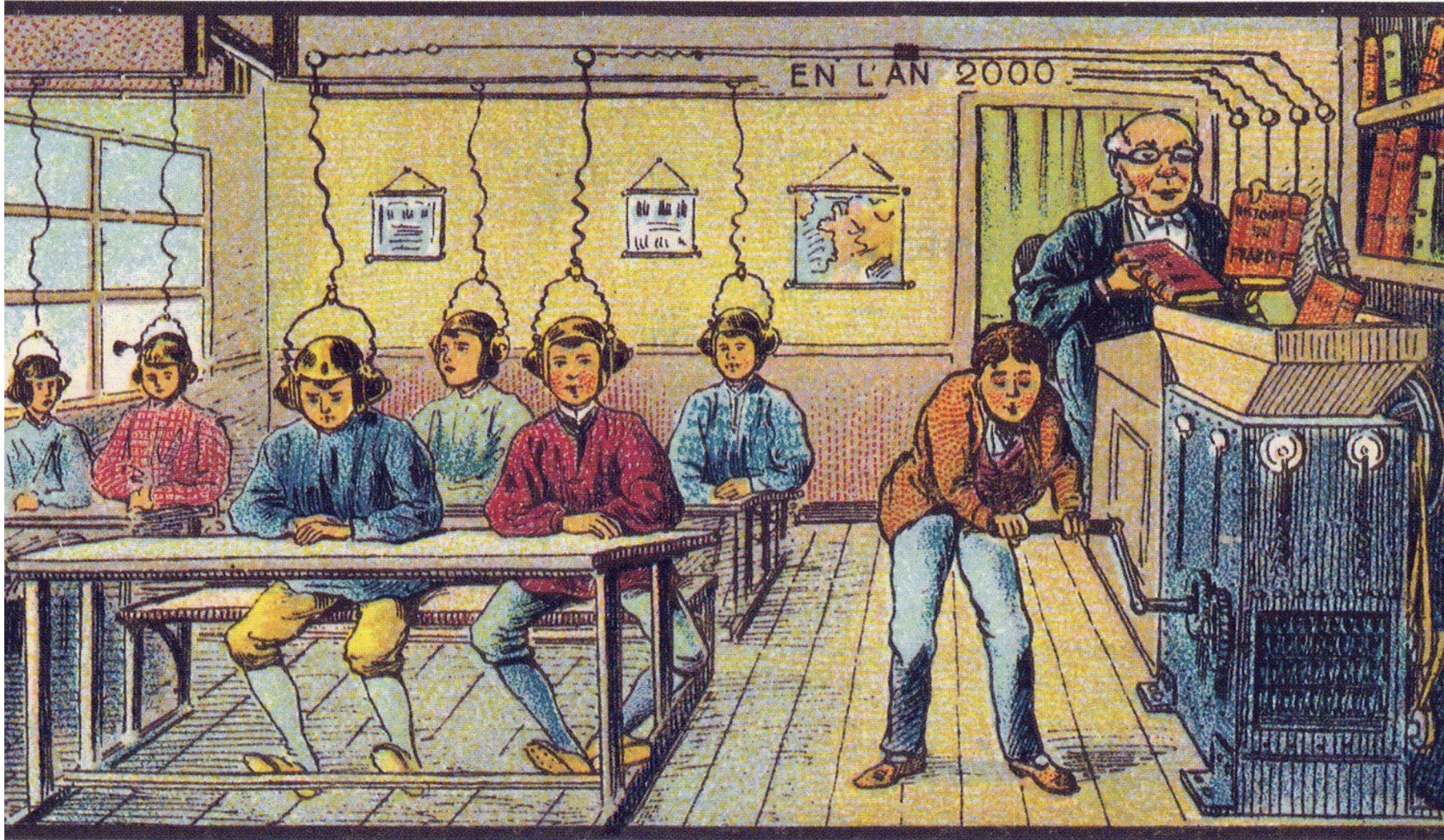


SRF

SRF Schweizer Radio und Fernsehen,
Zweigniederlassung der Schweizerischen Radio-
und Fernsehgesellschaft

[Kontakt](#) | [Datenschutz](#) | [Impressum](#)

What does it mean to learn? What does it mean to know?



At School

PART IV

DISCUSSION