

Talking about dementia

.....

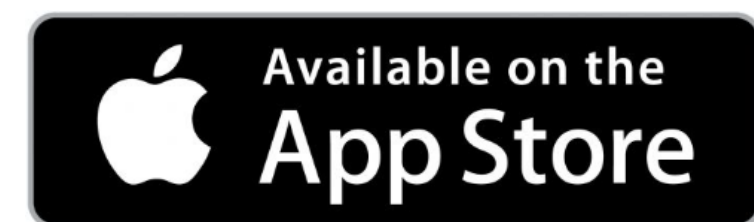
An LLM-based chat solution for
the people involved

Benedek Ágota, Barnabás Molnár, Csaba
Potyók, Tamás Mészáros (BME MIT),
Ákos Szőke

szoke.akos@multilogic.hu

<https://mental-fitness.hu>

27 November 2024



Agenda



- Looking for medical answers
 - classical way (search engine)
 - LLM approach (AI approach)
 - RAG method (improved AI approach)
- MentalFitness Chat - a RAG solution
- Experiments and Validation

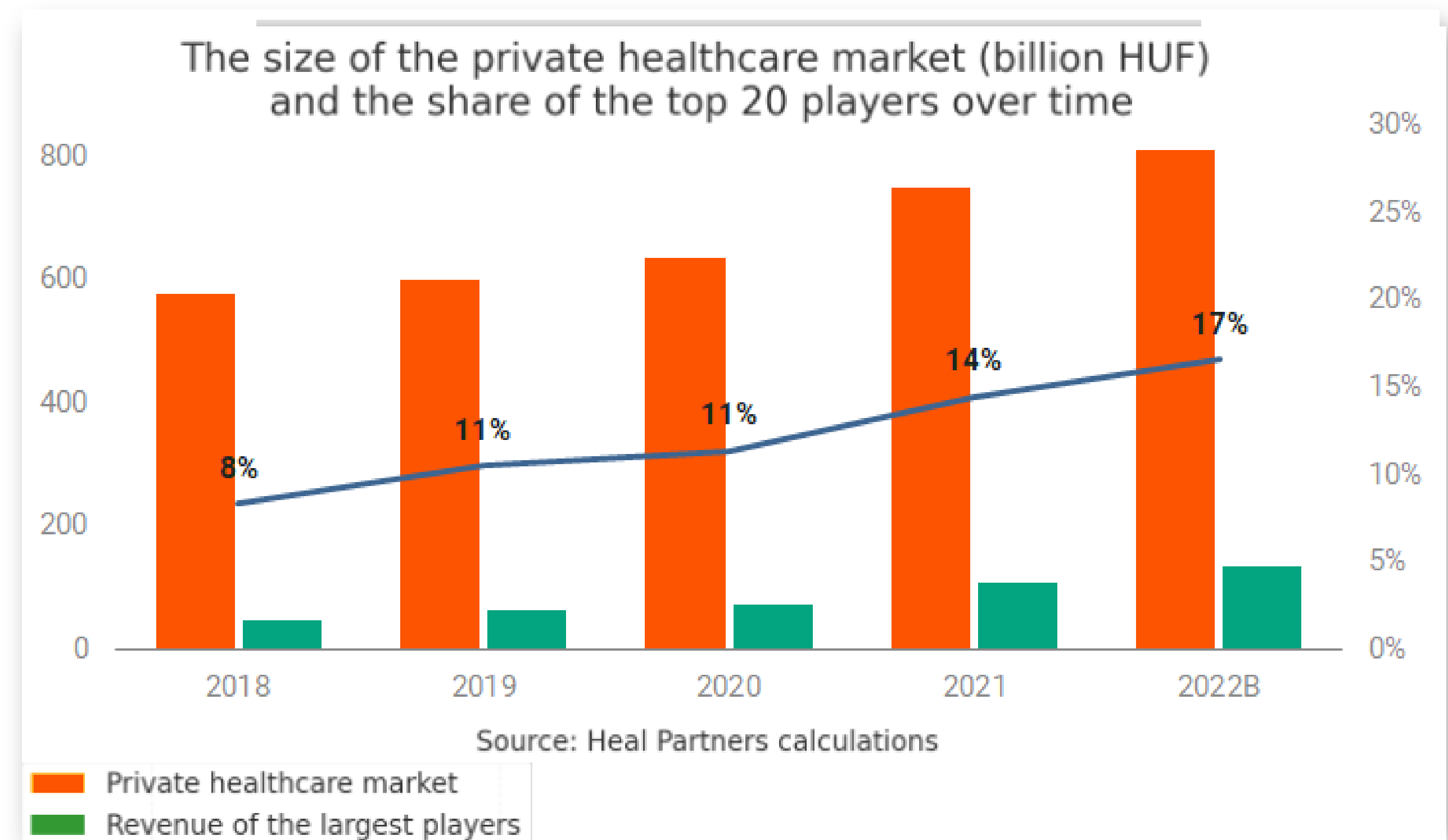
Most Popular Mental Health-related Questions



Vaidyanathan, U. et al. An evaluation of Internet searches as a marker of trends in population mental health in the US. *Sci Rep* 12, 8946 (2022).
Kalckreuth S et al. Mental health related Internet use among psychiatric patients: a cross-sectional analysis. *BMC Psychiatry*. 2014 Dec 24;14:368.

Everyone has medical questions

- To find reliable answers we need doctors (experts), but it's:
 - difficult to reach
 - time-consuming (queues, bureaucratic)
- Therefore, we often turn to
 - Private healthcare services (**expensive**)
 - The Web / our relatives (**unreliable**)
- Using web often results in:
 - Uncertainty
 - Stress
 - Medical mistreatment
 - Hypochondria



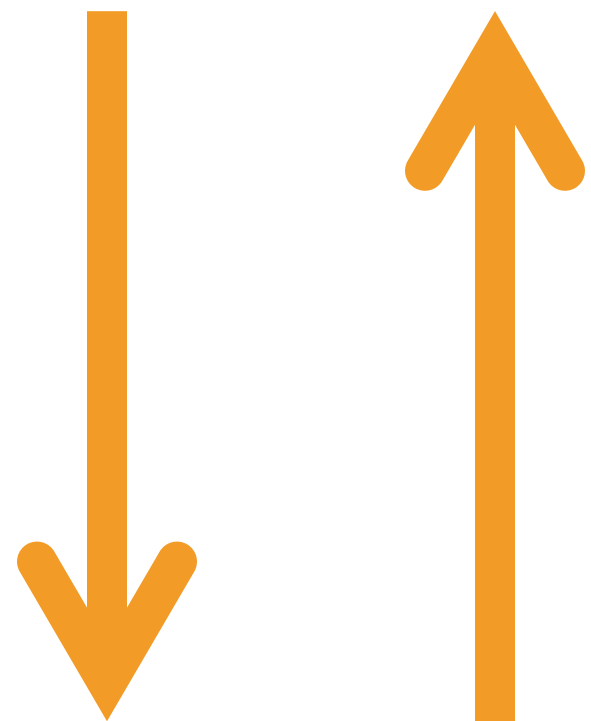
Website visitors in November 2024

webbeteg.hu	vs.	hazipatika.com	egeszsegkalauz.hu	informed.hu	benu.hu
2.210M		5.904M	1.430M	68,611	2.058M

What we expect is like a live chat

A fast and reliable answer to our medical questions

Q: Formulate
question in natural
language



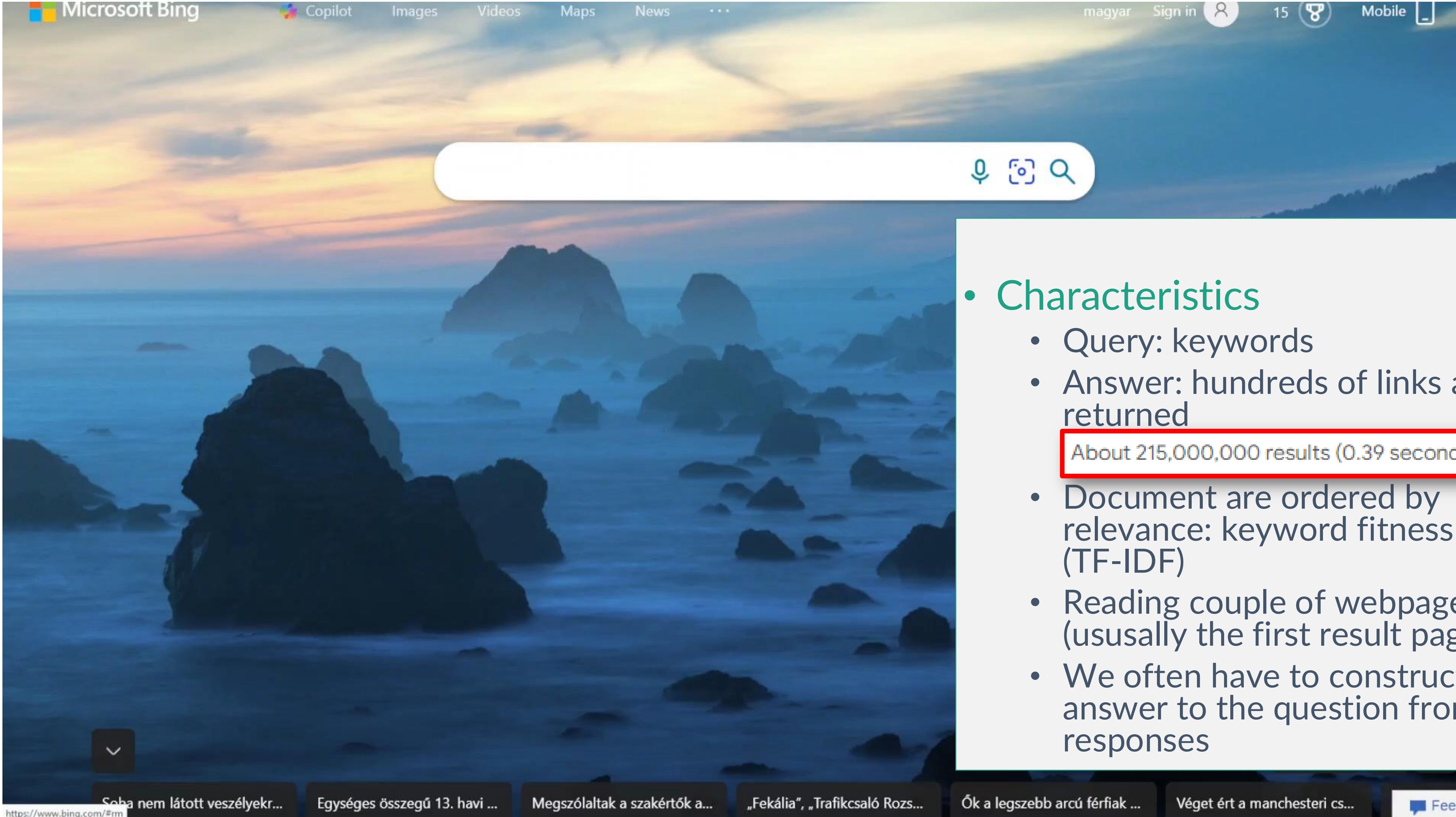
A: Natural
language answer
to the question



• Our expectations

- Questioning in natural language
- Natural language answers
- Efficiency and effectiveness
- Personalized assistance
- 24/7 availability
- Trustworthy answers

Demo #1 Let's „google” answer to „When is a good age to start taking action to reduce my risk for developing dementia?” A search engine based information retrieval



Microsoft Bing Copilot Images Videos Maps News magyar Sign in 15 Mobile

Search bar: [Empty]

- **Characteristics**
 - Query: keywords
 - Answer: hundreds of links are returned
 - About 215,000,000 results (0.39 seconds)
 - Document are ordered by relevance: keyword fitness (TF-IDF)
 - Reading couple of webpages (usually the first result page)
 - We often have to construct answer to the question from the responses

News ticker: Soha nem látott veszélyek... Egységes összegű 13. havi ... Megszóltak a szakértők a... „Fekália”, „Trafikcsaló Rozs... Ók a legszebb arcú férfiak ... Véget ért a manchesteri cs... Fee

URL: https://www.bing.com/#rr

Searching the Web Using Search Engines (~ 20 years old)

Q: Formulate question in natural language

1 Formulate the question using keywords

2 Select reliable sources from the responses

3 Evaluate the responses

4 Gather results

5 Reformulate the keywords

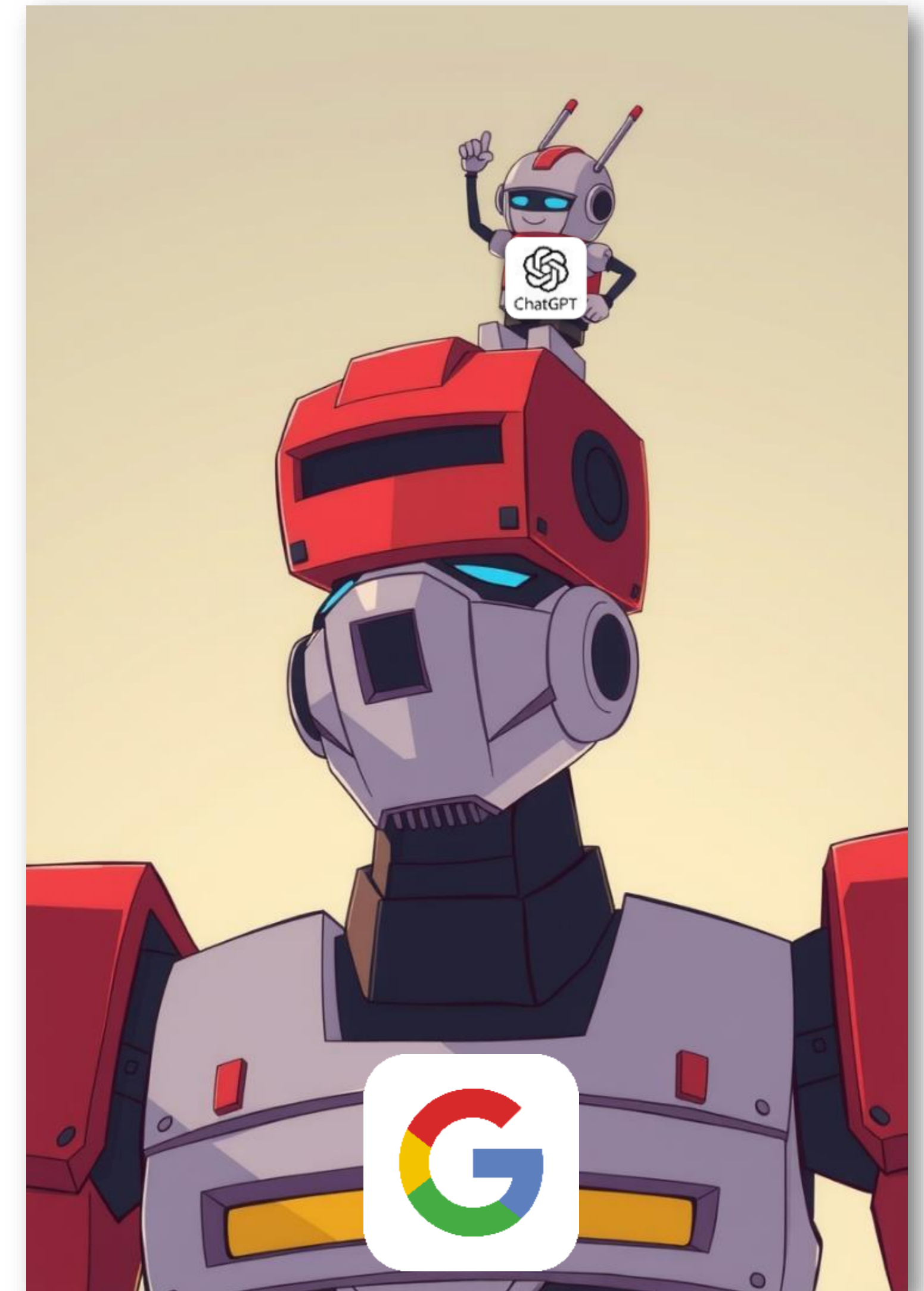
6 A: Answer to the question



2017: Google revolutionized text generation

The starting point of the revolution

- **Google introduced the Transformer***, which rapidly became the state-of-the-art approach to solve NLP problems.
- OpenAI's GPT (Generative Pre-trained Transformer – such as ChatGPT, DALL.E) reposes on Transformers.



*: Vaswani (2017), attention Is All You Need (doi.org/10.48550/arXiv.1706.03762)

Image, generated by Dall.E

What are Large Language Models?

A short intro

• Definition

- A large language model (LLM) is a type of artificial intelligence (AI) that is trained on a massive datasets of text and code. This allows the model to learn the statistical relationship between words and phrases, which in turn allows it to
 - Generate text
 - Translate languages
 - Write different kind of creative content, and
 - **Answer your question in an informative way (THIS IS OUR FOCUS)**

• Common LLMs

- GPT 3.5, 4.0
- Bard, Gemini
- LLama 3



How Do Large Language Models (LLMs) Work? #1

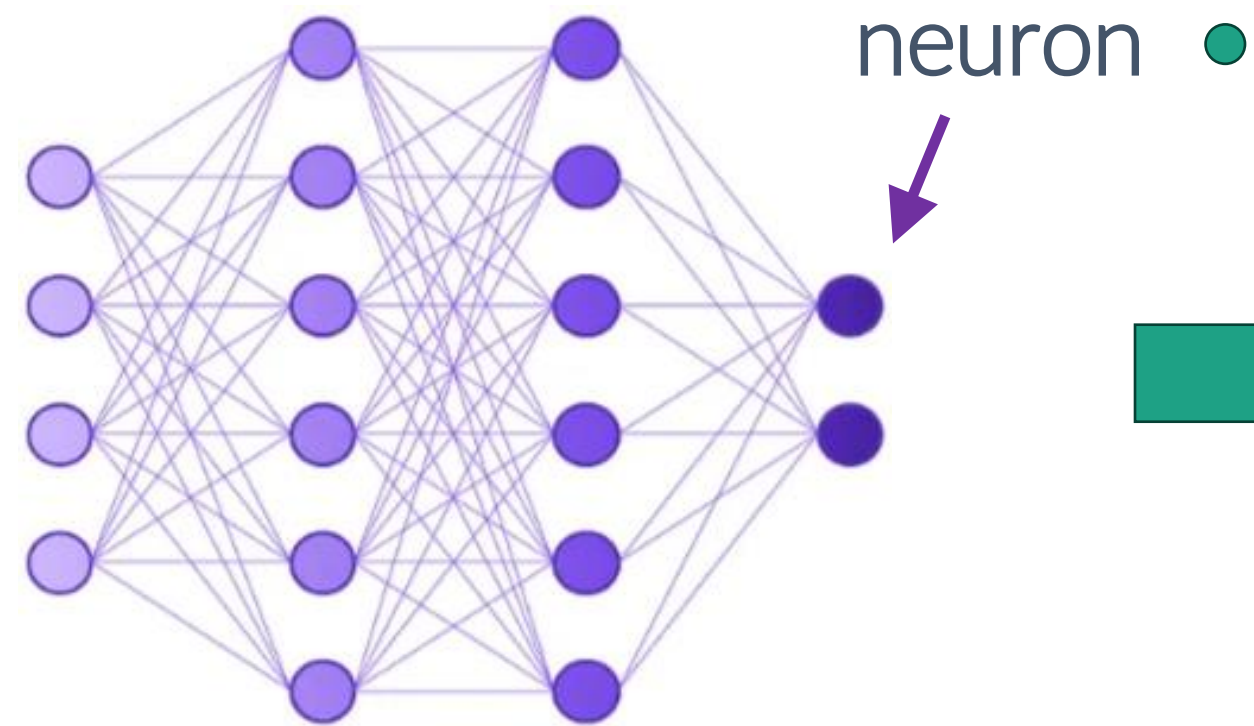
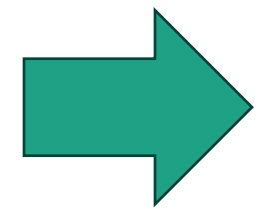
Main mechanics (in reality things are a little more complex...)

- Predicting the next word in a sentence

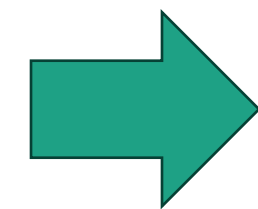


[Cats like to sleep in the ???]

Input

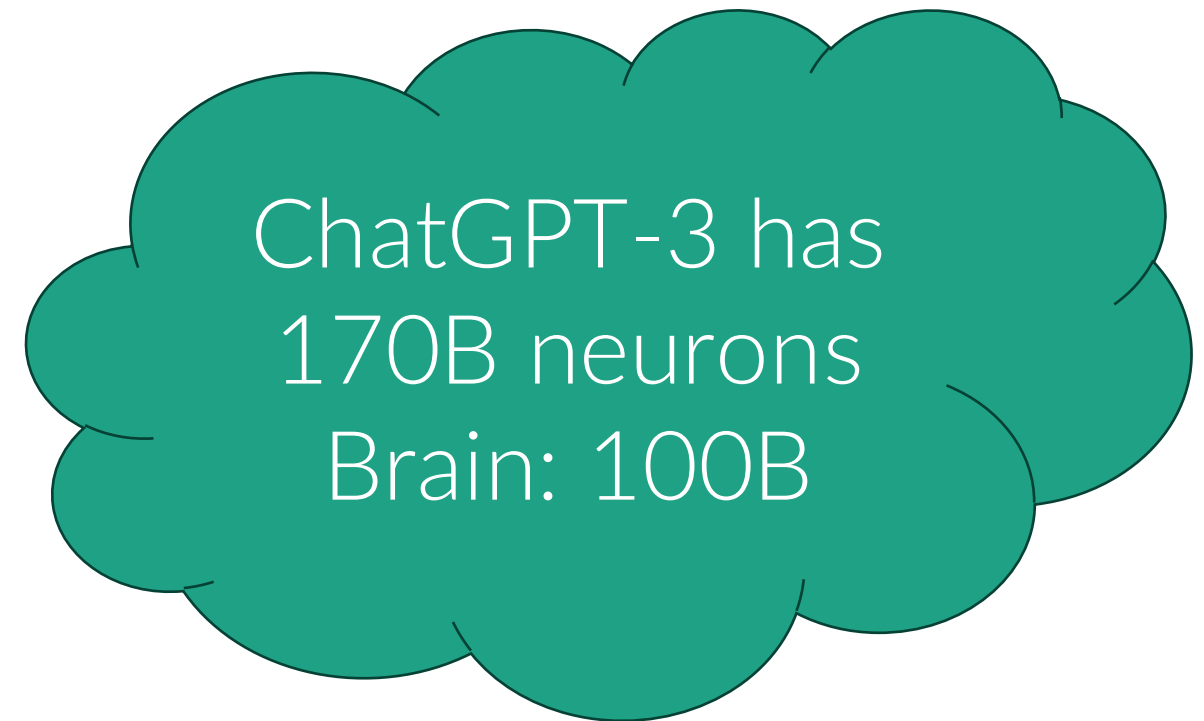


Neural network



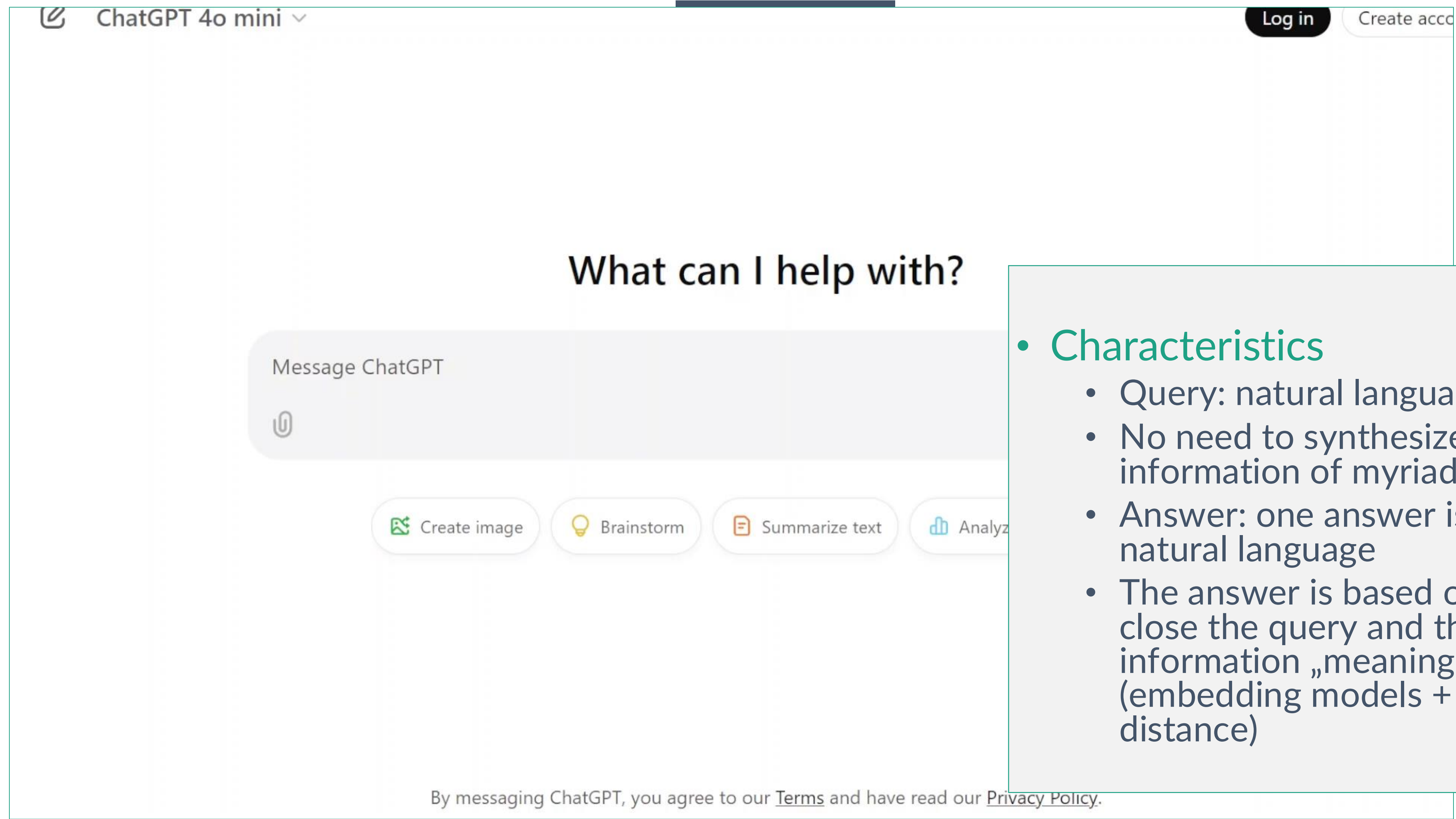
Word	Probability
ability	0.002
bag	0.071
box	0.085
...	...
zebra	0.001

Output



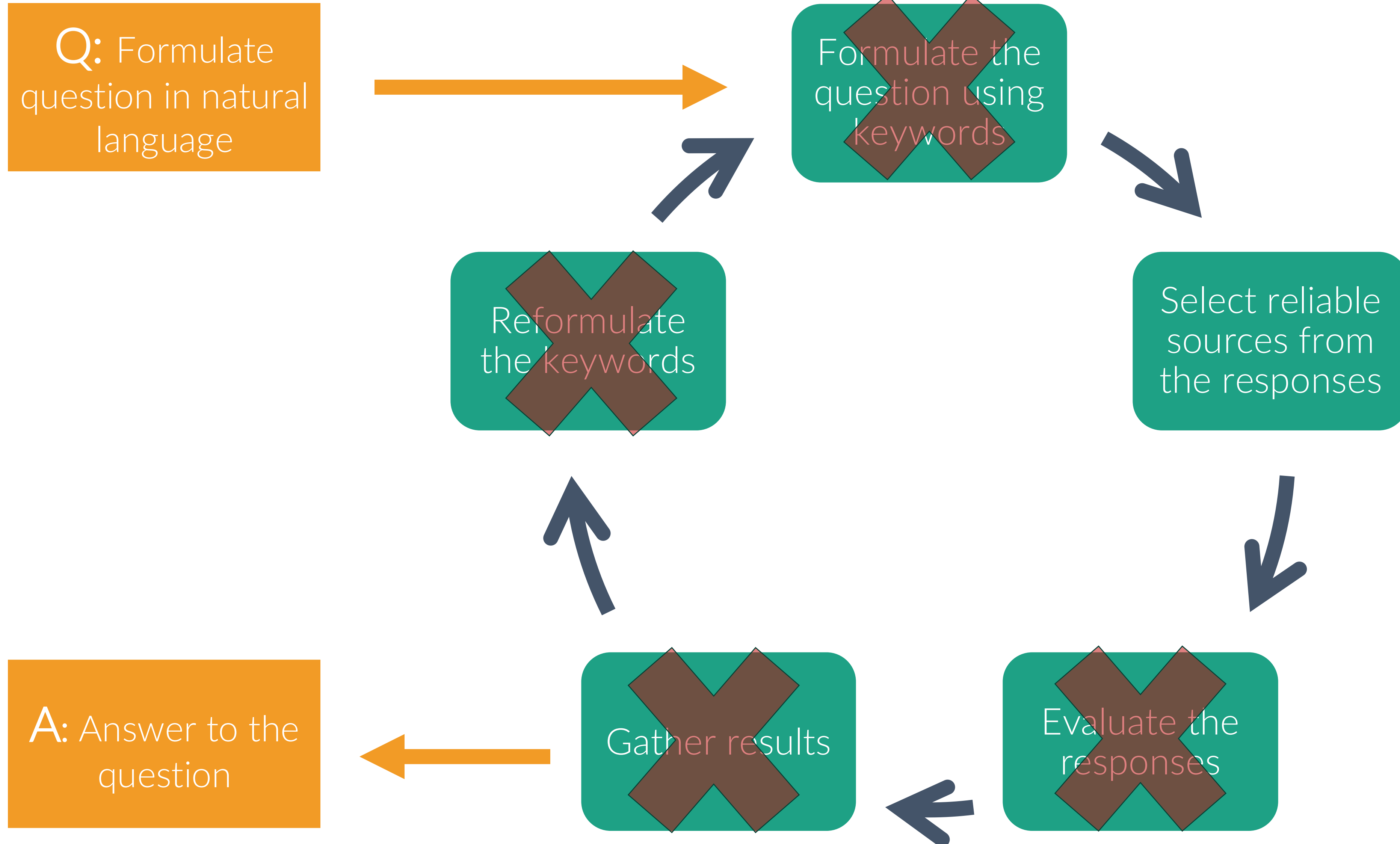
Demo #2 Let's ask ChatGPT to answer to „When is a good age to start taking action to reduce my risk for developing dementia?“

An LLM-based information retrieval



- **Characteristics**
 - Query: natural language
 - No need to synthesize information of myriad of pages
 - Answer: one answer is given in natural language
 - The answer is based on how close the query and the found information „meaning“ (embedding models + vector distance)

Get answer using LLMs: simplifies querying



Problem #1: Hallucinations

A recent Tidio survey found that 72% of users trust AI to provide factual and reliable information. Moreover, 75% of those respondents reported AI to mislead at least once.

• Hallucinations

- are confident statements that are False
- at the moment affects all known LLMs

• Main causes of hallucinations

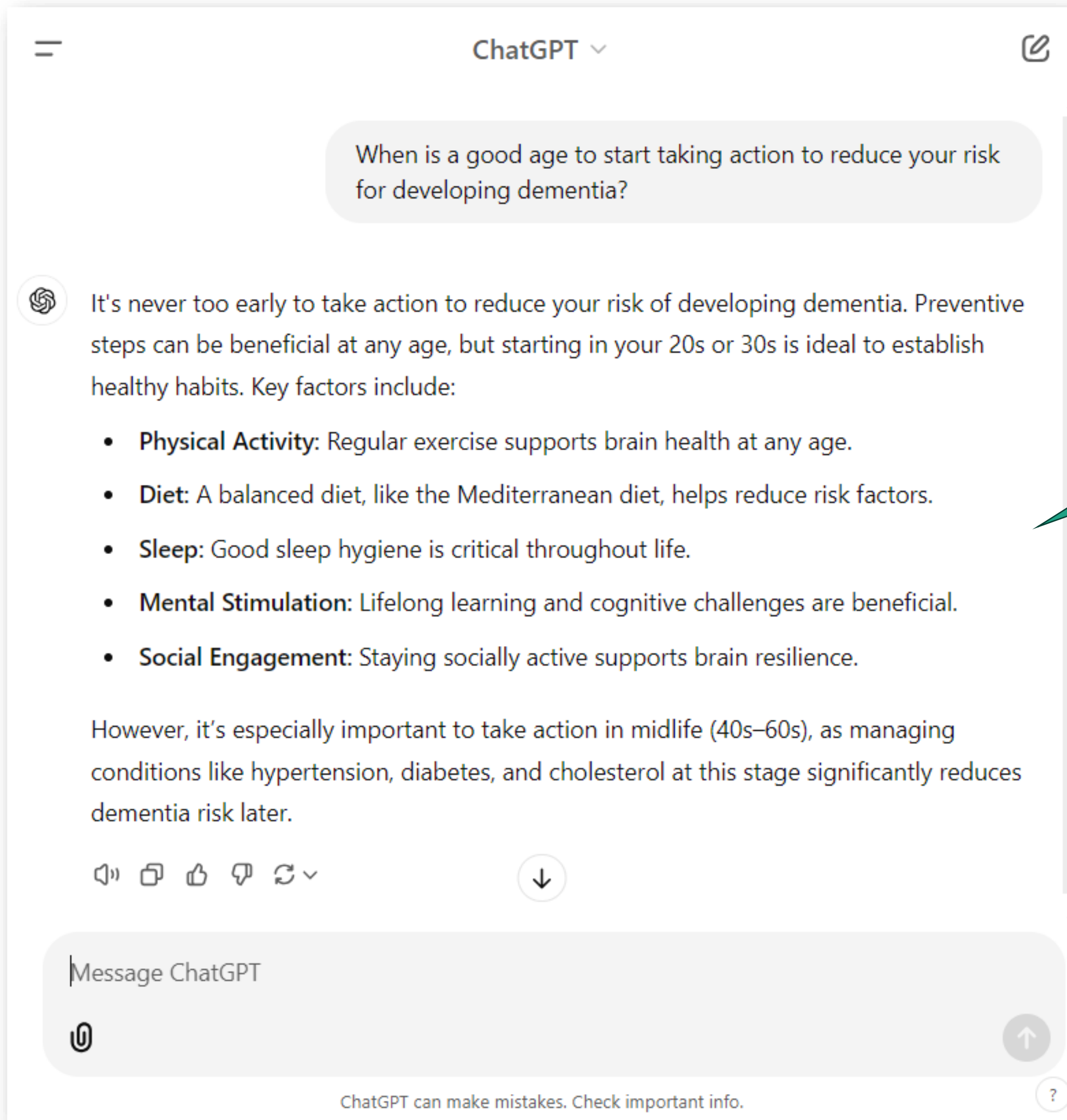
- **Factual Error:** language models cannot differentiate between a truth and a lie
- **Source Conflation:** model tries to combine extracted details from various sources

Falsely Accused Professor

ChatGPT response accused a law professor of sexual harassment. The LLM even cited a non-existent Washington Post article as the information source.



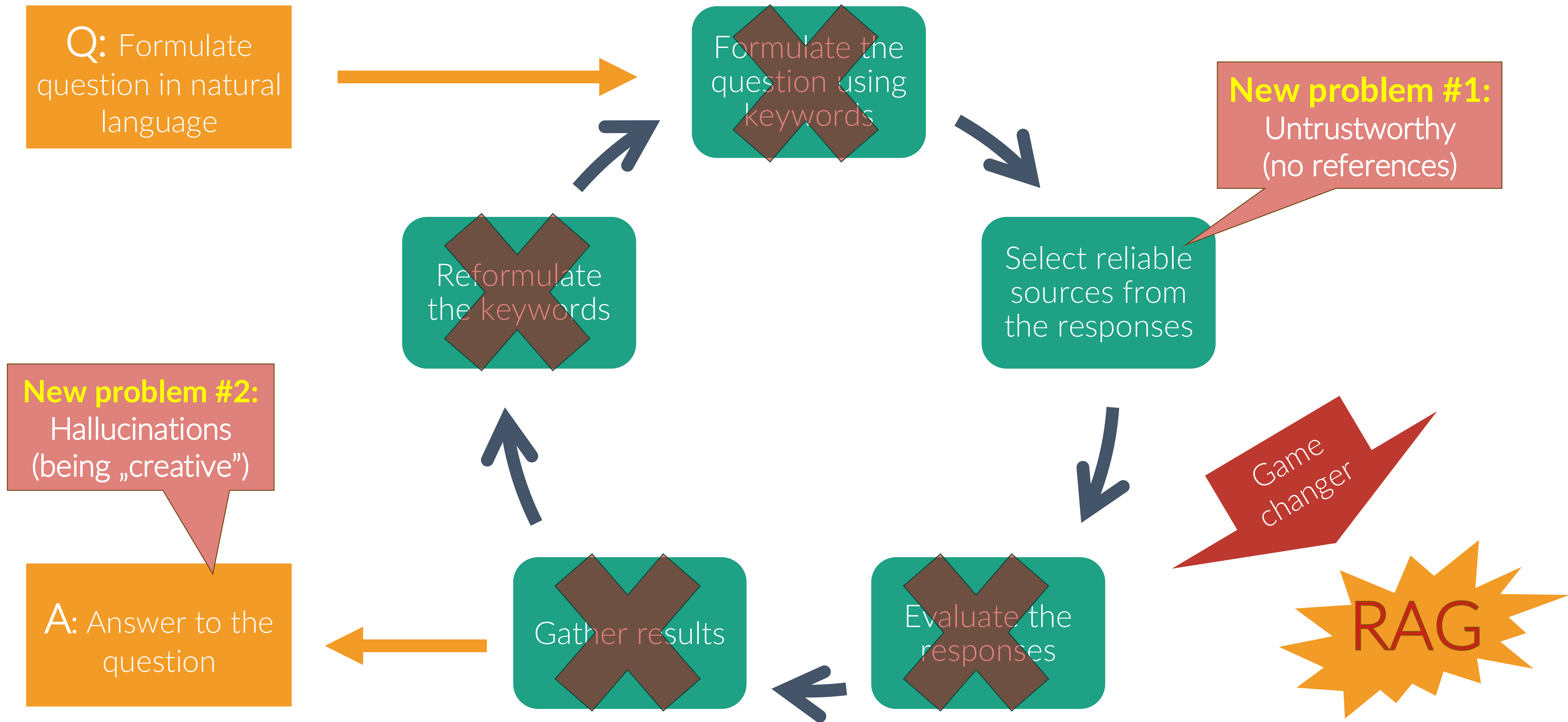
Problem #2: Untrustworthy



No reference

No information about the source information. Therefore, it cannot be checked.

Get answer using LLMs: new problems emerge



RAG architecture to solve LLMs problems

What is Retrieval-Augmented Generation?

A process of optimizing the output of a large language model, so it references an authoritative knowledge base outside of its training data sources before generating a response

Process

1. Enter question
2. Search for relevant information
3. Retrieve relevant information
4. Feed LLM with question and retrieved info
5. Answer the information
6. Present the answer



MentalFitness Chat follows RAG architecture

Cca. 250 pages

The screenshot shows the MentalFitness chat interface. On the left is a sidebar with navigation options: 'Explore dementia', 'Dementia FAQ', 'Chat about dementia', 'Chat history', 'TU', and 'S'. The main chat area is titled 'Chat about dementia' and 'Talk with an assistant'. It contains a user question: 'When is a good age to start taking action to reduce the risk of developing dementia?'. The bot's answer states: 'It is important to take action to reduce your risk of developing dementia at any age. Research indicates that lifestyle and environmental factors significantly influence dementia risk, and adopting healthy habits early in life can contribute to lifelong brain health. Engaging in activities such as staying socially active, being physically active, and managing medical conditions can help improve brain health throughout your life.' Below the answer, there is a 'Sources' section with two references: '10 ways to reduce your risk of dementia - Alzheimer Society of Canada' and '12 actions for a healthier brain at any age - Alzheimer Society of Canada'. A green callout box labeled 'Reference' points to these sources with the text 'Information about the source.' At the bottom of the chat area, there is a text input field with the placeholder 'Is there anything I can help you with?' and a send button. The footer of the interface says 'Get Help with Dementia Edition'.

User question

Bot answer

Reference
Information about the source.

• Source information

- www.alzheimer.ca
In English (Canadian)
Website



- www.felejte.hu
In Hungarian
Website



- www.demenszia.hu
In Hungarian
Youtube video



Cca. 5 hours audio

Demo #3 Let's ask MentalFitness Chat App to answer to „When is a good age to start taking action to reduce my risk for developing dementia?“

Explore dementia

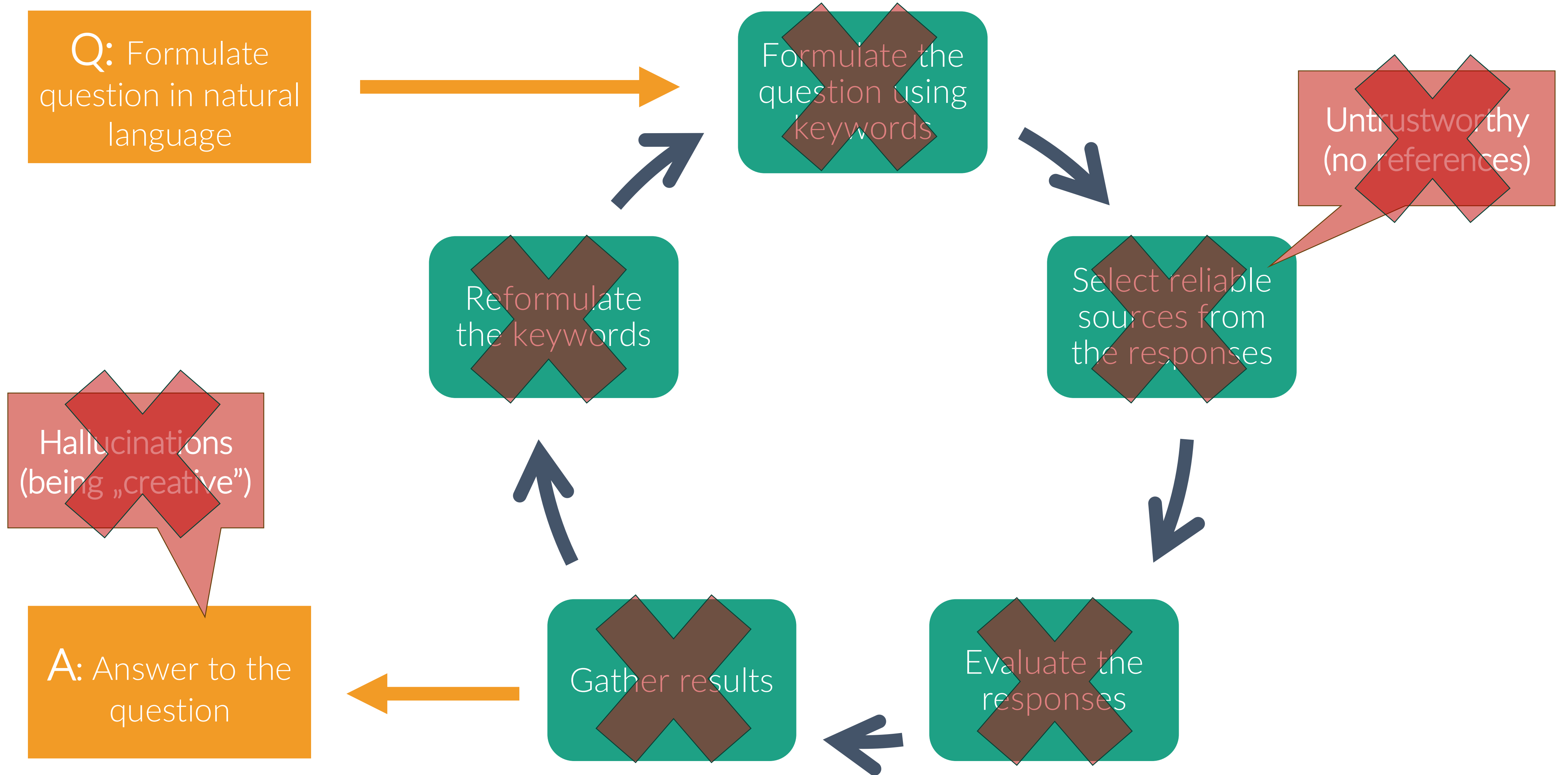
Find answers like never before

MentalFitness Explore searches information using natural language by understanding your intents and responding with the relevant (curated) sources and related topics.

 Explore does not have context memory. It answers only the last question.


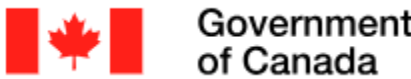


Wh| 

Get answer using RAG: it solves both the untrustworthiness and the hallucination problems



Preliminary Validation with Quizzes

- Publicly available dementia quizzes for the common people (28 questions)

	Link	Test result
Neuroscience Research Australia (NeuRA) 	https://neura.edu.au/quiz/the-dementia-quiz	5/5 ✓
Government of Canada 	https://www.canada.ca/en/public-health/services/diseases/dementia/risk-factors-prevention/quiz.html	10/10 ✓
AARP (All Older Americans) Nonprofit organization 	https://www.aarp.org/caregiving/health/info-2024/dementia-caregiver-quiz.html	8/8 ✓
Webbeteg 	https://www.webbeteg.hu/cikkek/demencia/22323/alzheimer-kor-kviz	5/5 ✓

Preliminary Validation with FAQs

- Example FAQs

Question	Ground truth
My mother has Alzheimer's disease. Will I get it?	It is only in a few rare cases where Alzheimer's disease runs in families. In these types of cases, there is a direct link between an inherited mutation in one gene and the onset of the disease. There are rare familial forms of dementia caused by genetic mutations such as familial Alzheimer's disease, frontotemporal dementia and familial vascular dementia, which are more likely to occur in people under the age of 65. For families where this is the case, family members, such as brothers, sisters and children, have a one in two chance of developing Alzheimer's disease.
How is dementia diagnosed?	Talk with your doctor about the unique changes you see in your person. There is no single test to diagnose any dementia. Diagnosis usually includes reviewing a person's medical history, physical exams, lab tests, imaging procedures, and cognitive assessments. The family doctor may refer the person to a specialist such as a neurologist, geriatrician, or psychiatrist. A diagnosis of probable dementia is made after excluding all other potential causes.

Preliminary Validation with FAQs

- Validated with RAGAS

- Relies on question and ground truth answer pairs (in our case **171 FAQ pairs**)
- It compares the Chatbot answers with the ground truth and calculates different metrics

Important note:
Only three, a small number of, sources are used!

Metrics	Definition	Validation result
Answer relevancy	Assessing how pertinent the generated answer is to the given prompt	94%
Faithfulness	How accurately the response generated by LLM reflects the information in the retrieved documents.	88%
Semantic similarity	The assessment of the semantic resemblance between the generated answer and the ground truth	92%

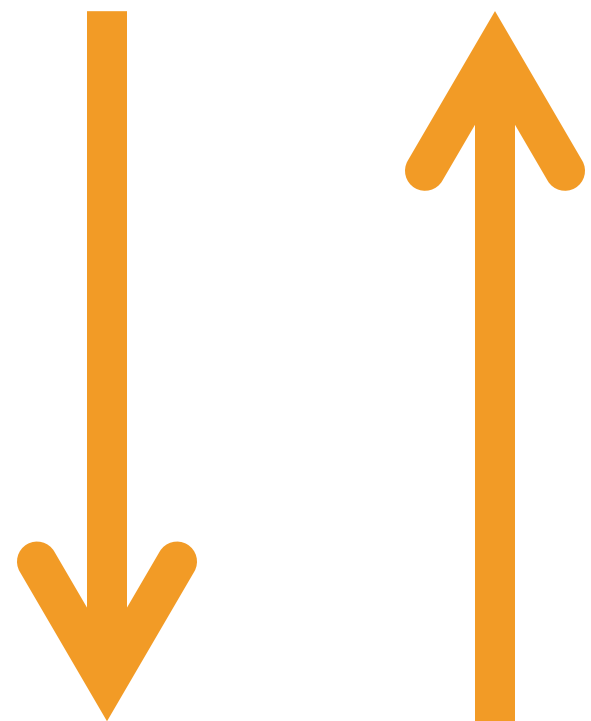
These are just quick validations, human expert oversight is still needed!

16 FAQ sites : alzheimercalgary.ca, yalemedicine.org, alzheimer.ca, psychiatry.org, alzheimersresearchuk.org, nia.nih.gov, dementiauk.org, wehelpcaregivers.com, ...

MentalFitness Chat characteristics



Q: Formulate question in natural language



A: Natural language answer to the question



- ✔ Questioning in natural language
- ✔ Natural language answers
- ✔ Efficiency and effectiveness
- ✔ Personalized assistance
- ✔ 24/7 availability
- ✔ Trustworthy answers