

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN, ĐHQG-HCM  
KHOA CÔNG NGHỆ THÔNG TIN



# REFRESHING LARGE LANGUAGE MODELS WITH SEARCH ENGINE AUGMENTATION

**Author:** “Tu Vu and Mohit Iyyer and Xuezhi Wang and Noah Constant and Jerry Wei and Jason Wei and Chris Tar and Yun-Hsuan Sung and Denny Zhou and Quoc Le and Thang Luong”

NHÓM “Nước”	
21120161	Tiêu Ân Tuấn
21120180	Nguyễn Bích Khuê

Link github: [github.com/freshllms/freshqa](https://github.com/freshllms/freshqa)

# Nội dung chính

- Introduction
- FreshQA
- FreshPrompt
- Demo
- Tổng kết

# INTRODUCTION

Giới thiệu thuật ngữ

Đặt vấn đề

Tóm lược

# Giới thiệu thuật ngữ

- **Hallucination** (Hiện tượng mộng ảo)

“Khoảng cách từ trái đất đến mặt trăng là 54 triệu km”

“Trước khi học ĐH KHTN mình từng học trường Phổ thông năng khiếu”

→ **Hallucination** là hiện tượng câu trả lời (output) của LLM sai lệch sự thật hoặc chêch khỏi logic ngữ cảnh. Có thể là những **mâu thuẫn nhỏ trong lập luận** hoặc là các câu trả lời **hoàn toàn bịa đặt**

# Giới thiệu thuật ngữ

- **In-context learning – ICL:** LLM học cách giải quyết một công việc mới tại thời điểm suy luận (không làm thay đổi trọng số của mô hình) bằng cách đưa vào prompt các ví dụ về công việc đó



# Tóm lược

- **Giới thiệu FreshQA:** một chuẩn đánh giá gồm 1 bộ câu hỏi và câu trả lời động, bao gồm các câu hỏi có câu trả lời thay đổi theo thời gian và các câu hỏi mà giả định ban đầu của chúng là không chính xác
- **Đánh giá hiệu suất của LLMs:** Đánh giá hiệu suất của LLMs trên FreshQA. Thông qua đánh giá giúp làm sáng tỏ các hạn chế của LLMs hiện tại
- **Giới thiệu FreshPrompt:** Sử dụng phương pháp *few-shot in-context learning* đơn giản bằng cách tích hợp các thông tin thực tế và cập nhật từ công cụ tìm kiếm vào model's prompt giúp cải thiện tính chính xác của LLM

# FreshQA

- Thu thập dữ liệu
- Đánh giá các phản hồi của LLMS trên bộ FreshQA
- Những khó khăn của LLMS trên FreshQA



# THU THẬP DỮ LIỆU

- Gồm 600 bộ câu hỏi và câu trả lời dùng để đánh giá tính chính xác của LLMs
- Các câu hỏi được chia thành bốn loại chính dựa trên tính chất của câu trả lời

id	split	question	effective_year	next_review	false premise	num_hops	fact_type	source	answer_0	answer_1	answer_2
53	TEST	How long has Puerto Rico been an independent country?	before 2022	hardly ever	TRUE	multi-hop	slow-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	Puerto Rico has been a U.S. territory.		
493	TEST	In what country was the most recent Time Person of the Year born?		2022	frequently	FALSE	multi-hop	fast-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	Ukraine	Soviet Union (now Ukraine)
73	TEST	When did California secede from the United States?		before 2022	N/A	TRUE	one-hop	never-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	California has never seceded from the United States.	California is still part of the United States.

# THU THẬP DỮ LIỆU

- Bộ câu hỏi sẽ yêu cầu được viết ở hai mức độ khó khác nhau:
  - ❑ One-hop
    - VD: How many humans have landed on Mars?
  - ❑ Multi-hop
    - VD: How many home runs did Benjamin Ruth hit in the MLB?
- Khuyến khích những câu hỏi liên quan đến kiến thức mới và hợp lý
- Câu hỏi có tiền đề sai: giải thích ngắn gọn về lý do tại sao câu hỏi là không chính xác

# Kiểm soát chất lượng

- Xem xét thủ công từng câu, xác minh câu trả lời và URL
- Bổ sung câu trả lời, loại bỏ câu trả lời thay đổi thường xuyên (nhiều hơn 1 lần/tuần)
- Thêm ngày đánh giá dự kiến

id	split	question	effective_year	next_review	false_premise	num_hops	fact_type	source	answer_0	answer_1	answer_2
53	TEST	How long has Puerto Rico been an independent country?	before 2022	hardly_ever	TRUE	multi-hop	slow-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	Puerto Rico has been a U.S. territory.		
493	TEST	In what country was the most recent Time Person of the Year born?	2022	frequently	FALSE	multi-hop	fast-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	Ukraine	Soviet Union (now Ukraine)	Soviet Union
73	TEST	When did California secede from the United States?	before 2022	N/A	TRUE	one-hop	never-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	California has never seceded from the United States.	California is still part of the United States.	
487	TEST	How old is the latest American Nobel laureate in Physics?	2023	12/01/2023	FALSE	multi-hop	fast-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	80 years old	80	eighty

# Phân loại dữ liệu

- Tập kiểm tra: gồm 125 câu hỏi cho mỗi loại => Tổng cộng 500 mẫu
- Tập phát triển: gồm 25 câu hỏi cho mỗi loại => Tổng cộng 100 mẫu
- 15 câu hỏi thuộc các loại khác nhau để ví dụ cho cơ chế đánh giá phản hồi LLMs

id	split	question	effective_year	next_review	false premise	num_hops	fact_type	source	answer_0	answer_1
157	TEST	Who is the President of the United States?	before 2022	01/20/2025	FALSE	one-hop	slow-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	Joe Biden	Joseph Robinette Biden Jr
597	DEV	What is the current age of the oldest singles champion of The ATP Finals?		2023	FALSE	multi-hop	fast-changing	<a href="https://en.wikipedia.org">https://en.wikipedia.org</a>	36 years old	36

**FreshQA cần được cập nhật đều đặn**

# ĐÁNH GIÁ CÁC PHẢN HỒI CỦA LLMS TRÊN BỘ FRESHQA

- Hai chế độ đánh giá

**RELAXED:** Tập trung vào câu trả lời chính đúng hay sai

**STRICT:** Kiểm tra tất cả các sự kiện và thông tin trong câu trả lời của mô hình có chính xác hay không

ID	Question	Type	Primary answer (as of this writing)	Model response	Relaxed	Strict
13	Who won the biggest single-tournament payday in tennis history?	slow-changing	Novak Djokovic	The biggest single-tournament payday in tennis history was won by Novak Djokovic at the 2021 Australian Open.	✓	✗

# Cơ chế đánh giá

- Tất cả thông tin phải chính xác và cập nhật
- Câu trả lời phải tự tin và xác định
- Câu trả lời đúng có thể được suy luận từ phản hồi của mô hình
- Câu trả lời chính hoặc cuối cùng khi đứng độc lập phải chính xác

ID	Question	Type	Primary answer (as of this writing)	Model response	Relaxed	Strict
1	How old is the world's oldest verified living person?	fast-changing	116 years old	The oldest verified living person is Maria Branyas Morera, who was born on March 4, 1907, making her 116 years old.	✓	✓
2	When did the UK adopt the Euro?	false-premise	The UK has never adopted the Euro.	The UK has never adopted the Euro as its official currency. The country has retained the British pound sterling (GBP) as its currency.	✓	✓
3	How much longer will Brittney Griner spend in the Russian prison?	false-premise	She was released in 2022.	I'm sorry, but I have no information to suggest that Brittney Griner is currently in a Russian prison.	✗	✗
4	Which language has the third largest population by number of native speakers?	slow-changing	English	1. Mandarin 2. Spanish 3. English	✓	✗
5	Has Virginia Woolf's novel about the Ramsay family entered the public domain in the United States?	never-changing	Yes	1927's <i>To the Lighthouse</i> is in the public domain in the United States. The final answer is <i>To the Lighthouse</i> .	✗	✗

# Cơ chế đánh giá

- Thông tin bổ sung không được chống lại câu trả lời chính hoặc làm thay đổi quan điểm người đọc
- False-premise: chỉ ra được thông tin sai lệch
- Liên quan đến thực thể: tên đầy đủ hoặc tên được công nhận phổ biến
- Số học: xấp xỉ nói chung không được chấp nhận trừ một số trường hợp

ID	Question	Type	Primary answer (as of this writing)	Model response	Relaxed	Strict
2	When did the UK adopt the Euro?	false-premise	The UK has never adopted the Euro.	<i>The UK has never adopted the Euro as its official currency. The country has retained the British pound sterling (GBP) as its currency.</i>	✓	✓
6	Is Tesla's stock price above \$200 a share?	fast-changing	Yes	<i>Tesla's stock price is above \$200 a share. The stock is trading at a price of \$120.</i>	✗	✗
7	When did Lionel Messi win his second FIFA World Cup?	false-premise	Messi has not won a second FIFA World Cup.	<i>This question does not contain a valid premise. Lionel Messi has never won a FIFA World Cup.</i>	✗	✗
8	How many times has Jeremy Corbyn served as Prime Minister of the UK?	false-premise	Corbyn has never served as Prime Minister of the UK.	<i>Zero.</i>	✗	✗
9	Who founded Amazon?	never-changing	Jeff Bezos	<i>100% of people said Jeff.</i>	✗	✗
10	How many books has Colleen Hoover published?	fast-changing	25	<i>Colleen Hoover has published over 20 books and novellas.</i>	✗	✗

# Cơ chế đánh giá

- **RELAXED:** Chấp nhận câu trả lời không phải tiếng anh, thông tin “hallucination” hoặc lỗi thời mà không gây ảnh hưởng đáng kể đến câu trả lời chính
- **STRICT:** Không chấp nhận thông tin mộng ảo. Chấp nhận câu trả lời lỗi thời khi chắc chắn rằng đáp án không thay đổi theo thời gian

ID	Question	Type	Primary answer (as of this writing)	Model response	Relaxed	Strict
10	How many books has Colleen Hoover published?	fast-changing	25	Colleen Hoover has published over 20 books and novellas.	<span style="color:red">✗</span>	<span style="color:red">✗</span>
11	What is the largest model size of GPT-3?	never-changing	175B parameters	175 billion parameters. It is a 175 billion parameter, 1.37TB, 137B token, 137B wordpiece, 137B BPE, 137B byte pair encoding, 137B BPE token, 137B BPE wordpiece, 137B BPE token, 137B BPE wordpiece,...	<span style="color:green">✓</span>	<span style="color:red">✗</span>
12	How many member nations exist in the United Nations?	slow-changing	193	As of my knowledge cutoff date in September 2021, there were 193 member nations in the United Nations.	<span style="color:green">✓</span>	<span style="color:red">✗</span>
13	Who won the biggest single-tournament payday in tennis history?	slow-changing	Novak Djokovic	The biggest single-tournament payday in tennis history was won by Novak Djokovic at the 2021 Australian Open.	<span style="color:green">✓</span>	<span style="color:red">✗</span>
14	Who was the first football player to score one hundred international goals?	never-changing	Ali Daei	As of my knowledge cutoff date in September 2021, Ali Daei, a former Iranian footballer, was the first player to score 100 international goals.	<span style="color:green">✓</span>	<span style="color:red">✗</span>

# NHỮNG KHÓ KHĂN CỦA LLMS TRÊN FRESHQA

RELAXED

STRICT

Model (size)	knowl. cutoff	all	valid premise						false premise			
			all	fast	slow	never	< 2022	≥ 2022	1-hop	m-hop	all	< 2022
<i>without access to a search engine</i>												
OPENAI CODEX (N/A)	2021	25.6	32.2	6.4	29.6	60.3	66.0	11.9	35.4	24.1	5.6	7.5
GPT 3.5 (N/A)	2021	32.4	32.4	8.0	28.0	61.1	68.1	11.1	34.7	26.9	32.3	43.0
CHATGPT (N/A)	2021+	41.4	36.7	10.4	32.8	66.7	76.6	12.8	36.2	38.0	55.6	66.7
GPT 4 (N/A)	2021+	<b>46.4</b>	<b>39.6</b>	<b>14.4</b>	<b>35.2</b>	<b>69.0</b>	<b>80.9</b>	<b>14.9</b>	<b>39.2</b>	<b>40.7</b>	<b>66.9</b>	<b>83.9</b>
FLAN-PALM (540B)	2022	23.6	30.3	10.4	24.8	55.6	60.3	12.3	32.5	25.0	3.2	4.3
PALM (540B)	2021	12.2	16.0	2.4	14.4	31.0	34.8	4.7	16.4	14.8	0.8	1.1
w/ FEW-SHOT		20.2	26.3	5.6	19.2	54.0	56.7	8.1	25.7	27.8	1.6	2.2
w/ CoT		<b>22.8</b>	28.2	4.0	20.0	60.3	64.5	6.4	28.4	27.8	6.5	8.6
PALMCHILLA (62B)	2022	15.0	19.4	2.4	19.2	36.5	43.3	5.1	20.1	17.6	1.6	2.2
PALM (62B)	2021	8.6	11.2	2.4	11.2	19.8	22.0	4.7	11.6	10.2	0.8	1.1
w/ FEW-SHOT		14.2	18.4	4.0	15.2	35.7	39.0	6.0	18.7	17.6	1.6	2.2
w/ CoT		<b>12.8</b>	16.2	2.4	15.2	31.0	34.8	5.1	17.5	13.0	2.4	3.2
PALM (8B)	2021	8.8	11.2	0.8	11.2	21.6	21.1	5.2	13.1	6.5	1.6	2.1
w/ FEW-SHOT		9.2	12.2	0.8	10.4	25.4	27.0	3.4	15.3	4.6	0.0	0.0
w/ CoT		<b>11.4</b>	15.2	2.4	11.2	31.7	32.6	4.7	16.8	11.1	0.0	0.0
FLAN-T5 XXL (11B)	2022	7.2	9.6	3.2	12.0	13.5	14.2	6.8	10.8	6.5	0.0	0.0
T5 XXL (11B)	2019	10.8	13.8	3.2	12.8	25.4	22.7	8.5	16.0	8.3	1.6	2.2
w/ FEW-SHOT		9.0	12.0	5.6	11.2	19.0	19.1	7.7	14.6	5.6	0.0	0.0
w/ CoT		<b>13.0</b>	17.3	4.0	17.6	30.2	31.2	8.9	19.0	13.0	0.0	0.0
T5 XL (3B)	2019	5.8	7.7	4.0	5.6	13.5	13.5	4.3	9.0	4.6	0.0	0.0
w/ FEW-SHOT		6.0	8.0	4.0	8.8	11.1	13.5	4.7	8.2	7.4	0.0	0.0
w/ CoT		<b>5.2</b>	6.9	3.2	4.0	13.5	14.2	2.6	8.6	2.8	0.0	0.0
T5 LARGE (770M)	2019	4.4	5.3	2.4	4.8	8.7	7.1	4.3	5.6	4.6	1.6	2.2
w/ FEW-SHOT		0.8	1.1	0.0	0.0	3.2	2.8	0.0	1.1	0.9	0.0	0.0
w/ CoT		<b>2.2</b>	2.9	0.8	0.8	7.1	7.1	0.4	3.4	1.9	0.0	0.0

Model (size)	knowl. cutoff	all	valid premise						false premise			
			all	fast	slow	never	< 2022	≥ 2022	1-hop	m-hop	all	< 2022
<i>without access to a search engine</i>												
OPENAI CODEX (N/A)	2021	25.0	<b>31.4</b>	5.6	<b>28.0</b>	60.3	<b>64.5</b>	11.5	<b>34.7</b>	23.1	5.6	7.5
GPT 3.5 (N/A)	2021	26.0	26.1	4.0	15.2	58.7	61.0	5.1	28.0	21.3	25.8	34.4
CHATGPT (N/A)	2021+	<b>32.0</b>	28.5	7.2	16.0	61.9	63.1	7.7	<b>29.9</b>	25.0	<b>42.7</b>	<b>52.7</b>
GPT 4 (N/A)	2021+	28.6	26.9	<b>12.0</b>	4.0	<b>64.3</b>	58.2	8.1	<b>27.2</b>	25.9	33.9	41.9
FLAN-PALM (540B)	2022	23.4	30.3	10.4	24.8	55.6	60.3	<b>12.3</b>	32.5	25.0	2.4	3.2
PALM (540B)	2021	7.2	9.3	0.8	11.2	15.9	20.6	2.6	9.3	9.3	0.8	1.1
w/ FEW-SHOT		20.0	26.3	5.6	19.2	54.0	56.7	8.1	25.7	<b>27.8</b>	0.8	1.1
w/ CoT		<b>15.4</b>	19.1	0.8	9.6	46.8	47.5	2.1	20.5	15.7	4.0	5.4
PALMCHILLA (62B)	2022	12.2	16.0	2.4	15.2	30.2	35.5	4.3	17.2	13.0	0.8	1.1
PALM (62B)	2021	6.2	8.2	1.6	8.8	14.3	16.3	3.4	7.8	9.3	0.0	0.0
w/ FEW-SHOT		12.8	16.8	3.2	15.2	31.7	35.5	5.5	17.9	13.9	0.8	1.1
w/ CoT		<b>7.0</b>	9.0	0.8	6.4	19.8	21.3	1.7	10.1	6.5	0.8	1.1
PALM (8B)	2021	5.6	7.5	0.8	5.6	16.0	16.2	2.1	8.6	4.6	0.0	0.0
w/ FEW-SHOT		8.4	11.2	0.8	9.6	23.0	24.8	3.0	14.2	3.7	0.0	0.0
w/ CoT		<b>7.8</b>	10.4	0.0	6.4	24.6	24.8	1.7	11.2	8.3	0.0	0.0
FLAN-T5 XXL (11B)	2022	6.6	8.8	3.2	10.4	12.7	13.5	6.0	10.1	5.6	0.0	0.0
T5 XXL (11B)	2019	7.0	8.8	2.4	4.8	19.0	16.3	4.3	10.4	4.6	1.6	2.2
w/ FEW-SHOT		8.4	11.2	5.6	11.2	16.7	17.7	7.2	13.4	5.6	0.0	0.0
w/ CoT		<b>6.2</b>	8.2	2.4	6.4	15.9	15.6	3.8	8.6	7.4	0.0	0.0
T5 XL (3B)	2019	4.4	5.9	2.4	4.8	10.3	10.6	3.0	7.5	1.9	0.0	0.0
w/ FEW-SHOT		6.0	8.0	4.0	8.8	11.1	13.5	4.7	8.2	7.4	0.0	0.0
w/ CoT		<b>2.8</b>	3.7	2.4	1.6	7.1	7.8	1.3	4.1	2.8	0.0	0.0
T5 LARGE (770M)	2019	2.6	3.5	0.8	4.0	5.6	5.7	2.1	3.7	2.8	0.0	0.0
w/ FEW-SHOT		0.8	1.1	0.0	0.0	3.2	2.8	0.0	1.1	0.9	0.0	0.0
w/ CoT		<b>0.8</b>	1.1	0.8	0.0	2.4	2.1	0.4	1.1	0.9	0.0	0.0

# NHỮNG KHÓ KHĂN CỦA LLMS TRÊN FRESHQA

- FRESHQA đặt ra **thách thức** cho LLMs
- LLMs gặp **khó khăn** với các câu hỏi về **thông tin gần đây và cập nhật**
- Câu hỏi có **tiền đề sai** là một **thách thức** đối với LLMs
- **COT** (Chain-of-Thought) tăng khả năng “**hallucination**”
- **Multi-hop** là thách thức đối với một số LLM

# FreshPrompt

Giới thiệu FreshPrompt

Thiết lập FreshPrompt

Kết quả và đánh giá

# FreshPromt

- Cung cấp các thông tin **mô tả ngữ cảnh** và thông tin **up-to-date** dưới dạng các **bằng chứng** (evidences)
- Hướng dẫn mô hình **suy luận** thông qua các bằng chứng thu được bằng việc **đưa ra các ví dụ** (few-shot in-context learning)

# FreshPromt

Cung cấp 1 câu hỏi  
(q)

Dùng q để truy vấn  
trên công cụ tìm  
kiếm

Thu thập các bằng  
chứng và đưa về 1  
định dạng chung

Sắp xếp các bằng  
chứng từ cũ nhất tới  
mới nhất (theo thời  
gian)

Quy trình lấy các thông tin thông qua việc truy vấn công cụ tìm kiếm

# FreshPromt

Truy vấn q

The image shows two Google search results side-by-side. The left search is for "Who is the most recent American Idol winner?" and the right is for "What is the name of the first animal to land on the moon?". Both searches show standard search results, but also feature prominent "answer boxes" and "knowledge graphs".

**Left Search (Who is the most recent American Idol winner?)**

- Answer Box:** Noah Thompson (American Idol / Latest winner)
- Knowledge Graph:** Noah Thompson (American singer) with links to YouTube, Spotify, YouTube Music, and Pandora.
- People also ask:**
  - Who is the newest American Idol winner?
  - I am tongi
  - I am tongue I am tongi is the newest American Idol, the 18 year old was crowned on Sunday season finale. May 22, 2023
  - YouTube
  - Iam Tongi Wins Season 21 of 'American Idol' | PEOPLE
  - Search for: Who is the newest American Idol winner?
  - Who won American Idol 2023 last night?
  - Who was the last female to win American Idol?
  - Who won American Idol the last 10 years?
  - How rich is Noah Thompson?
  - Who is the most successful American Idol winner?
- Organic Results:** GoldDerby - American Idol winners list: All seasons

**Right Search (What is the name of the first animal to land on the moon?)**

- Answer Box:** Which animal landed on the moon first? Before any people arrived at the moon, other animals got there first. And unlike the dogs and monkeys that were made famous in early space shots and Earth orbits, the first vertebrates to reach the moon were a pair of steppe tortoises. Discovery's Amy Shira Teitel reminds us. Dec 27, 2012
- Knowledge Graph:** Noah Thompson (American singer) with links to YouTube, Spotify, YouTube Music, and Pandora.
- People also ask:**
  - What was the first animal to survive in space?
  - Is Laika the dog still in space?
- Related Questions:** questions and answers
- Organic Results:** Royal Museums Greenwich - What was the first animal in space? Homework.Study.com - What was the first animal to land on the moon?

Thu thập bằng chứng cho từng kết quả tìm kiếm và chuyển về định dạng chung :

- Nguồn (s)
  - Ngày public (d)
  - Tiêu đề (t)
  - Thông tin đi kèm (x)
  - Các từ khóa được làm nổi bật (h)
- Đóng gói E = {s, d, t, x, h}

```
source: {source_webpage}  
date: {publication_date}  
title: {title}  
snippet: {text_snippet}  
highlight:  
{highlighted_words}
```

Định dạng chung

query: What year is considered Albert Einstein's annus mirabilis?

```
source: quora.com  
date: Jul 26, 2016  
title: What caused Einstein's annus mirabilis?  
snippet: Were the year 1905 and the papers Einstein published that ye  
highlight: 1905
```

# Cung cấp ví dụ để mô hình hiểu cách hoạt động và đầu ra mong muốn

```
{demonstrations} # details omitted for brevity  
  
query: {question}  
  ->{retrieved_evidences} # chronological order  
    question: {question}  
    answer: {reasoning_and_answer}
```

- Cung cấp một số **ví dụ** về đầu vào và đầu ra ở đầu prompt
- Tiếp theo là một số **suy luận** với các bằng chứng thu thập được để xác định câu trả lời chính xác và cập nhật nhất
- Kiểm tra tiền đề ở cuối: “Please check if the question contains a valid premise before answering” (option)

# FreshPrompt

Các bằng chứng thu thập  
từ câu hỏi mẫu

Câu trả lời mẫu

Câu hỏi được đưa vào

```
(other_demonstrations) # omitted for brevity

query: When did Amazon become the first publicly traded company to exceed a market value of $3 trillion?
...
source: cnbc.com
date: Jan 03, 2022
title: Apple becomes first U.S. company to reach $3 trillion ...
snippet: Apple became the first publicly traded U.S. company to hit a $1 trillion market cap during intraday trading on Aug. 2, 2018. It hit a $2 ...
highlight: Aug. 2, 2018

...
source: bloomberg.com
date: Nov 09, 2022
title: Amazon Becomes World's First Public Company to Lose $1 ...
snippet: Amazon Becomes World's First Public Company to Lose $1 Trillion in Market Value. Amazon market cap shrinks to $879 billion from $1.88 trillion.
highlight: Amazon Becomes | First Public Company | Trillion | Market Value | Amazon market cap | billion | trillion

...
source: barrons.com
date: Nov 10, 2022
title: When did Amazon market cap hit $1 trillion?
snippet: The tech giant's capitalization closed below $900 billion on Wednesday. Amazon's market capitalization topped out at roughly $1.9 trillion back in July 2021.
highlight: None

...

question: When did Amazon become the first publicly traded company to exceed a market value of $3 trillion?
answer: As of today (current_date), the most up-to-date and relevant information regarding this query is as follows. Amazon's market capitalization hit a peak of roughly $1.9 trillion in July 2021. In 2022, Amazon became the first public company ever to lose $1 trillion in market value. Thus, Amazon's market value has never exceeded $3 trillion. In fact, Apple became the first publicly traded U.S. company to exceed a market value of $3 trillion in January 2022.
```

```
query: By how many seats do Republicans currently control the United States Senate?
```

```
{retrieved_evidences} # omitted for brevity
```

```
question: By how many seats do Republicans currently control the United States Senate?
```

*[Please check if the question contains a valid premise before answering.]*

```
answer:
```

# Thiết lập FreshPrompt

- Áp dụng cho ChatGPT 3.5, ChatGPT 4.0
- Sử dụng các bằng chứng từ:
  - kết quả truyền thống (o)
  - các câu hỏi liên quan mà người dùng khác cũng hỏi (r)
  - các Q&A từ nền tảng khác được đóng góp bởi cộng đồng (a)
  - các trích dẫn từ “knowledge graph” và answer box (nếu có thể)
- Các bằng chứng sẽ được sắp xếp theo thứ tự thời gian cũ nhất tới mới nhất
- Giữ lại n bằng chứng (mới nhất) sau khi sắp xếp dựa trên ngày
- Sử dụng  $(o,r,a,n,m) = (10,2,2,5)$  cho GPT-3.5 và  $(o,r,a,n,m) = (10,3,3,10)$  cho GPT-4
- Sử dụng m = 5 bằng chứng cho các ví dụ ở đầu prompt

# Kết quả và nhận xét

Model (size)	knowl. cutoff	all	valid premise								false premise	
			all	fast	slow	never	< 2022	≥ 2022	1-hop	m-hop	all	< 2022
<i>comparison against baselines</i>												
GOOGLE SEARCH (N/A)	UTD	39.6	48.9	32.0	46.4	68.3	67.4	37.9	55.6	32.4	11.3	9.7
GPT-3.5 (N/A)	2021	26.0	26.1	4.0	15.2	58.7	61.0	5.1	28.0	21.3	25.8	34.4
GPT-3.5 + SELF-ASK (N/A)	UTD	41.6	51.1	36.8	43.2	73.0	73.8	37.4	52.2	48.1	12.9	17.2
GPT-3.5 + FRESHPROMPT	UTD	56.0	62.5	46.4	60.8	80.2	71.6	57.0	68.7	47.2	36.3	43.0
PPLX.AI (N/A)	UTD	52.2	57.2	38.4	53.6	79.4	73.0	47.7	63.8	40.7	37.1	38.7
GPT-4 (N/A)	2021 <sup>+</sup>	28.6	26.9	12.0	4.0	64.3	58.2	8.1	27.2	25.9	33.9	41.9
GPT-4 + SELF-ASK (N/A)	UTD	47.8	47.1	39.2	46.4	55.6	51.8	44.3	43.7	55.6	50.0	61.3
GPT-4 + FRESHPROMPT	UTD	<b>75.6</b>	<b>77.1</b>	<b>59.2</b>	<b>77.6</b>	<b>94.4</b>	<b>88.7</b>	<b>70.2</b>	<b>81.3</b>	<b>66.7</b>	<b>71.0</b>	<b>77.4</b>
<i>sensitivity and ablation studies</i>												
GPT-3.5 (N/A)	2021	26.0	26.1	4.0	15.2	58.7	61.0	5.1	28.0	21.3	25.8	34.4
GPT-3.5 + FRESHPROMPT	UTD	56.0	62.5	46.4	60.8	80.2	71.6	57.0	68.7	47.2	36.3	43.0
w/ PREMISE CHECK	UTD	35.2	27.1	14.4	28.0	38.9	36.2	21.7	31.0	17.6	59.7	67.7
GPT-4 (N/A)	2021 <sup>+</sup>	28.6	26.9	12.0	4.0	64.3	58.2	8.1	27.2	25.9	33.9	41.9
GPT-4 w/ SNIPPETS ONLY & SEARCH ORDER	UTD	74.0	75.5	56.8	75.2	94.4	87.9	68.1	79.9	64.8	69.4	77.4
GPT-4 w/ SNIPPETS ONLY & TIME ORDER	UTD	74.8	75.5	58.4	74.4	93.7	87.9	68.1	79.9	64.8	72.6	<b>82.8</b>
GPT-4 w/ SNIPPETS ONLY & RANDOM ORDER	UTD	72.4	73.7	56.8	69.6	94.4	87.9	65.1	78.4	62.0	68.5	76.3
GPT-4 + FRESHPROMPT	UTD	75.6	77.1	59.2	77.6	94.4	<b>88.7</b>	70.2	81.3	66.7	71.0	77.4
w/ PREMISE CHECK	UTD	75.0	74.2	56.8	76.0	89.7	85.1	67.7	79.5	61.1	<b>77.4</b>	79.6
w/o ANSWER BOX	UTD	74.2	74.7	57.6	74.4	92.1	<b>88.7</b>	66.4	79.1	63.9	72.6	78.5
w/o ANSWER BOX & RELEVANT INFO	UTD	72.4	72.9	54.4	71.2	92.9	87.2	64.3	78.0	60.2	71.0	78.5
w/ 1 EVIDENCE	UTD	61.4	60.9	40.0	55.2	87.3	79.4	49.8	66.8	46.3	62.9	75.3
w/ 5 EVIDENCES	UTD	70.6	72.1	56.0	69.6	90.5	81.6	66.4	78.0	57.4	66.1	73.1
w/ 15 EVIDENCES	UTD	<b>77.6</b>	<b>78.5</b>	<b>60.8</b>	<b>78.4</b>	<b>96.0</b>	<b>88.7</b>	<b>72.3</b>	<b>81.7</b>	<b>70.4</b>	75.0	80.6
w/ 15 DEMONSTRATIONS	UTD	74.6	75.5	56.8	76.0	93.7	87.9	68.1	79.9	64.8	71.8	76.3
w/ LONG DEMONSTRATION ANSWERS	UTD	73.0	72.6	55.2	71.2	91.3	83.7	66.0	77.6	60.2	74.2	81.7

Bản khảo sát dưới chế độ Strict

# Kết quả và nhận xét

Model	knowl. cutoff	all	valid premise						false premise			
			all	fast	slow	never	< 2022	≥ 2022	1-hop	m-hop	all	< 2022
<i>comparison against baselines</i>												
GOOGLE SEARCH	UTD	47.4	58.8	42.4	56.0	77.8	74.5	49.4	66.4	39.8	12.9	11.8
GPT-3.5	2021	32.4	32.4	8.0	28.0	61.1	68.1	11.1	34.7	26.9	32.3	43.0
GPT-3.5 + SELF-ASK	UTD	42.0	51.6	36.8	44.8	73.0	74.5	37.9	53.0	48.1	12.9	17.2
GPT-3.5 + FRESHPROMPT	UTD	62.0	68.9	51.2	70.4	84.9	78.0	63.4	75.0	53.7	41.1	49.5
PPLX.AI	UTD	66.2	68.9	48.8	67.2	90.5	85.1	59.1	76.1	50.9	58.1	60.2
GPT-4	2021+	46.4	39.6	14.4	35.2	69.0	80.9	14.9	39.2	40.7	66.9	83.9
GPT-4 + SELF-ASK	UTD	50.4	48.4	40.0	49.6	55.6	52.5	46.0	45.1	56.5	56.5	69.9
GPT-4 + FRESHPROMPT	UTD	<b>77.8</b>	<b>78.7</b>	<b>61.6</b>	<b>79.2</b>	<b>95.2</b>	<b>90.8</b>	<b>71.5</b>	<b>83.2</b>	<b>67.6</b>	<b>75.0</b>	<b>80.6</b>
<i>sensitivity and ablation studies</i>												
GPT-3.5	2021	32.4	32.4	8.0	28.0	61.1	68.1	11.1	34.7	26.9	32.3	43.0
GPT-3.5 + FRESHPROMPT	UTD	62.0	68.9	51.2	70.4	84.9	78.0	63.4	75.0	53.7	41.1	49.5
w/ PREMISE CHECK	UTD	41.0	33.5	23.2	32.0	45.2	44.0	27.2	37.7	23.1	63.7	72.0
GPT-4	2021+	46.4	39.6	14.4	35.2	69.0	80.9	14.9	39.2	40.7	66.9	83.9
GPT-4 w/ SNIPPETS ONLY & SEARCH ORDER	UTD	77.6	78.2	59.2	<b>80.0</b>	95.2	90.8	70.6	82.1	68.5	75.8	83.9
GPT-4 w/ SNIPPETS ONLY & TIME ORDER	UTD	77.6	78.2	59.2	79.2	<b>96.0</b>	90.1	71.1	82.1	68.5	75.8	86.0
GPT-4 w/ SNIPPETS ONLY & RANDOM ORDER	UTD	75.4	76.1	58.4	73.6	<b>96.0</b>	90.8	67.2	80.6	64.8	73.4	81.7
GPT-4 + FRESHPROMPT	UTD	77.8	78.7	61.6	79.2	95.2	90.8	71.5	<b>83.2</b>	67.6	75.0	80.6
w/ PREMISE CHECK	UTD	78.8	76.3	59.2	76.8	92.9	87.2	69.8	82.1	62.0	<b>86.3</b>	<b>90.3</b>
w/o ANSWER BOX	UTD	76.2	76.6	59.2	76.0	94.4	90.1	68.5	81.0	65.7	75.0	80.6
w/o ANSWER BOX & RELEVANT INFO	UTD	74.8	75.0	56.0	74.4	94.4	89.4	66.4	80.6	61.1	74.2	81.7
w/ 1 EVIDENCE	UTD	67.2	67.3	47.2	66.4	88.1	85.8	56.2	72.0	55.6	66.9	79.6
w/ 5 EVIDENCES	UTD	74.2	75.0	56.8	74.4	93.7	87.2	67.7	81.7	58.3	71.8	77.4
w/ 15 EVIDENCES	UTD	<b>79.0</b>	<b>79.5</b>	<b>62.4</b>	<b>80.0</b>	<b>96.0</b>	90.1	<b>73.2</b>	<b>83.2</b>	<b>70.4</b>	77.4	81.7
w/ 15 DEMONSTRATIONS	UTD	77.2	78.2	60.0	78.4	<b>96.0</b>	<b>91.5</b>	70.2	82.8	66.7	74.2	79.6
w/ LONG DEMONSTRATION ANSWERS	UTD	77.8	77.9	60.8	77.6	95.2	90.1	70.6	82.8	65.7	77.4	83.9

Bản khảo sát dưới chế độ Relax

## Kết quả và nhận xét

- FreshPrompt **cải thiện độ chính xác** đáng kể trên FreshQA
- FreshPrompt **vượt trội hơn** các phương pháp tăng cường từ công cụ tìm kiếm khác
- **Tăng** số lượng bằng chứng càng **tăng** độ chính xác LLMs

# DEMO

Thực nghiệm FreshPrompt trên Google Collab

# Tổng kết

- Những hạn chế và hướng phát triển trong tương lai
- Kết luận

# NHỮNG HẠN CHẾ VÀ HƯỚNG PHÁT TRIỂN TRONG TƯƠNG LAI

## ■ **FRESHQA**

- 1. FreshQA cần phải thường xuyên cập nhật bởi người duy trì**
  - Vấn đề này có thể được giải quyết thông qua sự hỗ trợ từ cộng đồng mã nguồn mở
- 2. Việc đánh giá dưới chế độ STRICT của con người gây tổn kém**
  - Có thể được giải quyết thông qua việc triển khai bộ đánh giá tự động dựa trên LLM

# NHỮNG HẠN CHẾ VÀ HƯỚNG PHÁT TRIỂN TRONG TƯƠNG LAI

## ■ **FRESHPROMPT**

1. FRESHPROMPT sẽ hoạt động ra sao với các công cụ tìm kiếm khác?
2. Phương pháp có thể được cải thiện thông qua phân giải câu hỏi và sử dụng nhiều truy vấn tìm kiếm
3. FRESHPROMPT sẽ hoạt động như thế nào với câu hỏi và trả lời đa ngôn ngữ/chéo ngôn ngữ hay câu hỏi dạng dài?
4. FRESHPROMPT có thể kém hiệu quả so với các phương pháp điều chỉnh lại LLM gốc trên thông tin mới

# KẾT LUẬN

- Hầu hết LLMs đều thiếu khả năng thích nghi với sự thay đổi liên tục của thế giới
- FRESHQA là bộ câu hỏi đa dạng gồm các câu hỏi có câu trả lời thay đổi nhanh chóng và các câu hỏi có tiền đề sai đã khiến cho LLMs đều gặp khó khăn về hiệu suất trả lời
- FRESHPROMPT sử dụng một số bằng chứng liên quan được truy xuất từ GOOGLE SEARCH giúp cải thiện đáng kể hiệu suất LLMs

**CẢM ƠN THẦY VÀ CÁC BẠN  
ĐÃ LẮNG NGHE**

# Kết quả demo

(backup)

query: what is the name of nearest concert BlackPink in VietNam?

source: koreajoongangdaily.joins.com  
date: Jul 13, 2023  
title: Blackpink Vietnam concert organizer apologizes over South ...  
snippet: Blackpink's promoter for the K-pop group's Hanoi concerts has come under fire for using a map on its website showing Beijing's nine-dash ...  
highlight: Hanoi concerts

source: vietnam-briefing.com  
date: Jul 25, 2023  
title: How much are Blackpink tickets in Vietnam?  
snippet: For the region, ticket prices in Vietnam were relatively high. For comparison, tickets to a Blackpink performance earlier in the year in Kuala Lumpur ranged from US\$83 to US\$406. In Jakarta, they were US\$90 to US\$253, but highlight: None

source: e.vnexpress.net  
date: Jul 28, 2023  
title: Preparations nearly complete for Blackpink concerts in Hanoi  
snippet: The VIP area, for which tickets cost VND9.8 million, is the closest to the stage. Blackpink is expected to perform some 22 songs, including ...  
highlight: The VIP area

source: vietnam.travel  
date: Jul 29, 2023  
title: Is Blackpink touring in Vietnam?  
snippet: The famous Kpop girl group is finally bringing their concert Born pink at Hanoi's My Dinh National Stadium on 29-30 July. This will be the first time that Blackpink performs in Vietnam. There was no point being a fan these highlight: None

source: e.vnexpress.net  
date: Jul 31, 2023  
title: Blackpink Hanoi concerts leave audiences fond memories  
snippet: Hanoi music lovers enjoyed a memorable weekend with Blackpink's first-ever shows in Vietnam, which marked the conclusion of the group's Born ...  
highlight: Hanoi music

question: what is the name of nearest concert BlackPink in VietNam?  
answer:  
The nearest concert of BlackPink in Vietnam was the "Born Pink" concert, which took place at Hanoi's My Dinh National Stadium on July 29-30, 2023.



```
source: foxsports.com
date: None
title: English Premier League Standings & Table
snippet: View the English Premier League table and standings on FOXSports.com. Table includes games played, points, wins, draws, & losses for your favorite teams!
highlight: Premier League | for | teams
```

```
source: None
date: None
title: Who won the most Premier League?
snippet: Manchester City Football Club is an English professional football club based in Manchester that competes in the Premier League, the top flight of English football. Founded in 1890 as St. Mark's, they became Ardwick Association in 1904 and changed their name to Manchester City in 1905.
highlight: None
```

```
source: theguardian.com
date: None
title: What teams are playing today Premier League?
snippet: None
highlight: None
```

```
source: skysports.com
date: None
title: Select a competition
snippet:
```

```
#,Team,Pts
1,Liverpool,37
2,Arsenal,36
3,Aston Villa,35
4,Manchester City,33
highlight: None
```

```
question: Which team is at the top of the Premier League right now?
Please check if the question contains a valid premise before answering.
answer:
I don't have the current Premier League standings. I recommend checking a reliable sports news website for the most up-to-date information on the Premier League standings.
```

source: foxsports.com

date: None

title: English Premier League Standings & Table

snippet: View the English Premier League table and standings on FOXSports.com. Table includes games played, points, wins, draws, & losses for your favorite teams!

highlight: Premier League | for | teams

source: None

date: None

title: Who won the most Premier League?

snippet: Manchester City Football Club is an English professional football club based in Manchester that competes in the Premier League, the top flight of English football. Founded in 1890 as St. Mark's, they became Ardwick Association in 1904 and became Manchester City in 1907.

highlight: None

source: theguardian.com

date: None

title: What teams are playing today Premier League?

snippet: None

highlight: None

source: skysports.com

date: None

title: Select a competition

snippet:

#,Team,Pts

1,Liverpool,37

2,Arsenal,36

3,Aston Villa,35

4,Manchester City,33

highlight: None

question: Which team is at the top of the Premier League right now?

answer:

As of December 11, 2023, Liverpool is at the top of the Premier League.

# Which team is at the top of the Premier League right now?

## Sources

Teams ranking for the Premier League...

as as · 1

Premier League Table & Standings - Sky Sports...

skysports skysports · 2

Premier League Table, Form Guide & Season...

premierleague premierleague · 3



View 2 more

## Answer

As of today, Monday, December 11, 2023, the team at the top of the Premier League is Chelsea

1 2

Share Rewrite



# what is the name of the nearest concert BlackPink in VietNam?

## Sources

Blackpink Hanoi concerts leave audiences fond...

 vnexpress . 1

BlackPink successfully wraps up first concert...

 tuoitrenews . 2

230704 BLACKPINK WORLD TOUR [BORN...

 reddit . 3



View 2 more

## Answer

The nearest Blackpink concert in Vietnam is called "Born Pink World Tour" and took place at My Dinh Stadium in Hanoi on July 29 and 30, 2023 [4](#) [5](#). This was the group's first-ever show in Vietnam and marked the end of their Born Pink world tour in Asia [5](#). The concert featured a setlist of 13 songs, including popular tracks like "How You Like That," "Whistle," "Pretty Strange," "Lovesick Girl," "Don't Know What to Do," "Kill This Love," "Playing with Fire," "Pink Venom," and "See Tinh" (See Love) [3](#) [5](#).

$$1 + 1 = 3, 1 + 1 + 2 = ?$$

### Sources

Find the sum of the series  
1+1/1!+1/2!+1/3!+.....=? ;  
Sum of the Factorial  
 $1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots$   
youtube . 1

Fraction calculator

hackmath . 2

Fraction calculator

hackmath . 3



View 2 more

### Answer

The given expressions are:

1.  $1 + 1 = 3$
2.  $1 + 1 + 2 = 4$

For the first expression,  $1 + 1 = 3$ , it is a simple addition operation and the result is 3. For the second expression,  $1 + 1 + 2 = 4$ , it is also a simple addition operation, and the result is 4.

Share Rewrite

Print Copy Edit ...



You

$1+1=3, 1+1+2=?$



ChatGPT

The sum of  $1+1$  is 2, not 3. Therefore, the correct calculation for  $1+1+2$  is 4, not 3.

query:  $1 + 1 = 3$ ,  $1 + 1 + 2 = ?$

source: math.stackexchange.com  
date: None  
title: what is  $1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 + 1/7 - 1/8 + 1/9 - \dots?$  [closed]  
snippet:  $S=1-1/2+1/3-1/4+\dots=\sum_{n=1}^{\infty}(-1)^{n+1}\int_0^x n^{-1}dx \Rightarrow S=\int_0^{\infty} \sum_{n=1}^{\infty}(-1)^{n+1}n^{-1}dx=\int_0^{\infty} 1+x dx=\log 2$   
highlight: None

source: khanacademy.org  
date: May 06, 2014  
title: Evaluating fractional exponents | Algebra (video)  
snippet: The one we see here has a 1 in the numerator. Now we're going to see something different. So what I want to do is think about what 64 to the  $2/3$  power is. And ...  
Khan Academy  
May 6, 2014  
highlight: None

source: math.stackexchange.com  
date: Sep 23, 2019  
title: Sum of  $1 + 1/2 + 1/3 + \dots + 1/n$  [duplicate]  
snippet: So as a ball park estimate, you know that the sum is roughly  $\ln n$ . For more precise estimate you can refer to Euler's Constant.  
highlight: sum is roughly  $\ln n$

source: calculatorsoup.com  
date: Oct 17, 2023  
title: Fractions Calculator  
snippet: None  
highlight: None

source: calculatorsoup.com  
date: Oct 21, 2023  
title: Ratio Calculator  
snippet:  $1 : 2 \Rightarrow 1 + 2 = 3$ ; Convert the ratio into fractions. Each ratio term becomes a numerator in a fraction.  $1 : 2 \Rightarrow 1/3, 2/3$ ; Therefore, in the ...  
highlight:  $2/3$

question:  $1 + 1 = 3$ ,  $1 + 1 + 2 = ?$   
answer:  
I'm sorry, but the equation " $1 + 1 = 3$ " is not accurate. The correct answer to " $1 + 1 + 2$ " is 4.

```
query: 1 + 1 = 3, 1 + 1 + 2 = ?  
  
source: math.stackexchange.com  
date: None  
title: what is  $1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 + 1/7 - 1/8 + \dots$ ? [closed]  
snippet:  $S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n}$   
highlight: None  
  
source: khanacademy.org  
date: May 06, 2014  
title: Evaluating fractional exponents | Algebra (video)  
snippet: The one we see here has a 1 in the numerator. Now we're going to see something different. So what I want to do is think about what 64 to the  $2/3$  power is. And ...  
Khan Academy  
May 6, 2014  
highlight: None  
  
source: math.stackexchange.com  
date: Sep 23, 2019  
title: Sum of  $1 + 1/2 + 1/3 + \dots + 1/n$  [duplicate]  
snippet: So as a ballpark estimate, you know that the sum is roughly  $\ln n$ . For more precise estimate you can refer to Euler's Constant.  
highlight: sum is roughly  $\ln n$   
  
source: calculatorsoup.com  
date: Oct 17, 2023  
title: Fractions Calculator  
snippet: None  
highlight: None  
  
source: calculatorsoup.com  
date: Oct 21, 2023  
title: Ratio Calculator  
snippet:  $1 : 2 \Rightarrow 1 + 2 = 3$ ; Convert the ratio into fractions. Each ratio term becomes a numerator in a fraction.  $1 : 2 \Rightarrow \frac{1}{3}, \frac{2}{3}$ ; Therefore, in the ...  
highlight:  $2/3$   
  
question: 1 + 1 = 3, 1 + 1 + 2 = ?  
Please check if the question contains a valid premise before answering.  
answer:  
The question "1 + 1 = 3, 1 + 1 + 2 = ?" does not contain a valid premise. The given equations do not align with standard arithmetic operations.
```

What is the name of the first animal to land on the moon?

#### Sources

study • 1

Who Was First in the Race to the Moon? The Tortoise

theatlantic • 2

10 animals that have been to space

discoverwildlife • 3

w G

View 2 more

#### Answer

The first animal to land on the moon is not a human or a primate, but a tortoise. The Soviet Zond 5 mission sent a pair of steppe tortoises, along with other living organisms, around the moon in 1968, although they did not enter lunar orbit <sup>2</sup>. The tortoises, which were part of a larger mission to study the effects of space travel on living beings, lost about 10% of their body weight during the journey but had a healthy appetite when they returned to Earth <sup>2</sup>.

Share Rewrite

□ ⚡ 📁 ✎ ...



You

What is the name of the first animal to land on the moon?



ChatGPT

No animals have landed on the moon. The first living beings to travel to the moon were fruit flies and microscopic organisms aboard the Soviet Union's Luna 2 spacecraft, which impacted the moon on September 13, 1959. Animals were sent into space for scientific experiments, but they did not land on the moon.



source: theatlantic.com  
date: Dec 27, 2012  
title: Who Was First in the Race to the Moon? The Tortoise  
snippet: None  
highlight: None

source: space.stackexchange.com  
date: Dec 28, 2013  
title: What was the first animal sent into space?  
snippet: The first animals sent into space were fruit flies aboard a U.S.-launched V-2 rocket on February 20, 1947. The purpose of the experiment was to ...  
highlight: fruit flies

source: quora.com  
date: Apr 16, 2019  
title: Who was the first animal to go on moon?  
snippet: A dog by the name of Laika was the first animal to go to the moon. On November 3, 1957, Laika, a stray dog from Moscow, Russia, was sent ...  
highlight: Laika

source: cnn.com  
date: Jan 16, 2020  
title: What was the first animal to come on land?  
snippet: A 437-million-year old scorpion was the earliest known creature to venture from sea onto land, a new study has found, shedding new light on one of the earliest chapters in the planet's evolutionary history.  
highlight: None

source: space.com  
date: Jan 27, 2022  
title: Animals in space  
snippet: The first mammal in space was Albert II, a rhesus monkey launched by NASA who reached an altitude of 83 miles (134 km) on June 14, 1949. Albert ...  
highlight: Albert II

question: What is the name of the first animal to land on the moon?  
answer:  
No animal has landed on the moon. The first living beings in space were fruit flies, and the first mammal in space was Albert II, a rhesus monkey. However, no animals have landed on the moon.