

LLMs and Agents

Brainstorming



July 11 2025

Where should we begin?

Ask anything

+



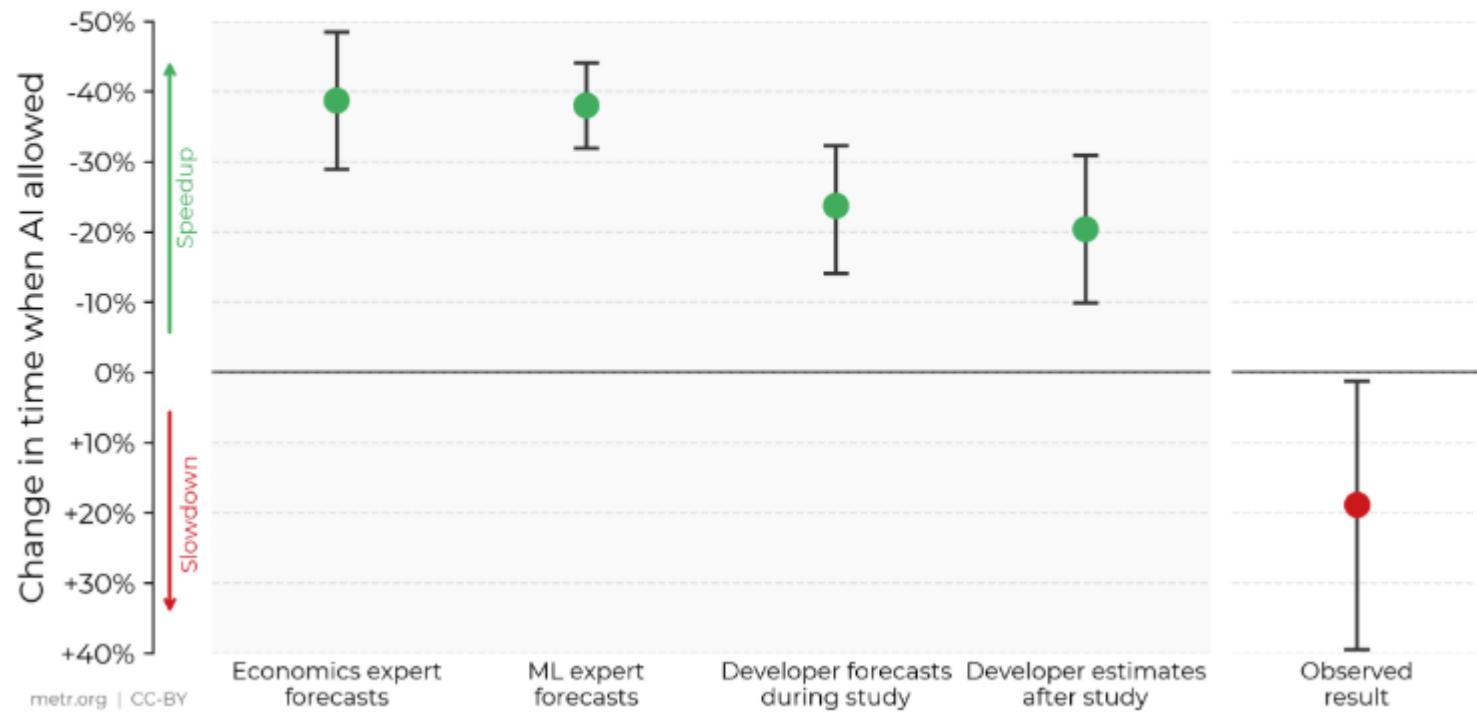
Tools



Against Expert Forecasts and Developer Self-Reports, Early-2025 AI Slows Down Experienced Open-Source Developers



In this RCT, 16 developers with moderate AI experience complete 246 tasks in large and complex projects on which they have an average of 5 years of prior experience.





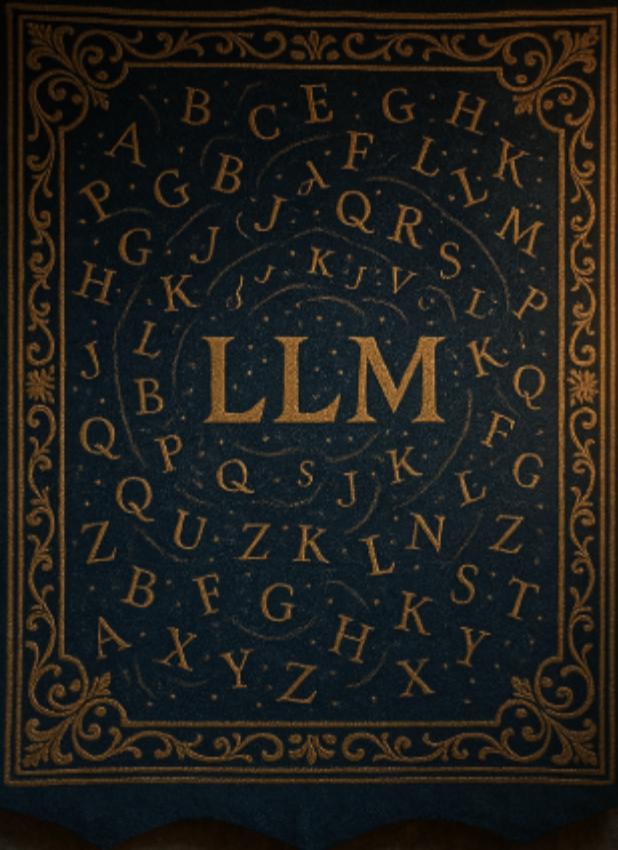
WORD WIZARDS

B

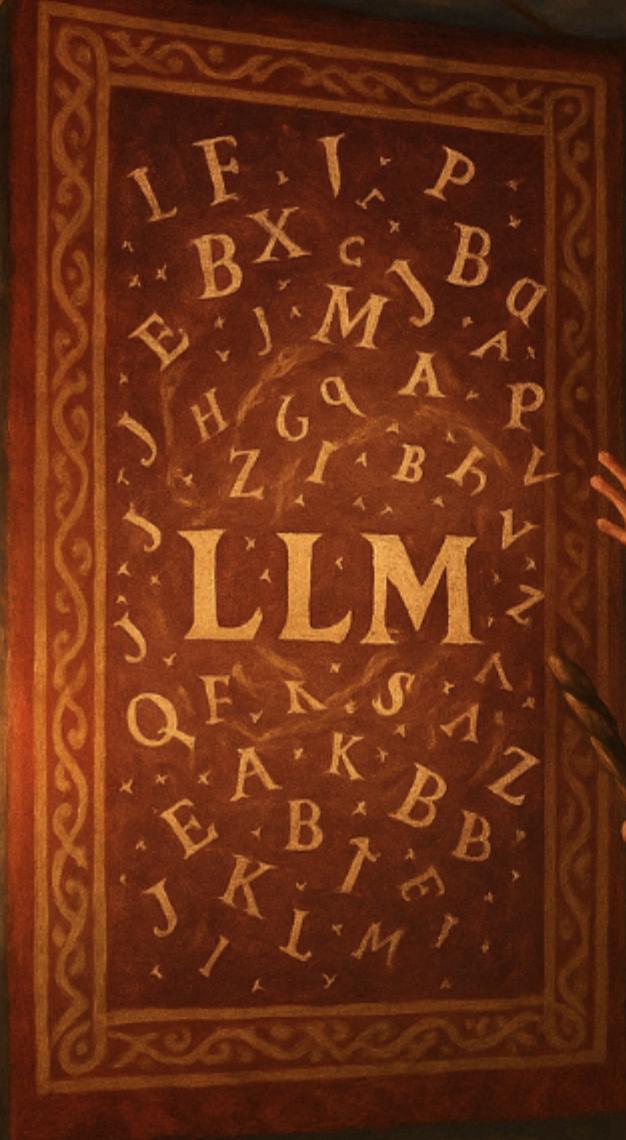
Artificial Intelligence \neq LLM

Artificial Intelligence \neq Agentic

Let's focus on definitions...

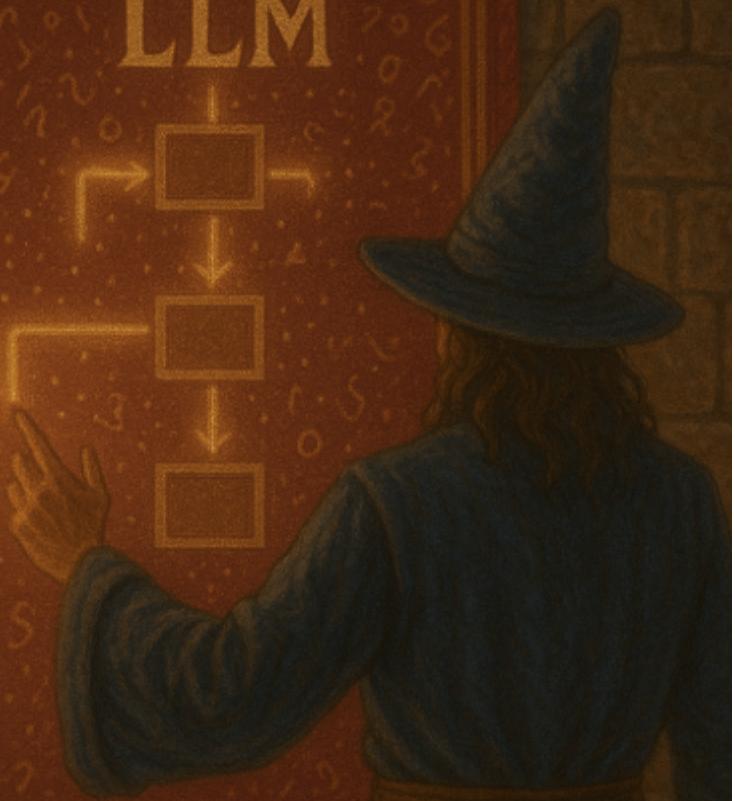
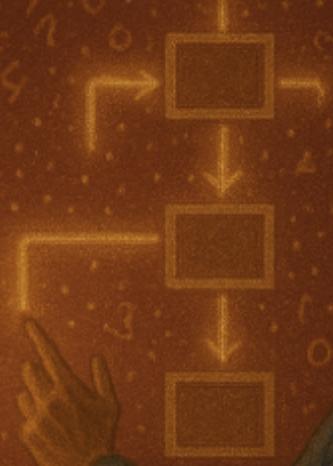


B C E G H
A G B F L K
R G J Q R S M
G J K J V S L P
H K J V L P Q
J L B K Q F G
Q P Q S J K L N Z
Q U Z K L N S T
Z B F G H K Y
A X Y Z X Y



L F J P
B X C B Q
E Y M A P
H G Q A L
Z I B S V
LLM
Q F R S A
E A K Z
J B T B
K I G A
L M I

LLM



Prompt

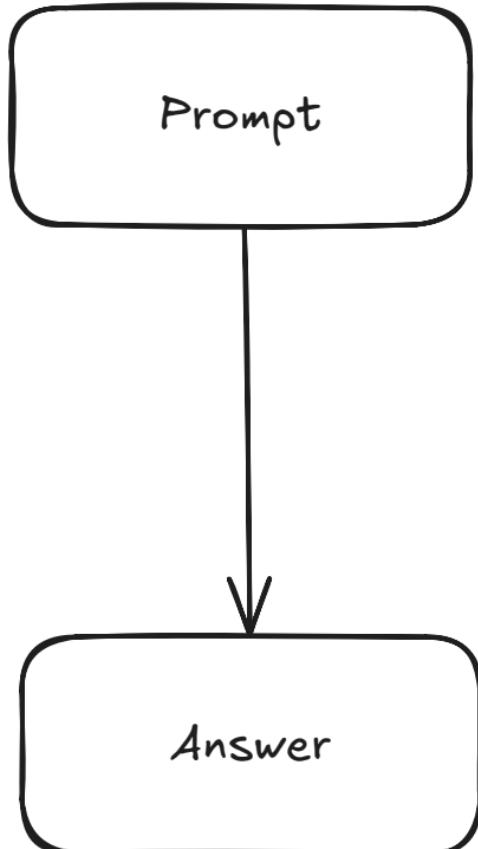


LLM

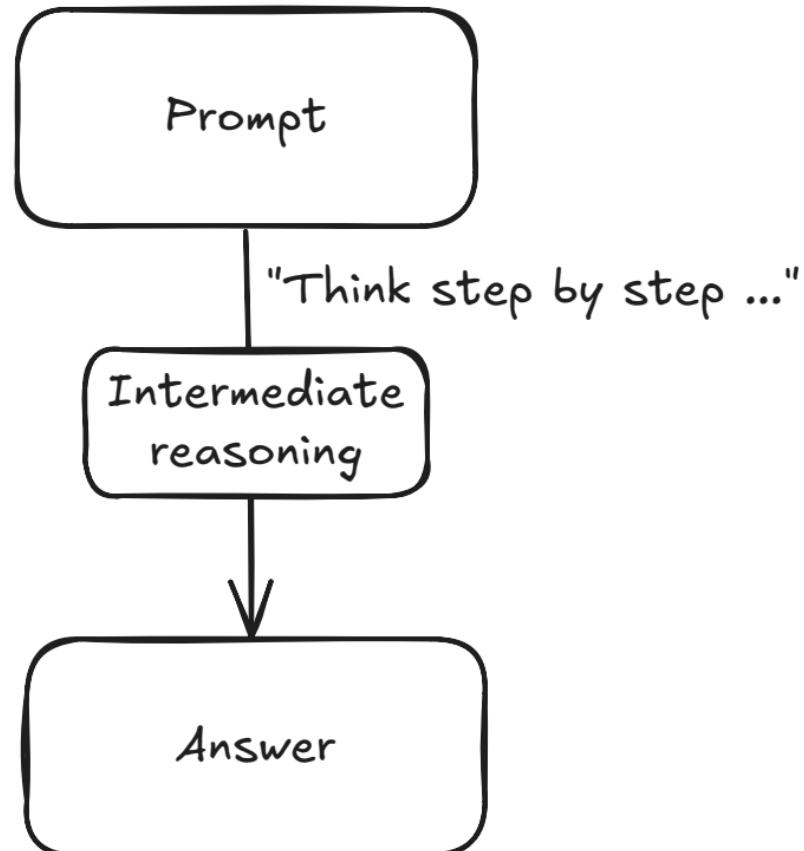


Answer

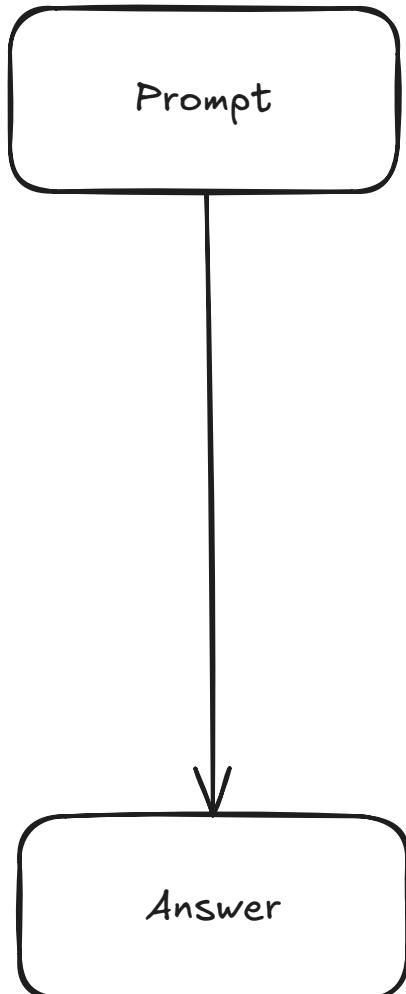
"direct prompting"



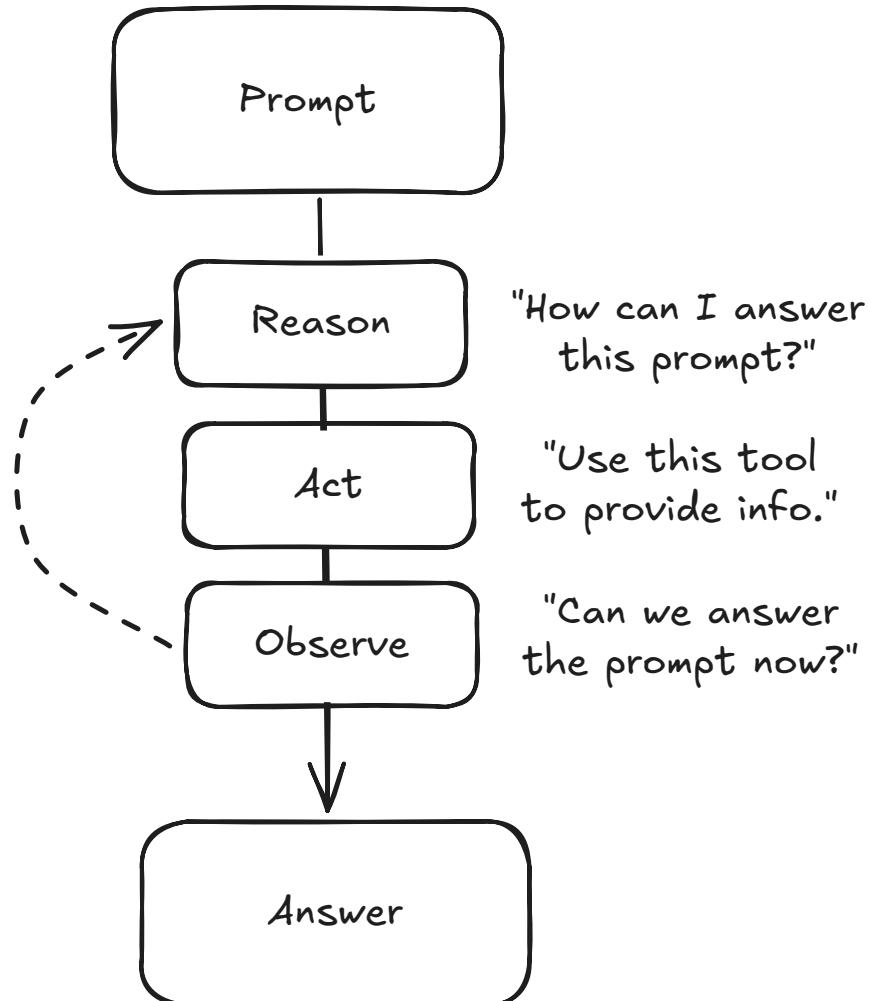
"chain of thought"

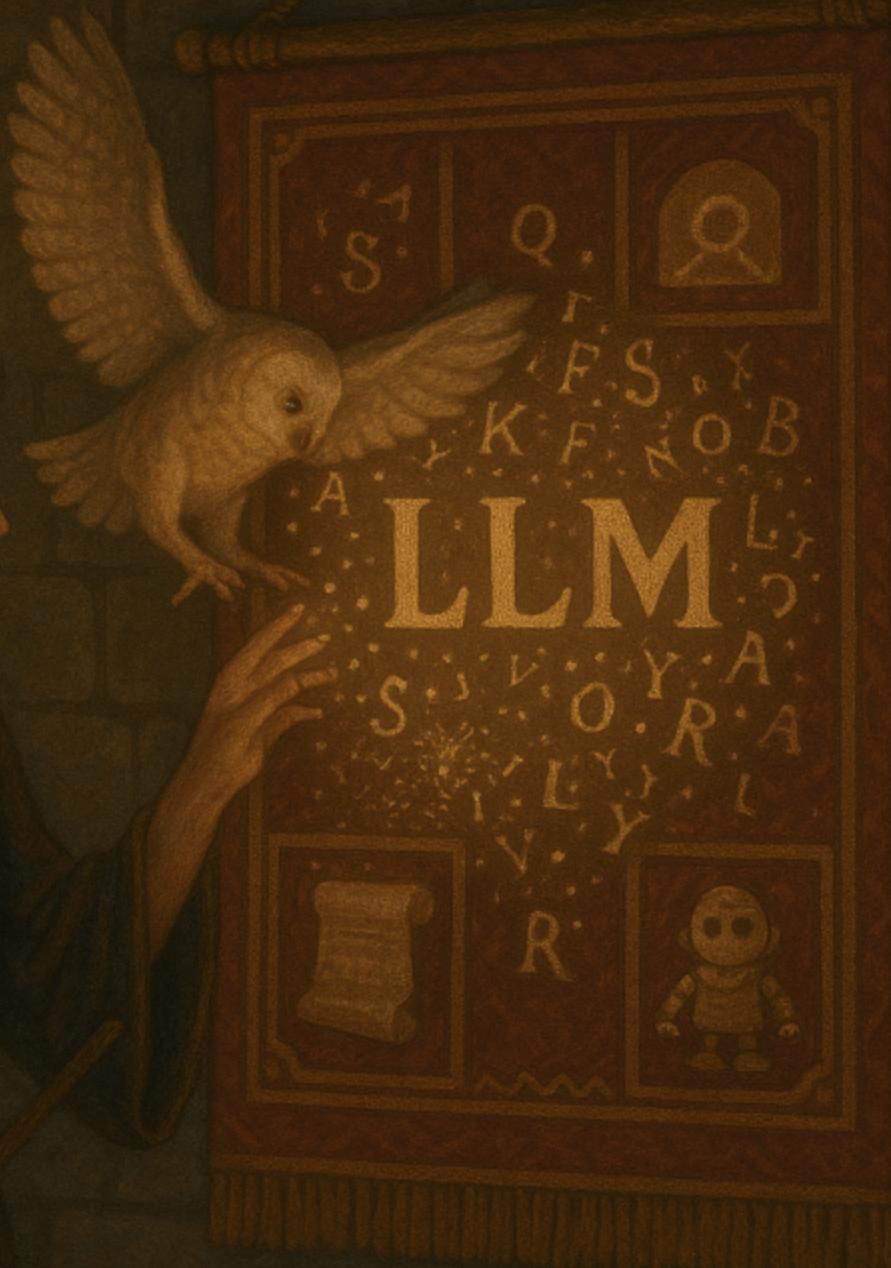


"direct prompting"

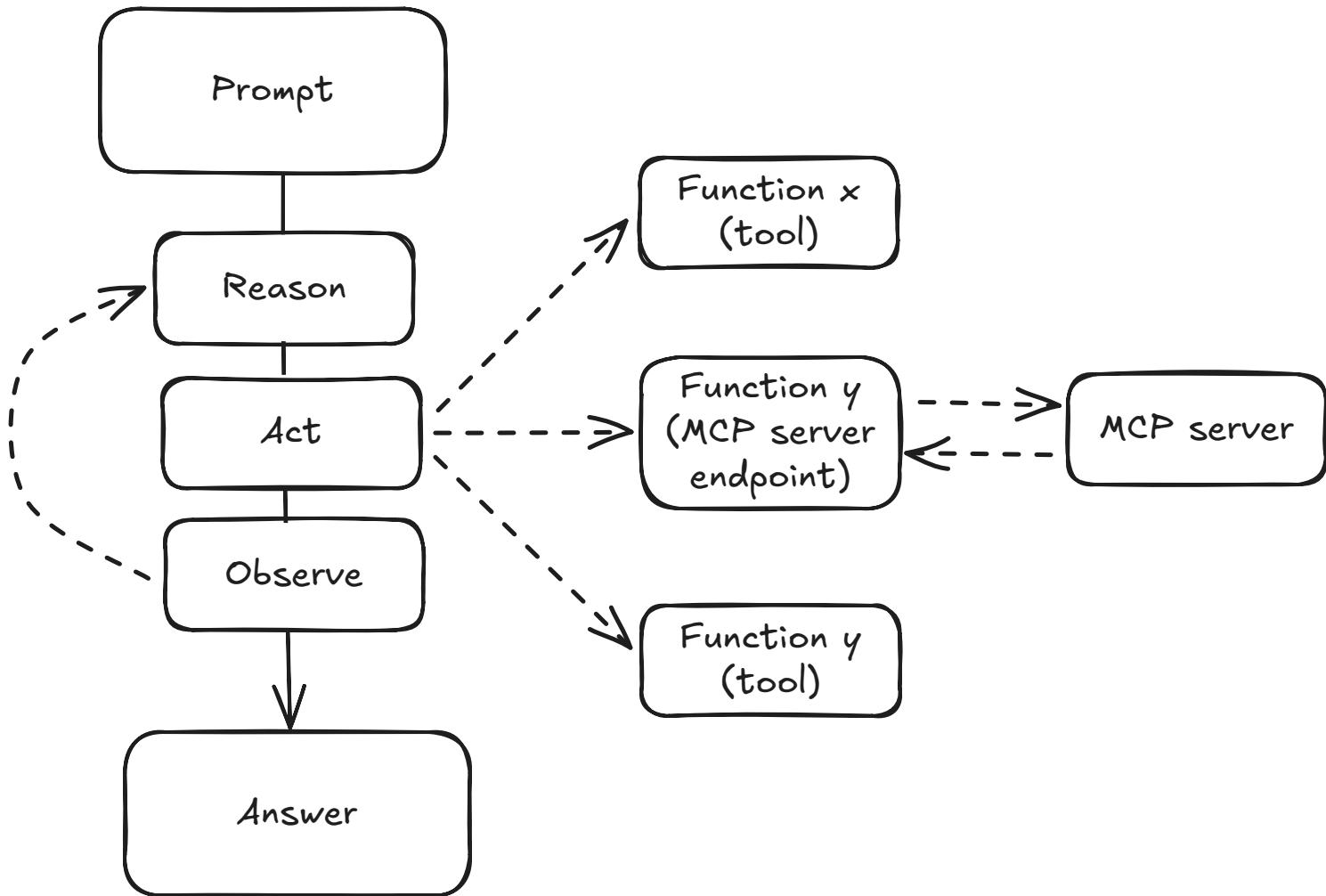


"ReAct"



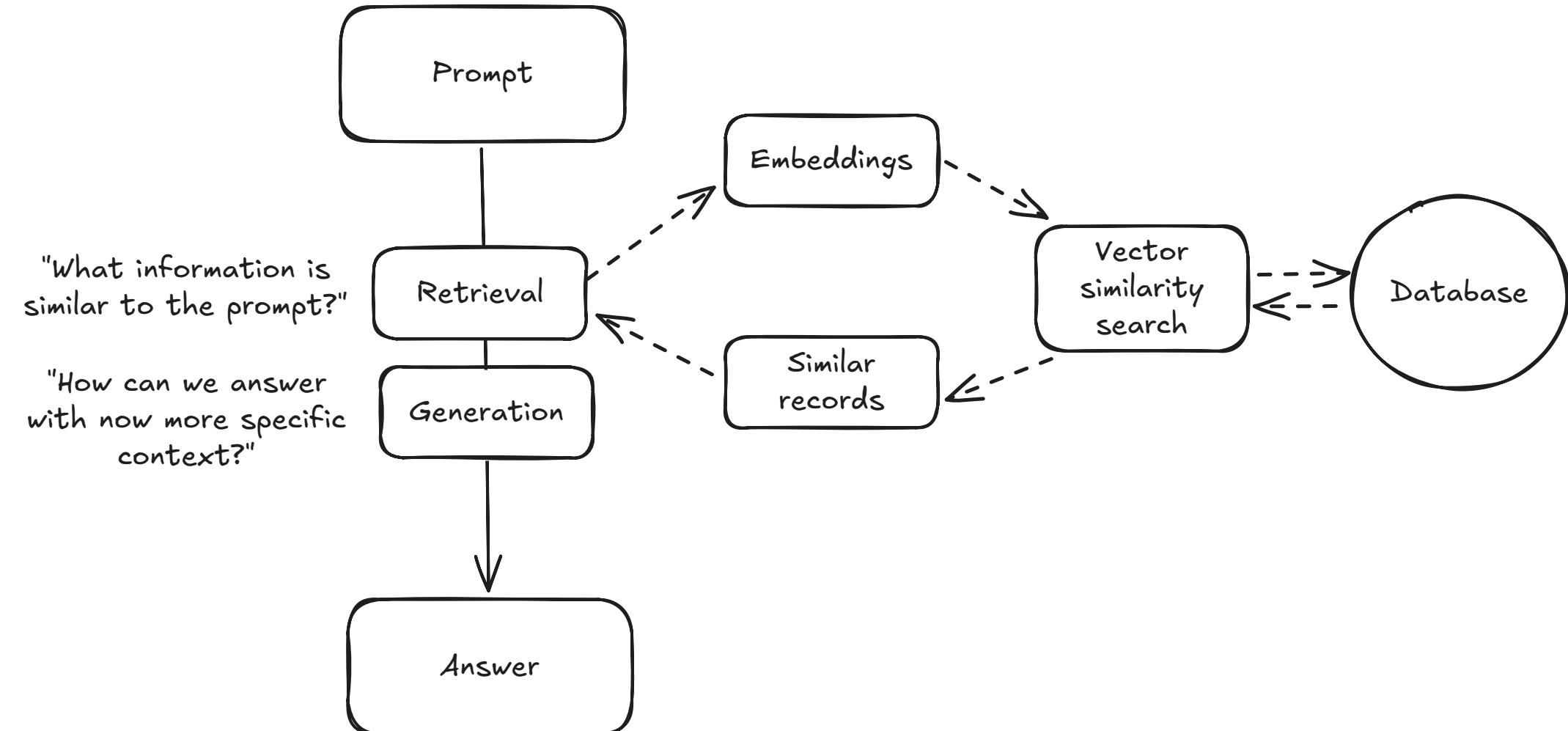


"ReAct"

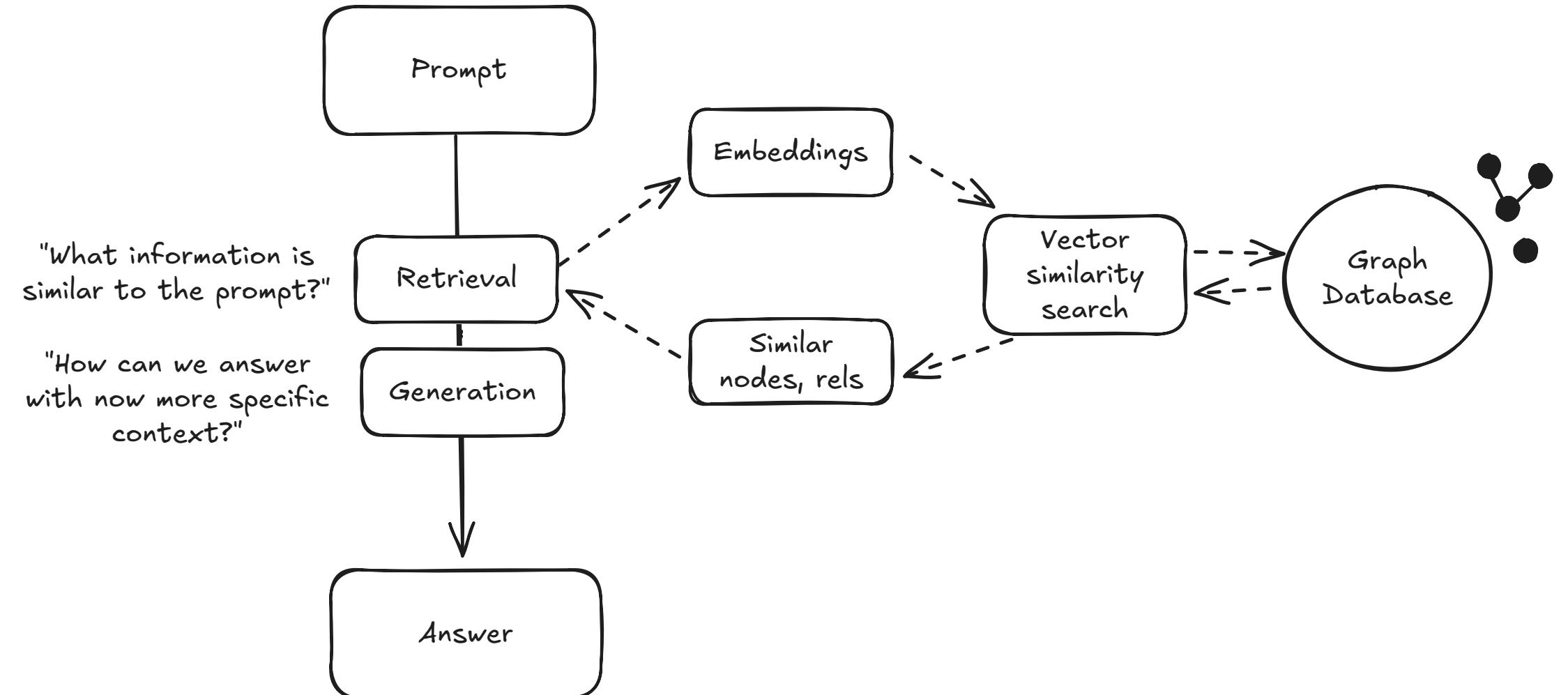




"RAG"



"RAG"



EXTERNAL LUM



C B N I K A
S P Q Z

LOCAL LLM

P R C C L O



A Q

L U V S A N C U S

L U V S A N C U S

L U V S A N C U S



Prompt



Answer

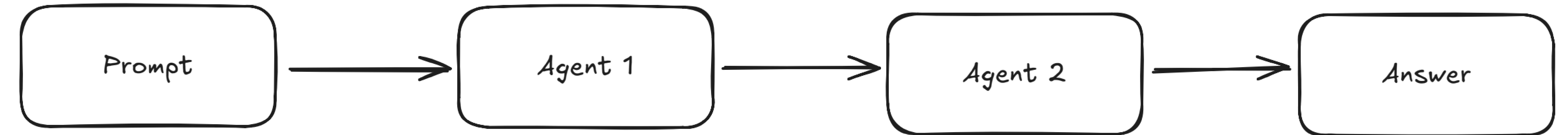
Prompt

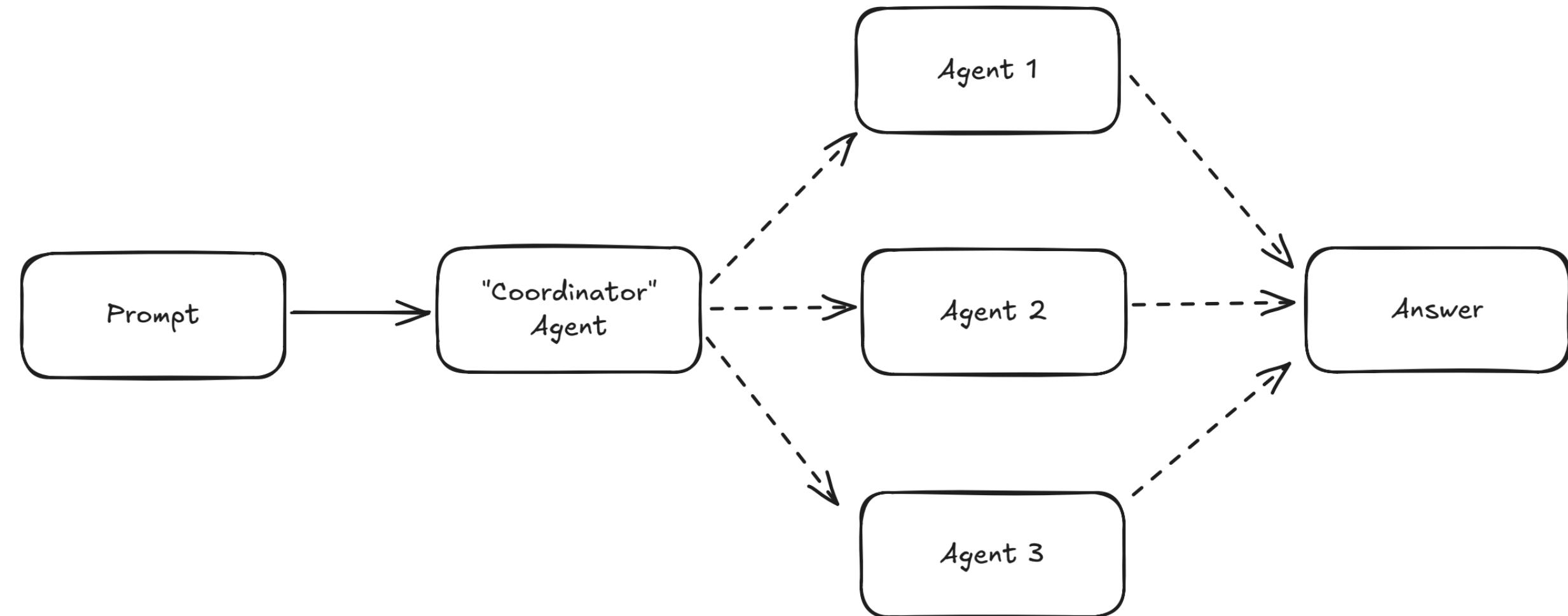
"castle"



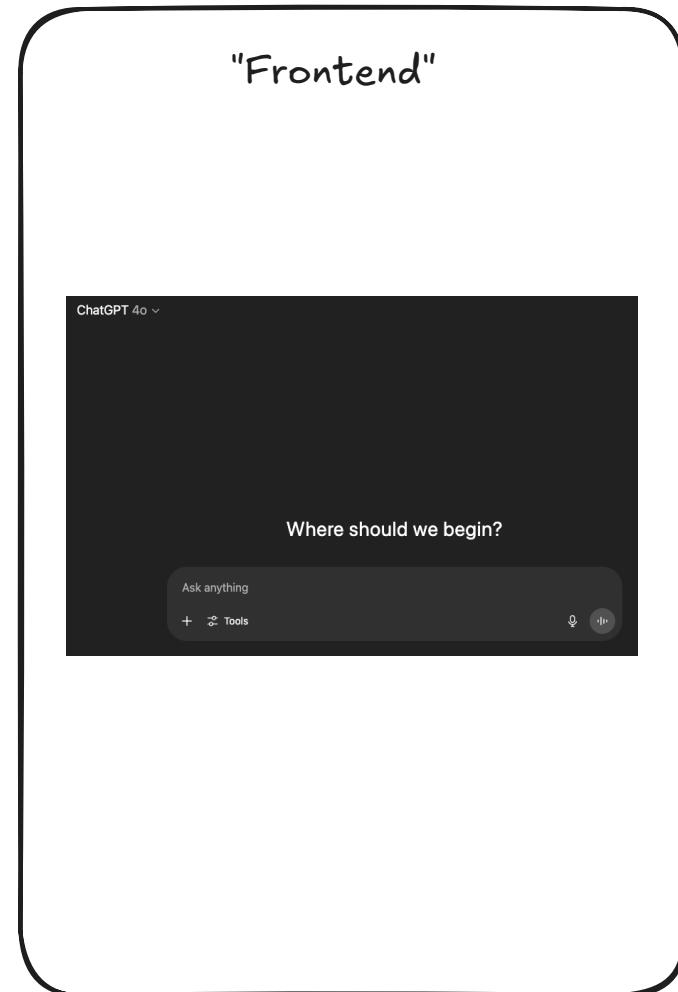
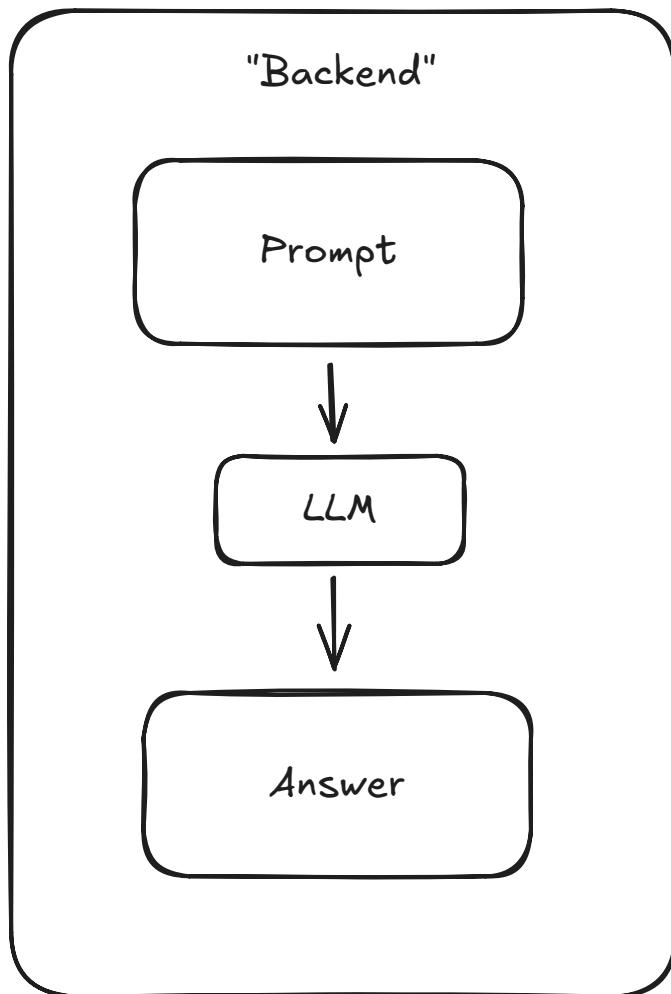
Answer





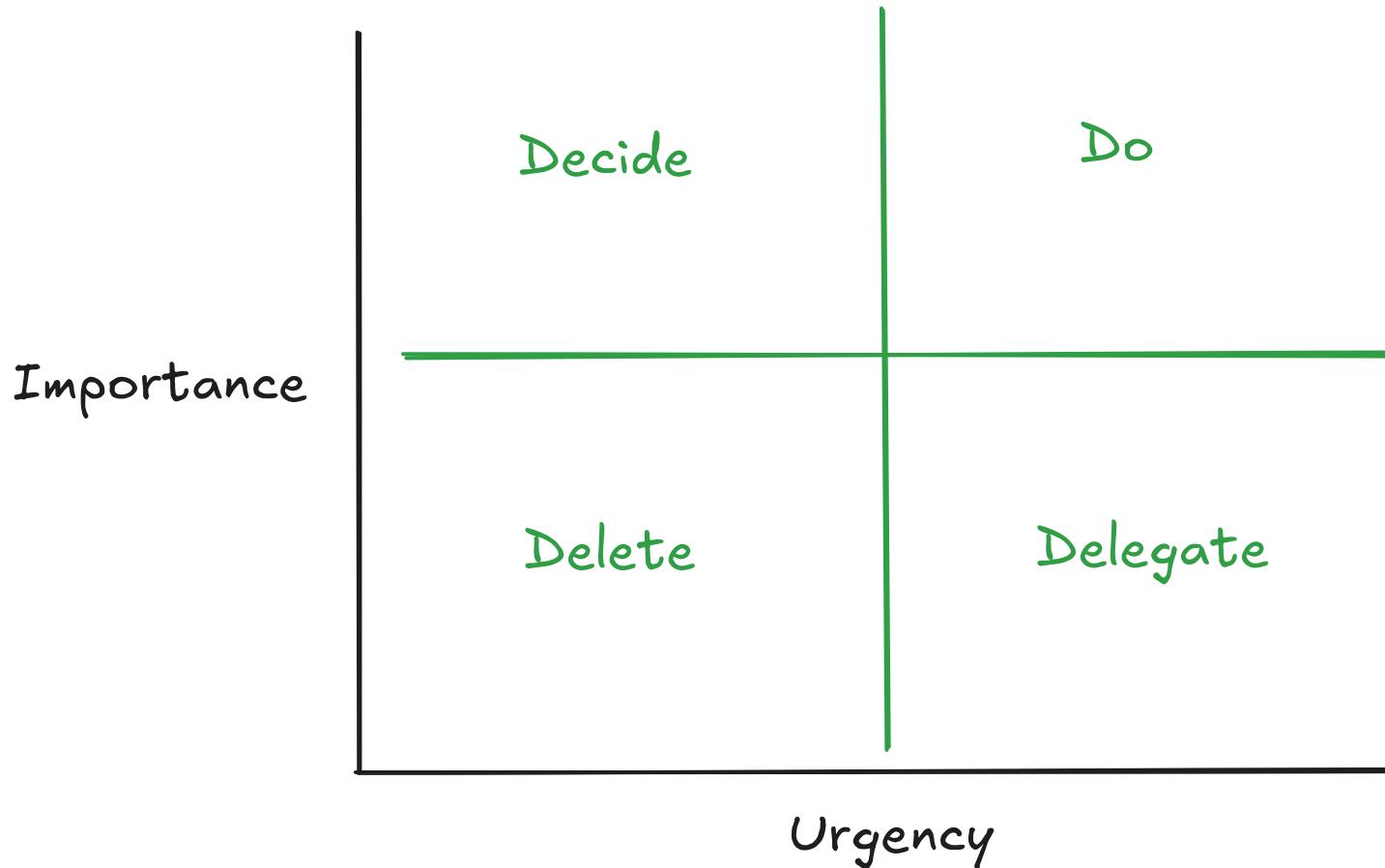


Web interface packages (Streamlit, Gradio, etc.)

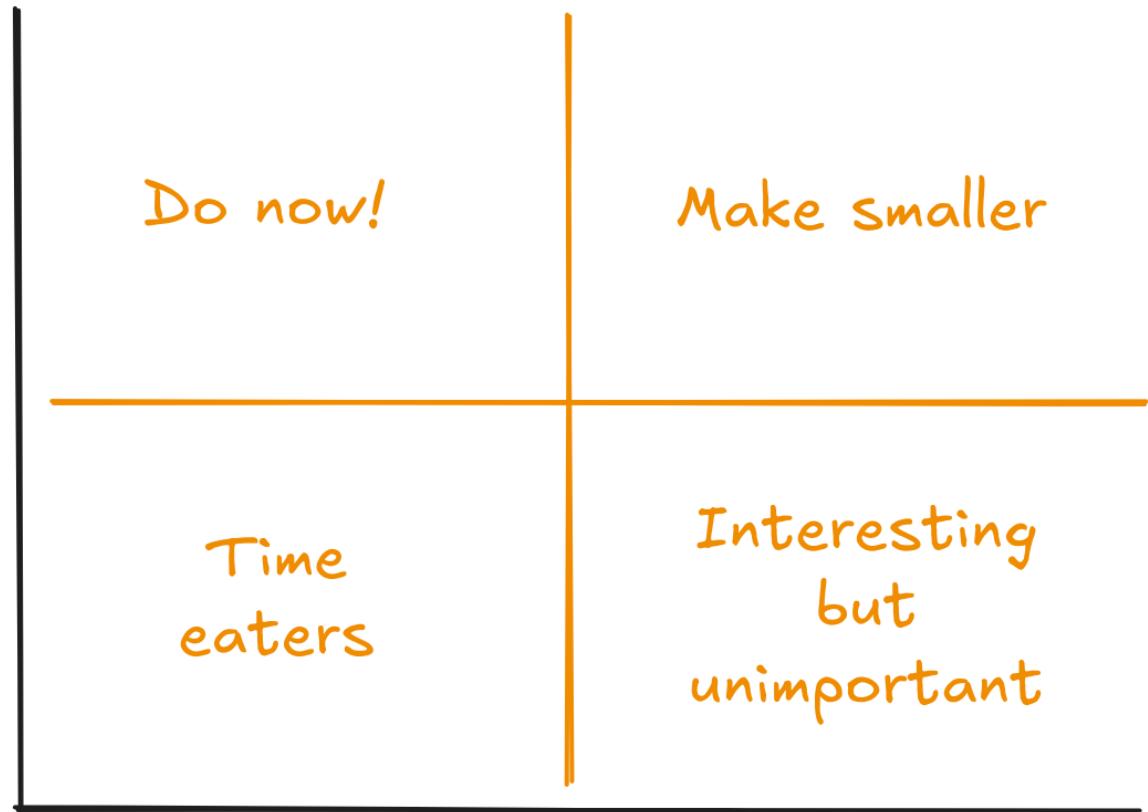


Demos!

Eisenhower matrix



Value



Complexity

Brainstorm!

What could agents do for your science?