

New Directions in Generative AI for Strategic Growth

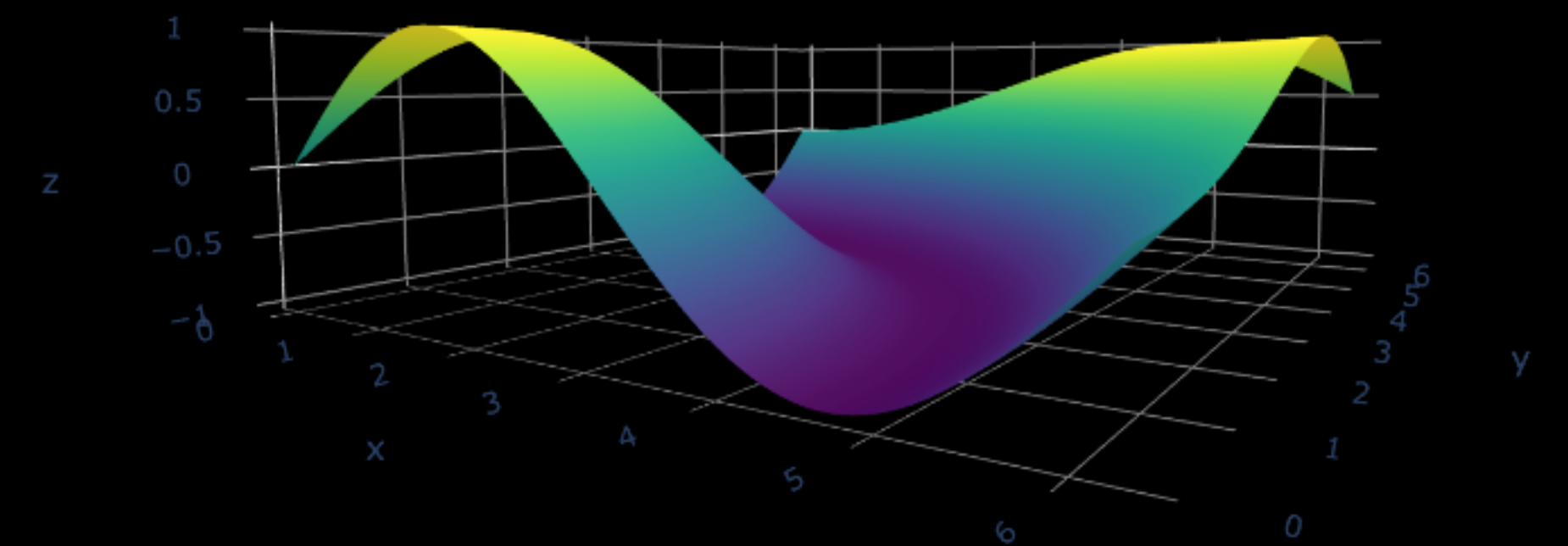
Monishwaran Maheswaran

UC Berkeley Master's Student (Also BS'23 EECS + Math)

Efficient Machine Learning

LLM Optimizations

Large Scale Machine Learning

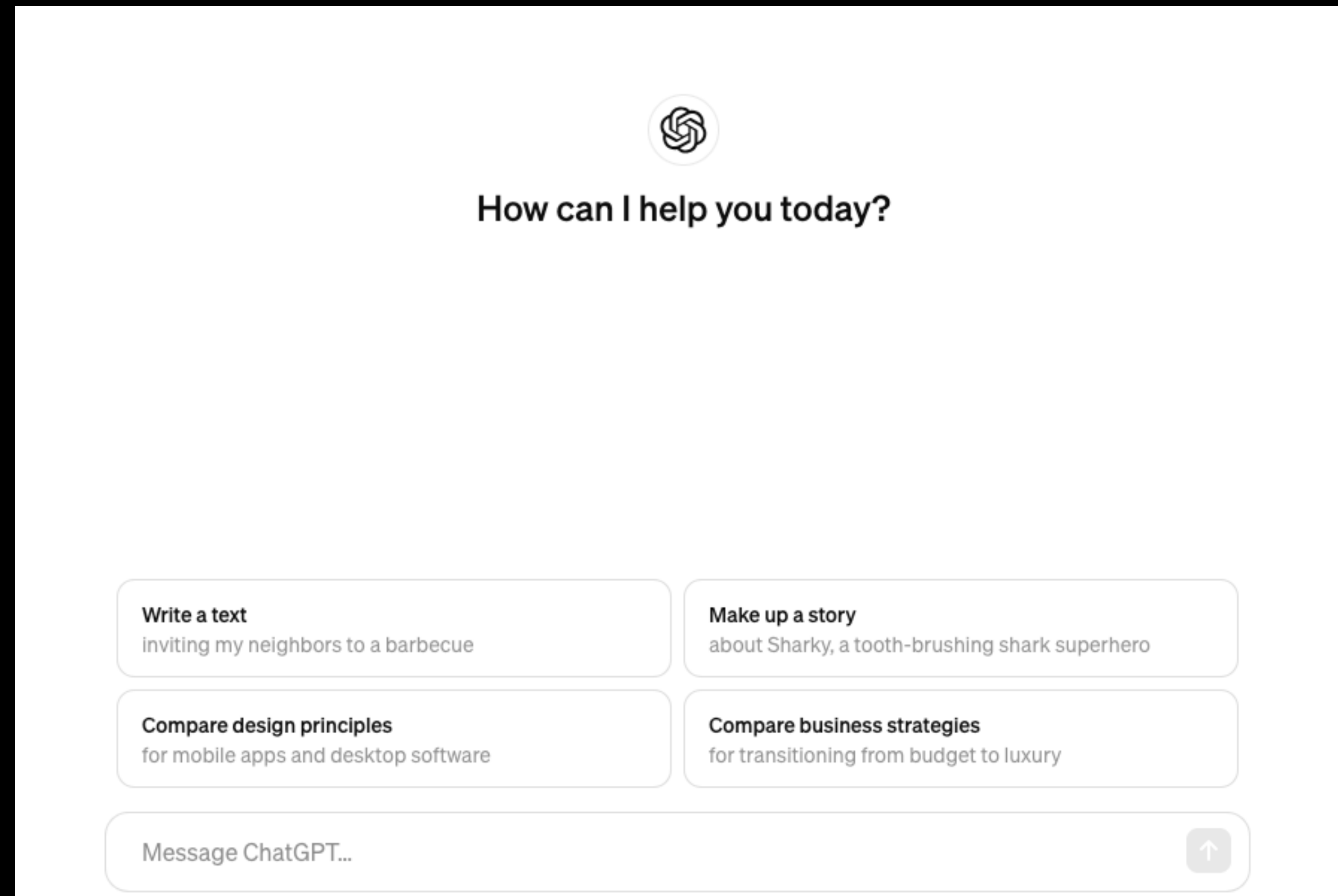


How many of you have used ChatGPT?

Early Days Large Language Models as text generators

What could they do?

- Great at writing stories
- Summarization tasks



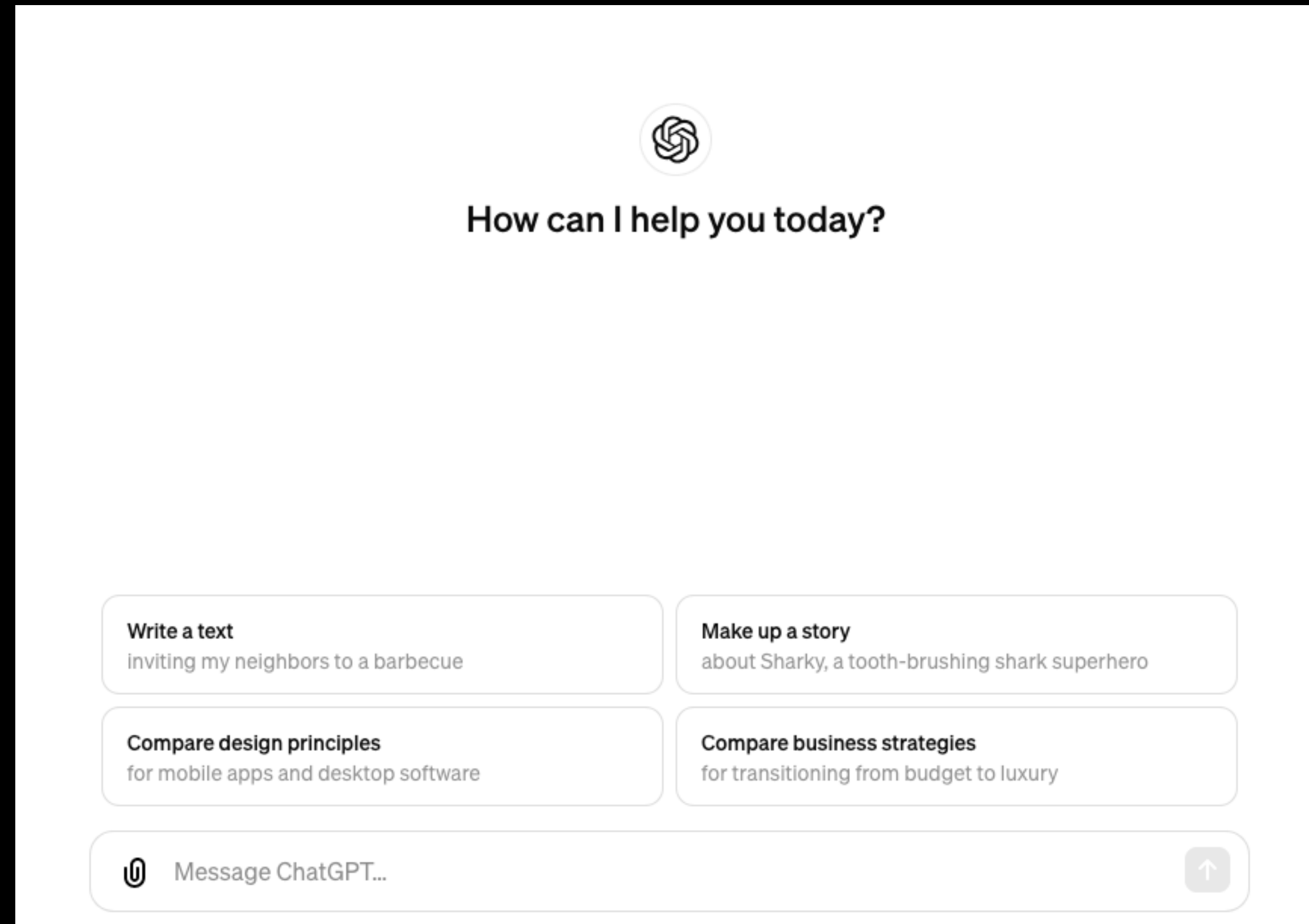
ChatGPT 3.5

Post Early Days Large Language Models with inputs besides text

What could they do?

- Great at writing stories
- Summarization tasks
- Ingest PDFs, images, and text
- Surf the web

But still lacked something!



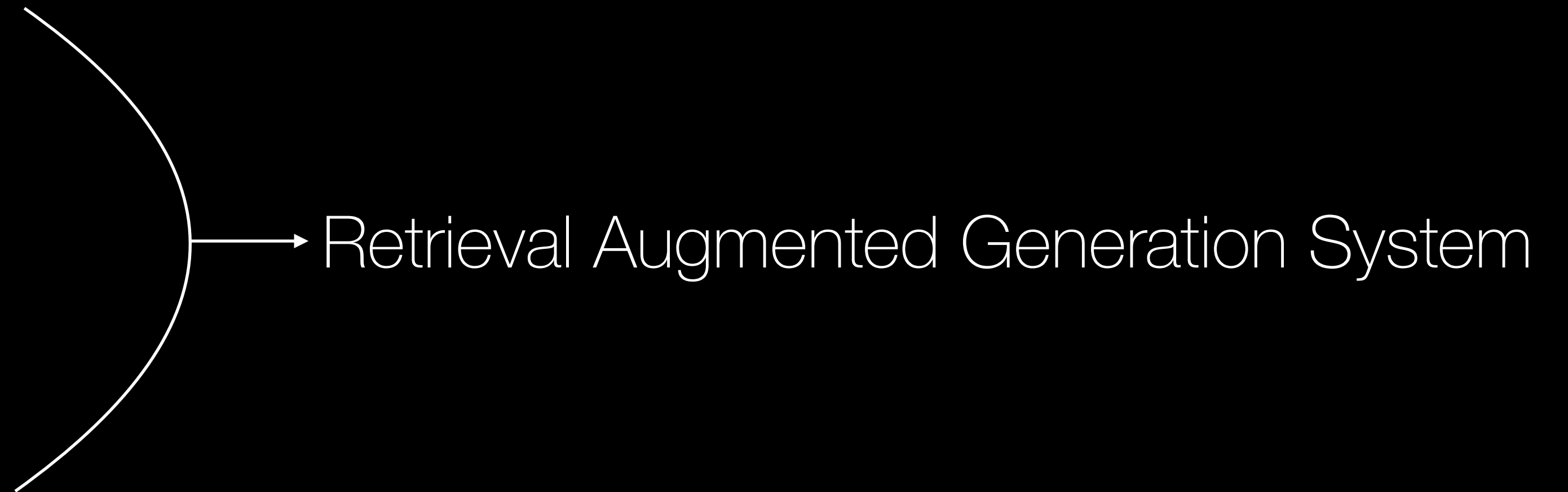
ChatGPT 4

Now

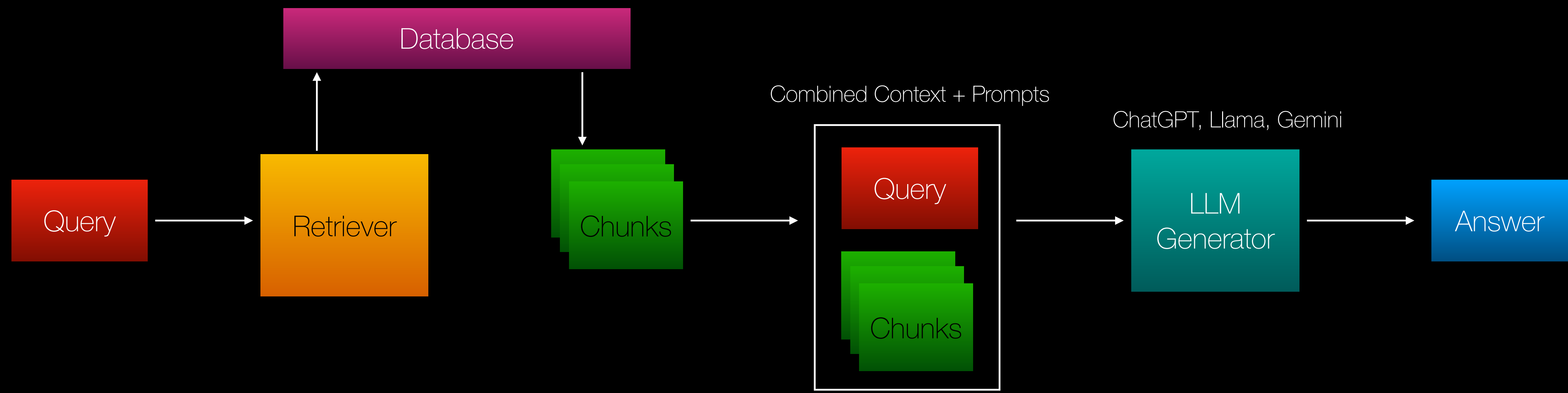
Large Language Models with memory

What could they do?

- Great at writing stories
- Summarization tasks
- Ingest PDFs, images, and text
- Surf the web
- Query a database

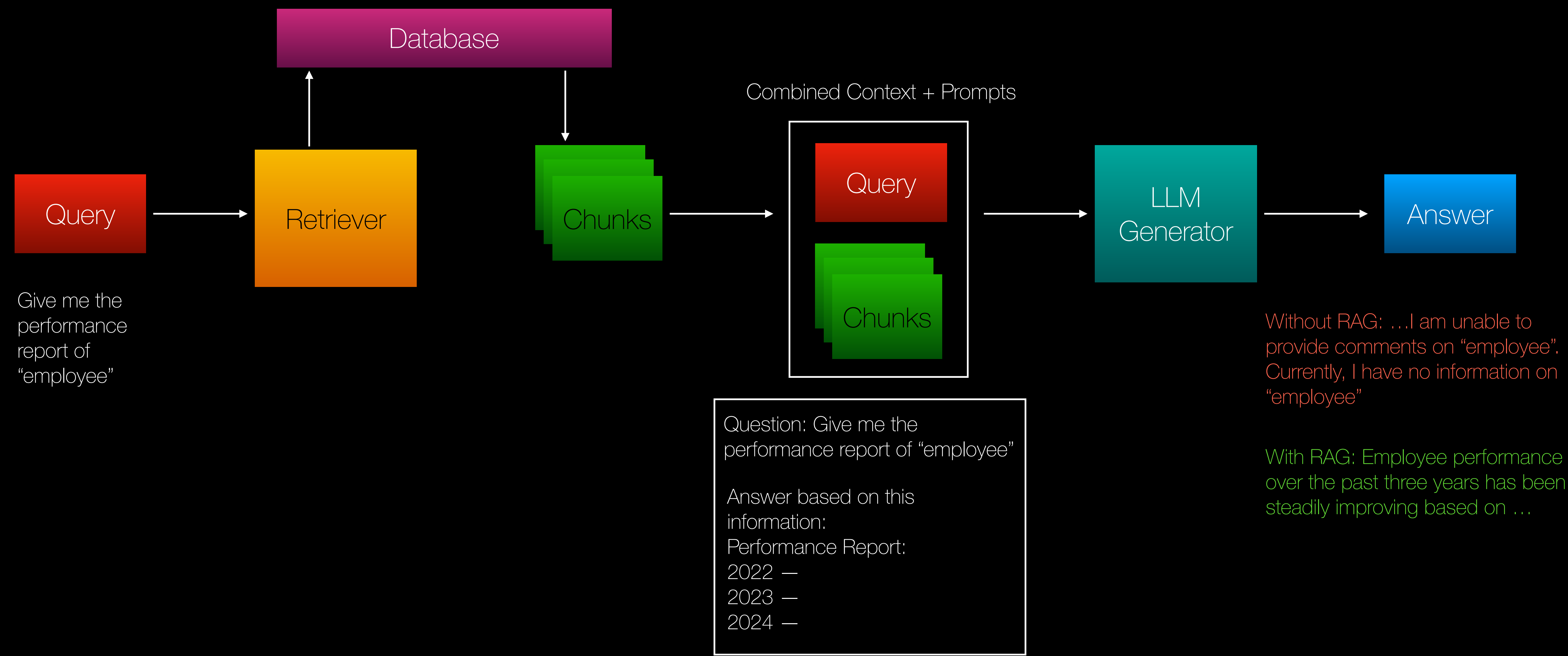


Retrieval Augmented Generation System



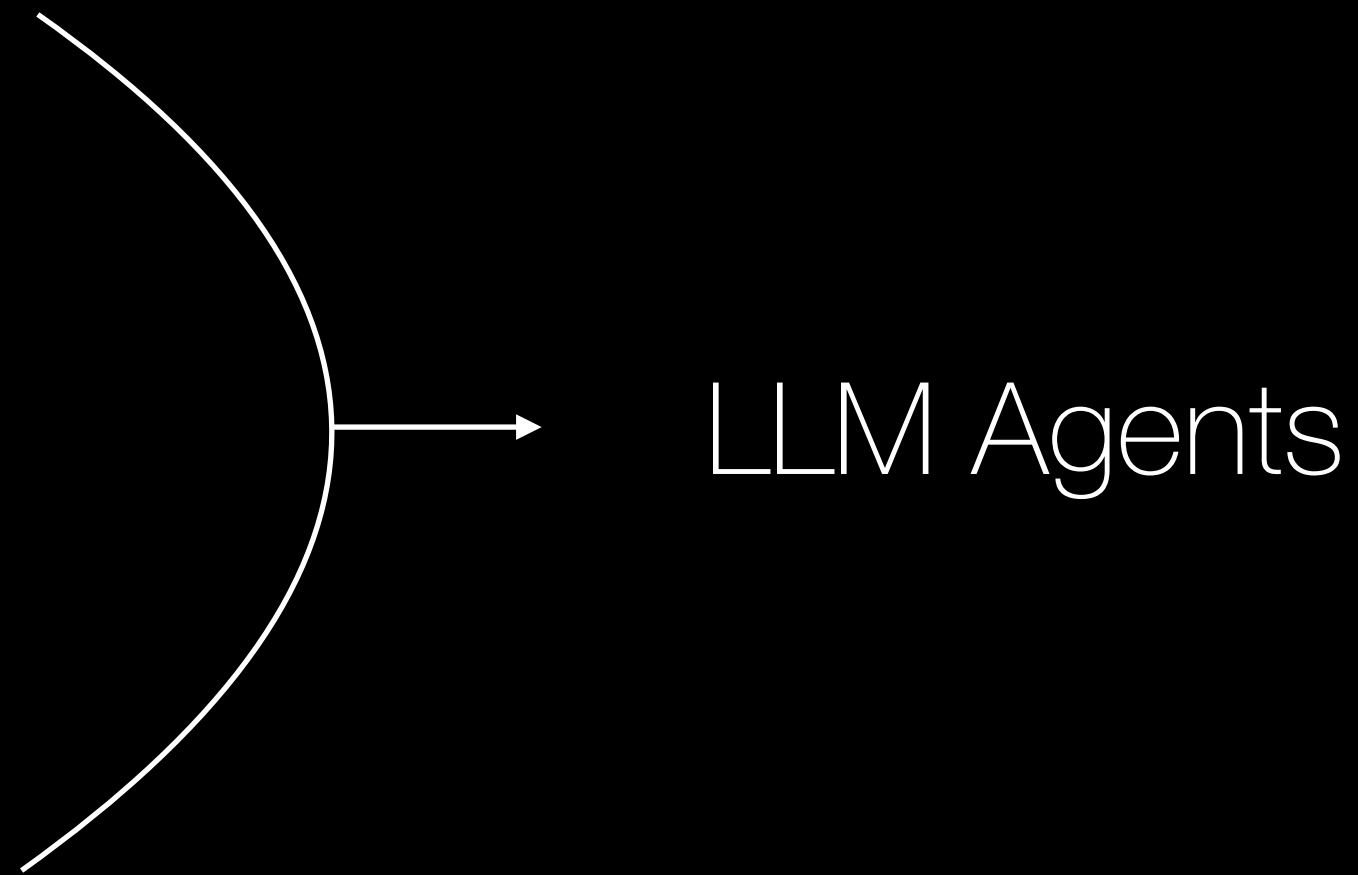
Small Language Models are better!

Retrieval Augmented Generation System: Example Walkthrough

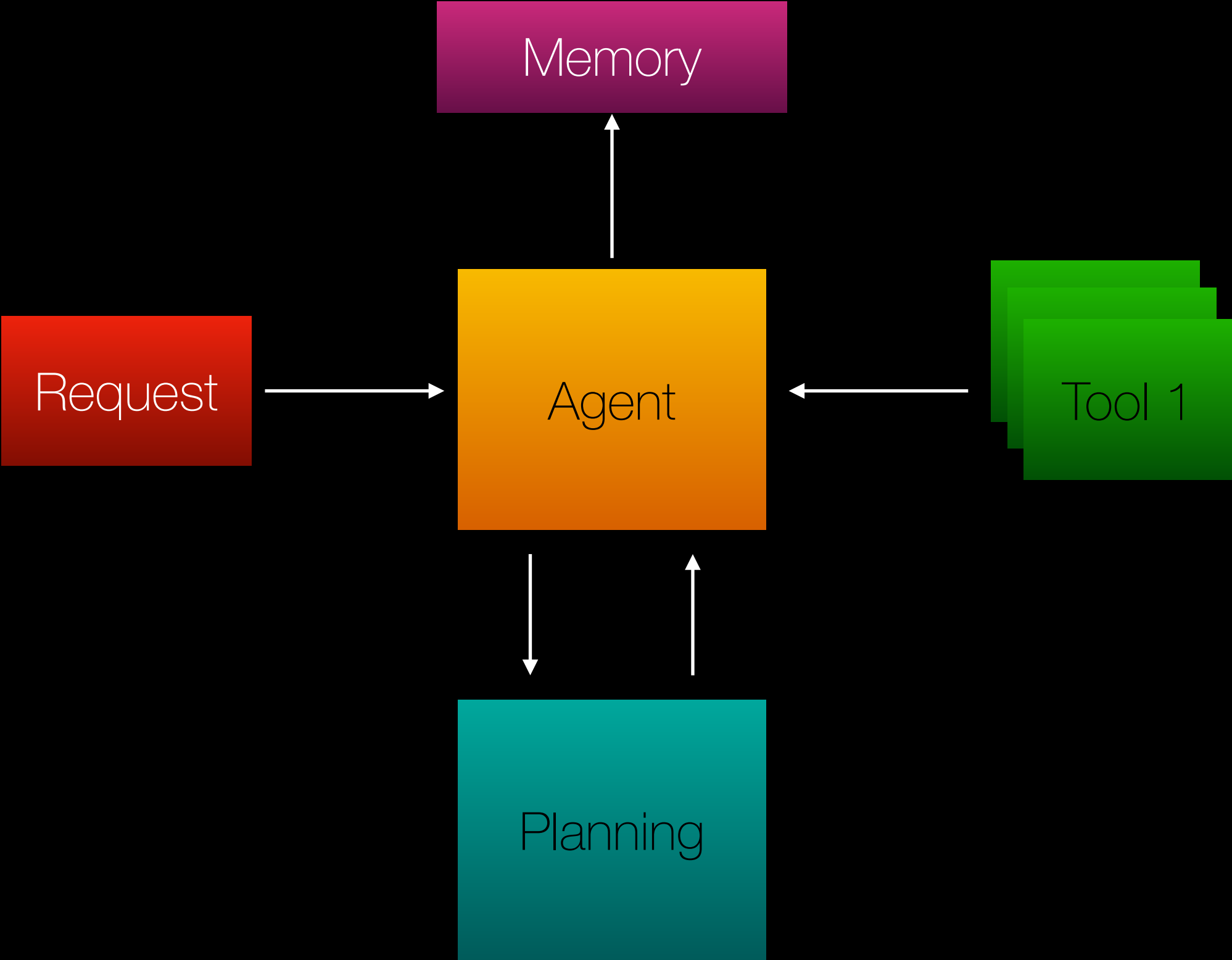


What could they do?

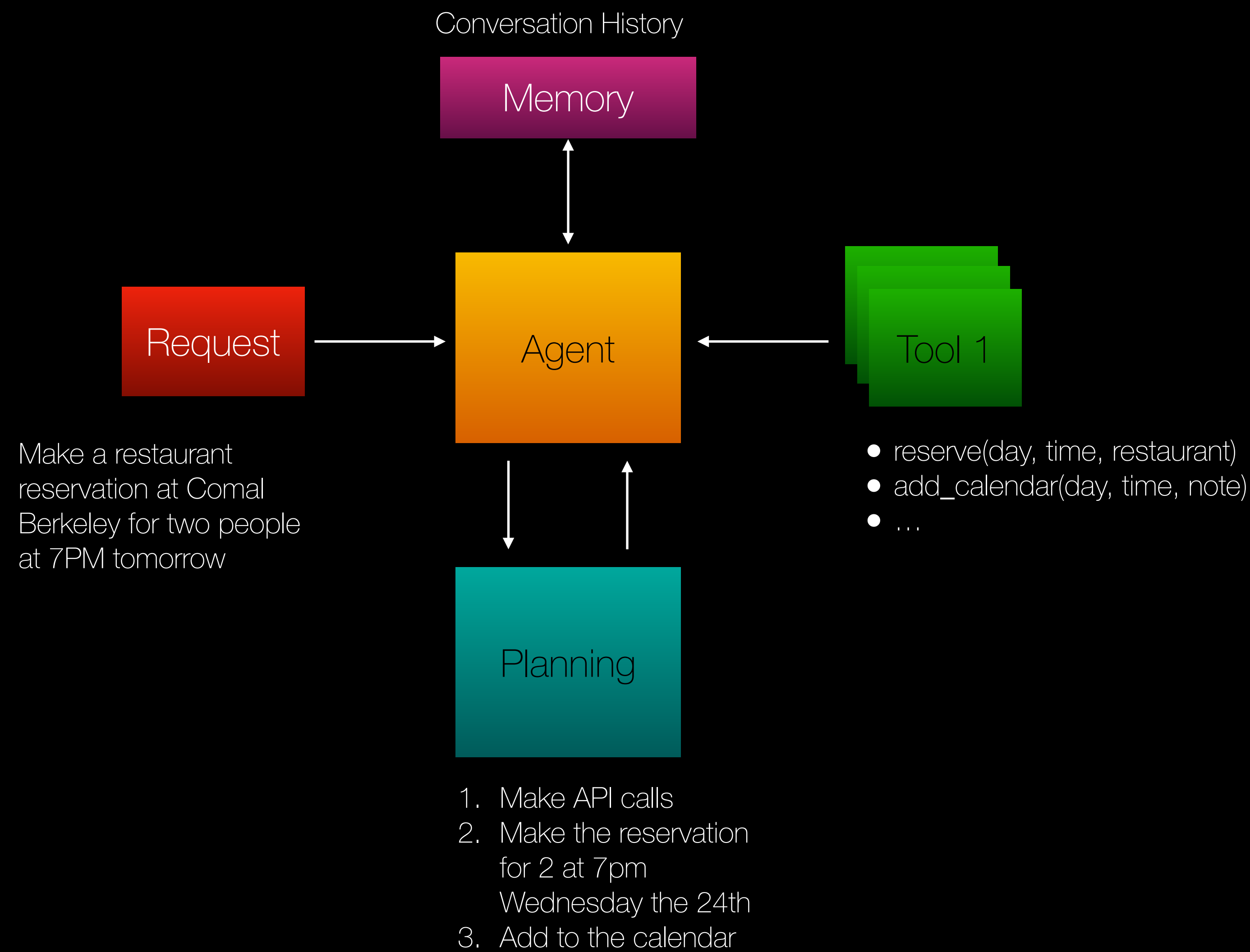
- Great at writing stories
- Summarization tasks
- Ingest PDFs, images, and text
- Surf the web
- Query a database
- Plan job
- Execute a job



LLM Agent



LLM Agent: Example Walkthrough



With agents, the possibilities are limitless

- Multi-agents: agents that communicate with one another
- Swarm of agents that work on a common goal
- Personal agent like a personal assistant: managing calendar and meetings

Some overarching benefits

- Improved productivity — (no humans actually doing the work)
- Reduced costs — (no humans actually doing the work)
- Informed decision making
- Improved customer experience

Communicative Agents for Software Development

Chen Qian♠ Xin Cong♠ Wei Liu♠ Cheng Yang♣ Weize Chen♠ Yusheng Su♠
Yufan Dang♠ Jiahao Li♦ Juyuan Xu^ Dahai Li★ Zhiyuan Liu♠✉ Maosong Sun♠✉
♠Tsinghua University ♣Beijing University of Posts and Telecommunications
♦Dalian University of Technology ^Brown University ★Modelbest Inc.
qianc62@gmail.com liuzy@tsinghua.edu.cn sms@tsinghua.edu.cn

Thank you