

PARIS
21



AI and Data Workshop in Rwanda

Large Language Models

Shanghai Jiao Tong University

Aug 5–7, 2025

Kigali, Rwanda



Content

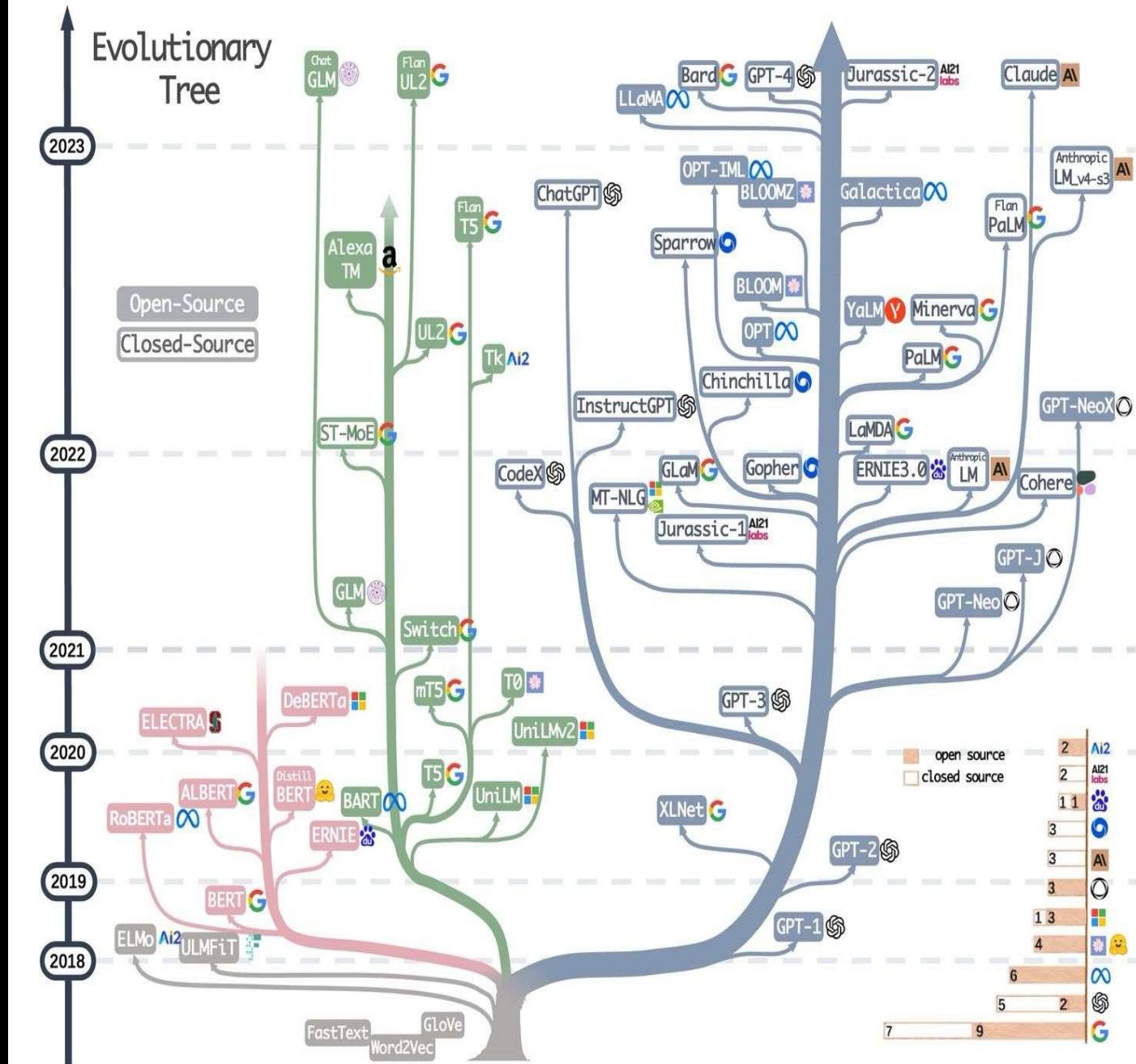
- Review: What's LLM ?
- Features and Comparison between LLMs.
- Practices and Challenges.



Section 1:

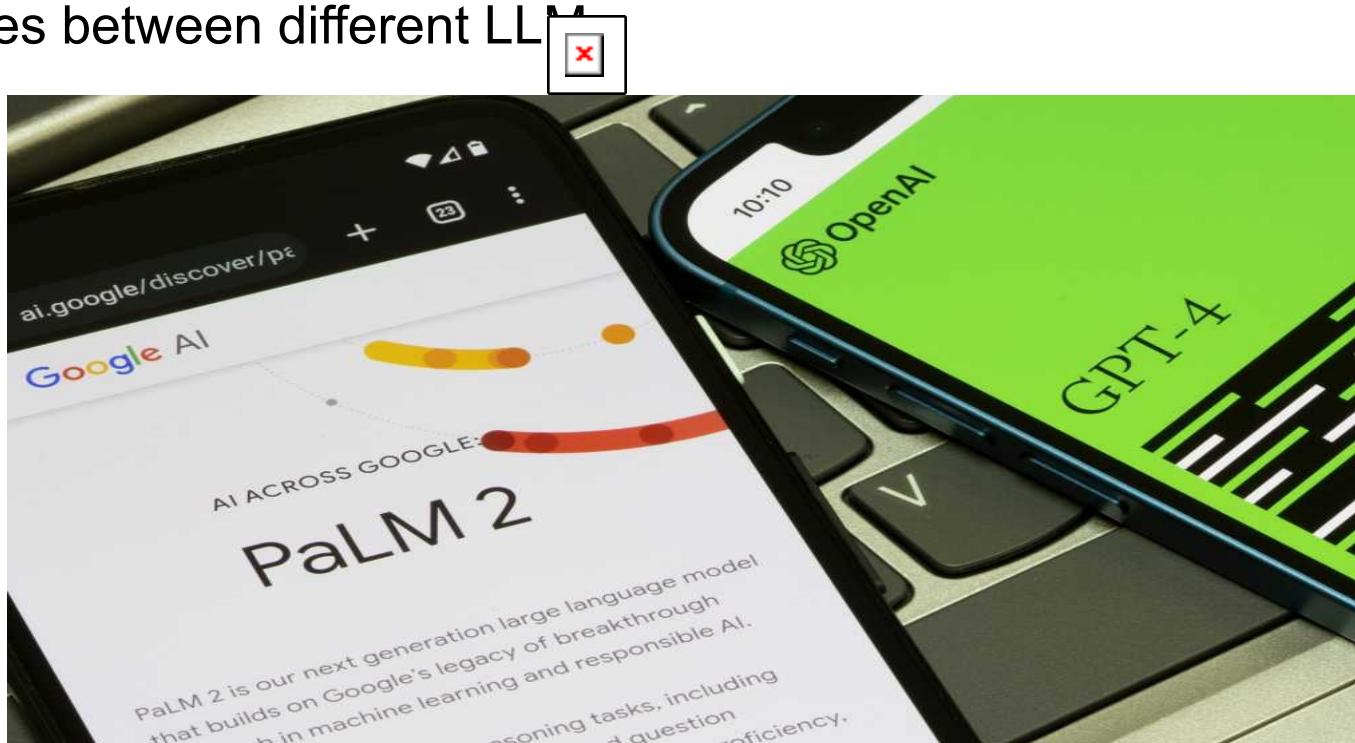
Review:

What's LLM?



What's LLM?

- LLM is short for Large Language Model, which is the most advanced and practical tool for writing, data analysis, coding and etc.
- Basically, it learns from the data and predicts the words according to the user's prompt.
- To specialize a model's ability in a specific field (coding, reasoning and etc.), Researchers adopt techniques like **pre-training**, **redesigning the inner mechanism** and **fine-tuning**, resulting in the differences between different LLMs



Section 2: Features and Comparison between LLMs

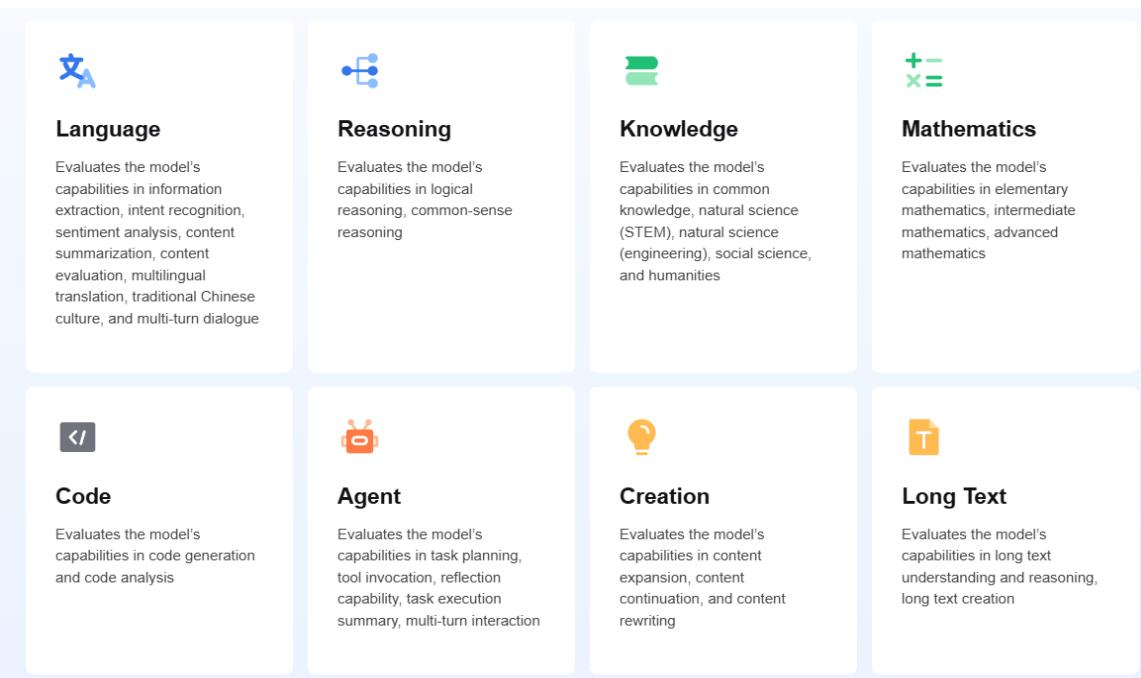


What matters to a LLM ?

- People use LLM in many different ways, we can divide the LLM ability in fields like **long text, coding, reasoning**

- Discussion: What do you think the ability that matters to LLMs in your daily tasks?

- Researchers evaluate LLMs by...
 - Designing test-problem (take exams like humans)
 - Rating by users (Arena)



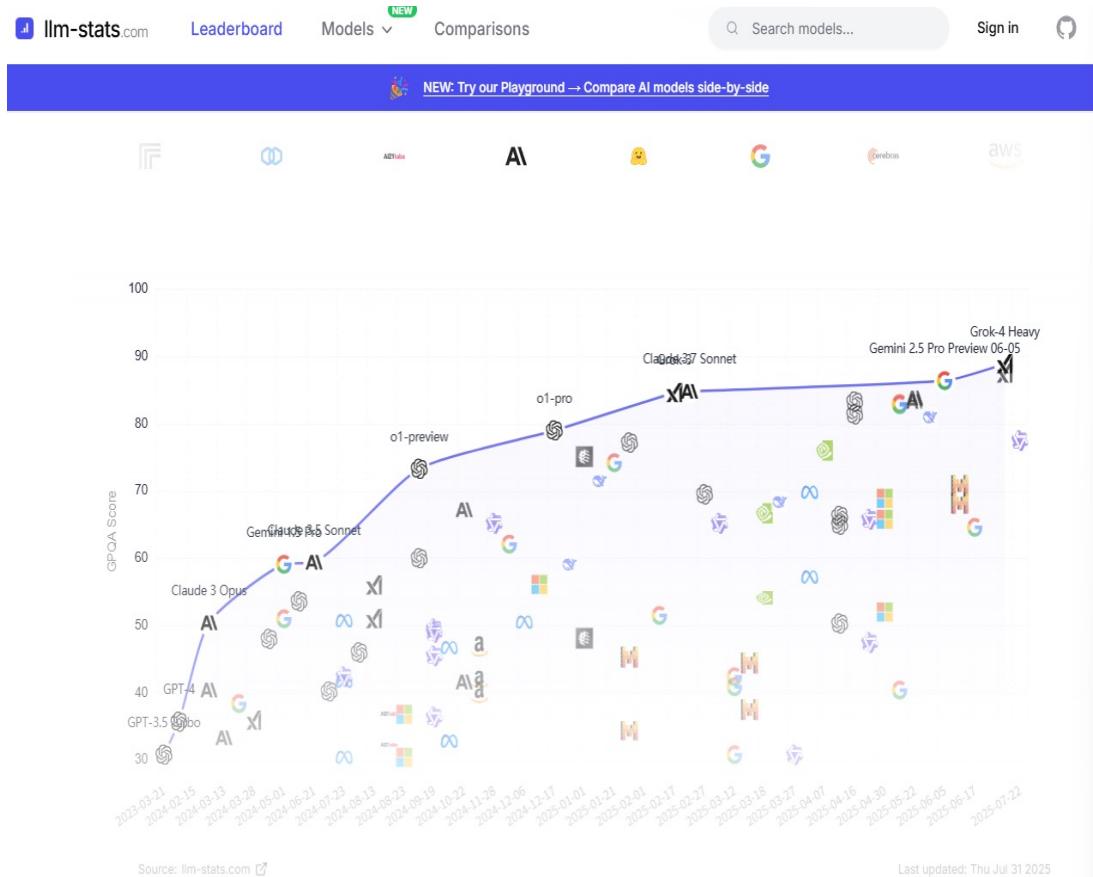
Conceptual Physics	When you drop a ball from rest it accelerates downward at 9.8 m/s^2 . If you instead throw it downward assuming no air resistance its acceleration immediately after leaving your hand is (A) 9.8 m/s^2 (B) more than 9.8 m/s^2 (C) less than 9.8 m/s^2 (D) Cannot say unless the speed of throw is given.	✓ ✗ ✗ ✗
College Mathematics	In the complex z -plane, the set of points satisfying the equation $z^2 = z ^2$ is a (A) pair of points (B) circle (C) half-line (D) line	✗ ✗ ✗ ✓

Figure 4: Examples from the Conceptual Physics and College Mathematics STEM tasks.

MMLU test-dataset problems



Leaderboard



LLM-stats: <https://llm-stats.com/>
Test-based. Can check out the SOTA and corresponding test score.

Text		2 days ago
Rank (UB) ↑ Model ↑		
1	gemini-2.5-pro	
1	o3-2025-04-16	
2	chatgpt-4o-latest-20250326	
3	gpt-4.5-preview-2025-02-27	
3	grok-4-0709	
5	claude-opus-4-20250514-think...	

WebDev		2 days ago
Rank (UB) ↑ Model ↑		
1	Gemini-2.5-Pro	
1	DeepSeek-R1-0528	
1	Claude Opus 4 (20250514)	
1	Qwen3-Coder	
4	Claude Sonnet 4 (20250514)	
4	Claude 3.7 Sonnet (20250219)	

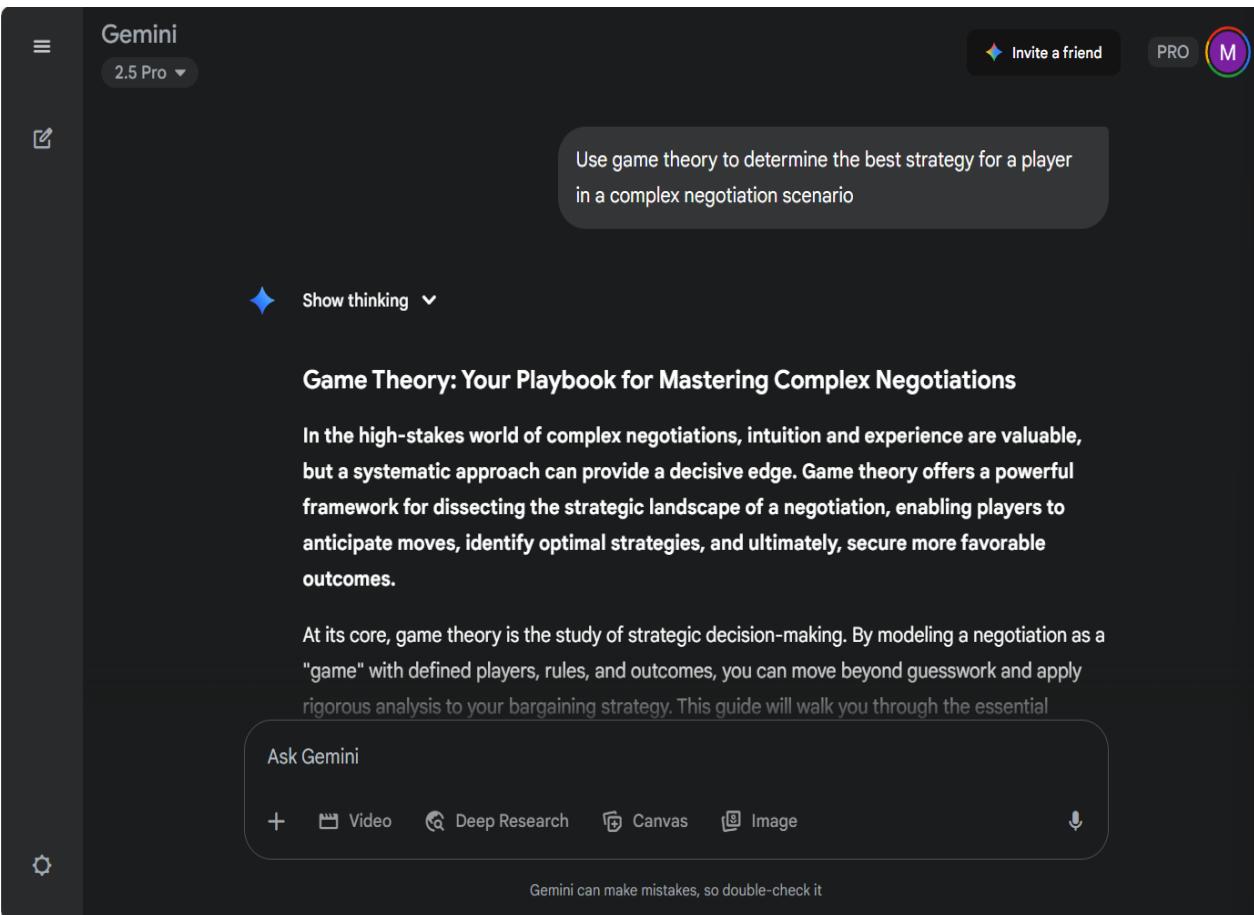
LMArena: <https://alpha.lmarena.ai/leaderboard>
Arena-based. Users rate the LLMs by comparing the answers. Closer to real situations



The mainstream LLM tool

Gemini

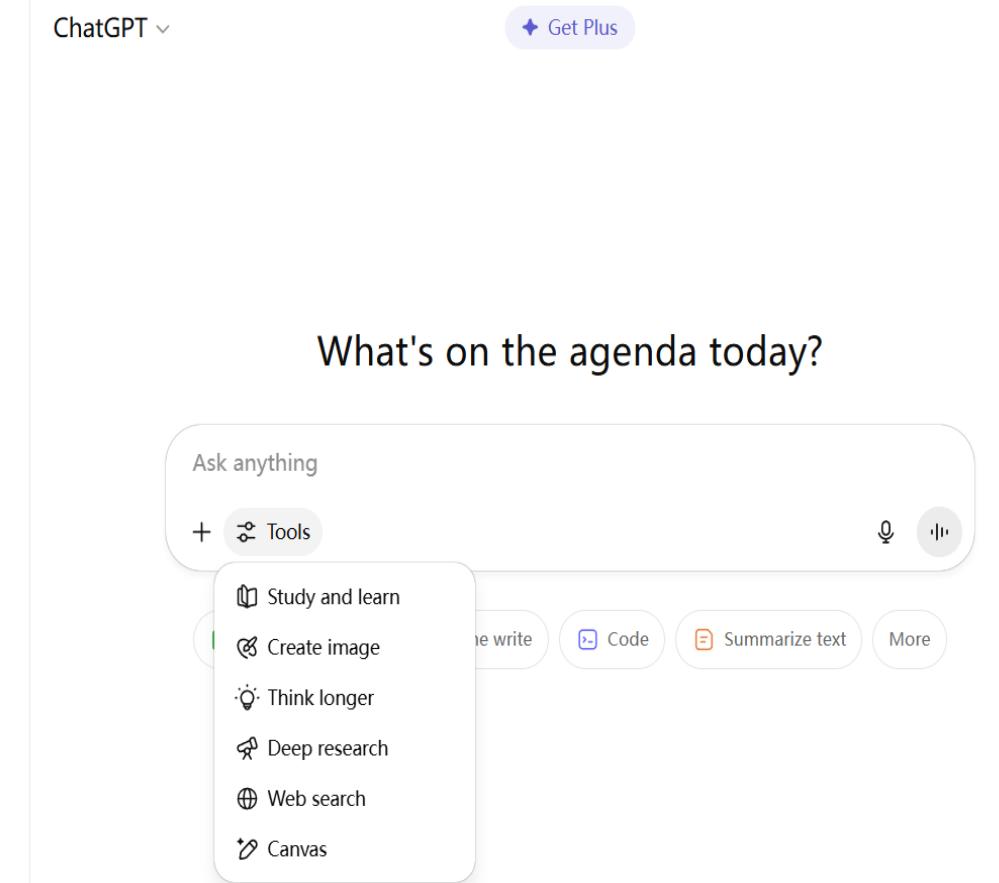
- Developed by Google. SOTA for me.
- Two models open to be used:
 - 2.5 pro: strong reasoning, coding and mathematical ability (but need to subscribe)
 - 2.5 flash: fast respond but weaker in reasoning.
- Complete assisting functions:
 - **Canvas: generating Google docs according to the prompt**
 - Deep Research: Conducting large-range searching and deep analysis
 - Generate codes during answering
 - Can execute Python code and give figures; Support files processing like pdf,xls,csv.....



The mainstream LLM tool

ChatGPT

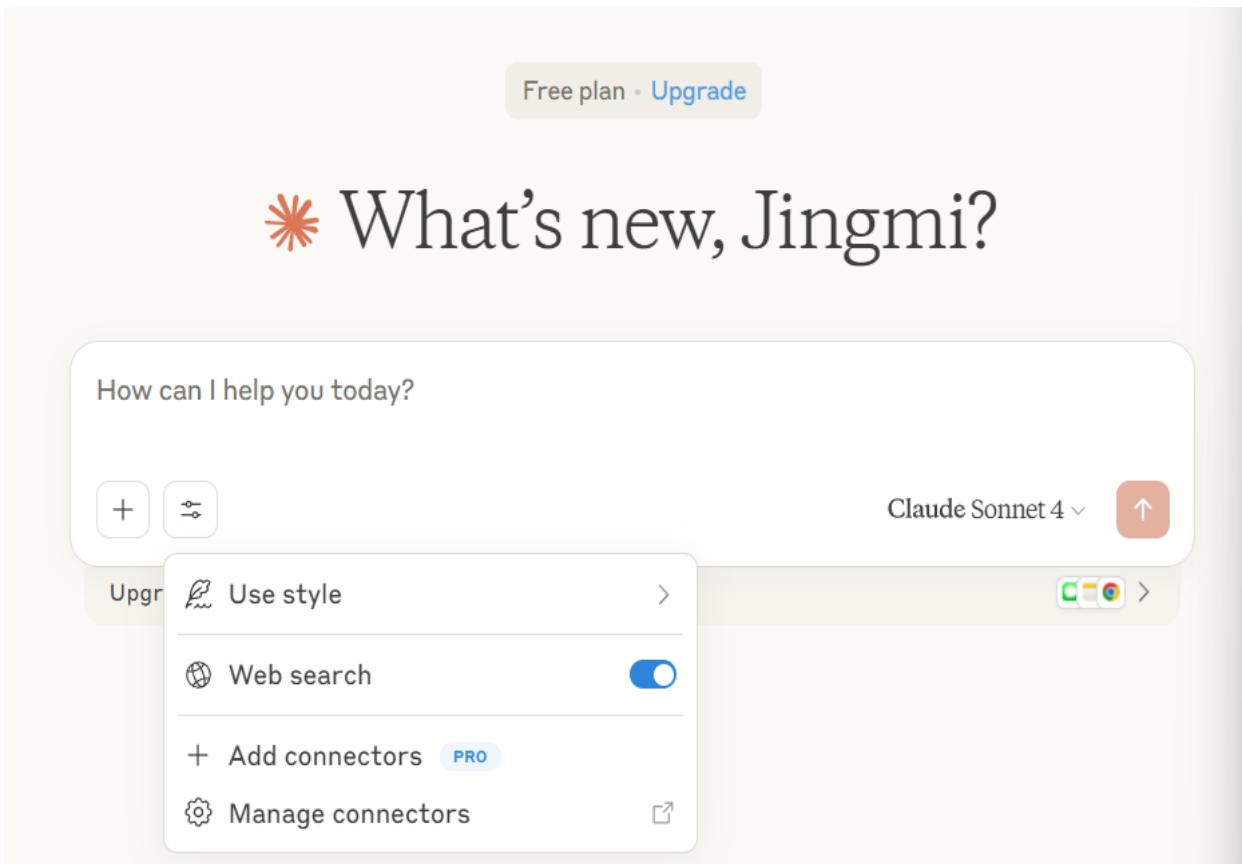
- Developed by OpenAI.
- A lots of models:
 - For free users: 40 mini only, basic reasoning and knowledge, but much slower respond
 - Stronger reasoning and less hallucination with higher-level models -- need to subscribe
- Complete assisting functions
 - Study and learn: new functions just released, give detailed instruction during answering like a teacher
 - Deep research, canvas, file-reading support
 - Python execution and figure generation



The mainstream LLM tool

Claude

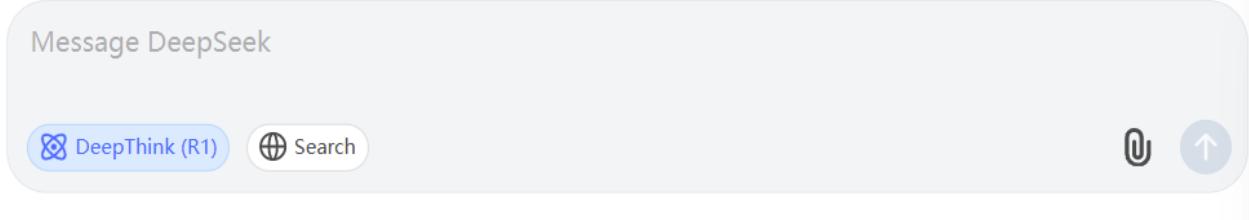
- Developed by Anthropic.
- Famous and Specialized in coding.
- Assisting functions:
 - A special function called ‘connector’, can use small plugins developed by others
 - Less file-processing support (eg. do not support xlsx)



The mainstream LLM tool

Deepseek

- Developed by Deepseek.
- Open source and free-- you can make your own deepseek
- Two models open to be used:
 - R1: strong reasoning, give the reasoning process during answering
 - o3 - fast searching but weaker in reasoning
- File support but no canvas and no figure generation



Section 3:

Practices and Challenges



Demostration: Analysis with Rwanda Agriculture Data

- Data Source: NISR - Agricultural Household Survey 2020
- Adopt Gemini-2.5 Flash to solve
 - Re-format the docs into clean data (Data with variables as columns)
 - Plot a figure with readable dataframe
 - Generate a Google docs using Canva in Gemini

List of tables

Table 53: Monthly lactating cows per Province, 2019/20 Agricultural Year

Month	Kigali	South	West	North	East	Total
2019						
September	5,512	43,214	29,803	27,903	44,354	150,787
October	5,812	41,380	29,751	26,354	45,040	148,337
November	5,532	42,655	29,518	27,004	46,046	150,754
December	5,545	43,952	33,665	24,604	50,937	158,703
2020						
January	6,250	46,005	36,483	24,480	50,153	163,371
February	6,053	46,963	35,100	25,695	53,423	167,235
March	6,665	49,210	36,528	30,377	47,005	169,784
April	6,582	52,958	38,512	31,106	46,694	175,852
May	6,356	49,645	37,626	30,665	44,999	169,290
June	6,054	46,083	38,007	31,915	43,134	165,194
July	6,089	45,012	35,643	34,044	44,380	165,169
August	6,080	46,074	39,529	40,392	45,384	177,459

Source: NISR, AHS 2020

Table 46 Table 47 Table 48 Table 49 Table 50 Table 51 Table 52 Table 53 Table 54 Table 55

Example: NISR Agricultural Household survey2020,
format as .xlsx



Try yourself !

■ Data Source: EICV7

■ Tasks

- Try out the LLMs mentioned above
- Find the appropriate LLM
- write prompt for LLM to read the sex ratio data in Page 20 (Table 1.2) and make it plotable ('tidy data')
 - hint: you may finish steps by steps using multiple prompts
- Plot a figure
- (optional) Generate a report and further analysis

Table 1.2: Sex Ratio⁵, by area of residence, province, and consumption quintile

EICV7	Sex ratio	Female population (000s)	Total
All Rwanda	91.8	7,064	13,549
Residence area			
Urban	93.0	1,990	3,841
Rural	91.3	5,074	9,708
Province			
City of Kigali	97.0	944	1,860
Southern	91.4	1,583	3,030
Western	89.0	1,537	2,906
Northern	90.8	1,100	2,099
Eastern	92.4	1,899	3,654
Quintile			
Q1	93.2	1,403	2,710
Q2	92.2	1,410	2,711
Q3	90.5	1,422	2,709
Q4	90.9	1,420	2,711
Q5	92.2	1,409	2,709

Source: NISR, EICV7

NISR EICV Table 1.2





Thank you!



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY