

# **Unleashing Efficiency**

**Leveraging LLM Best Practice Principles and Frameworks for  
Research Engineering Productivity**

**Prof. Nathan Kutz and Mel Mashiku, June 27 2023**

# Agenda

- Prompt Engineering 101
- Live Coding: ChatGPT API 101
- Live Coding: LangChain 101
- Live Coding: GPT-Engineer 101

# Types of LLMs

- Base LLM
  - Predicts next word, based on text training data.
- Instruction Tuned LLM
  - Follows instructions.
  - Fine-tune on instructions and good attempts at following those instructions.
    - RLHF: Reinforcement Learning with Human Feedback

# Principles of Prompting

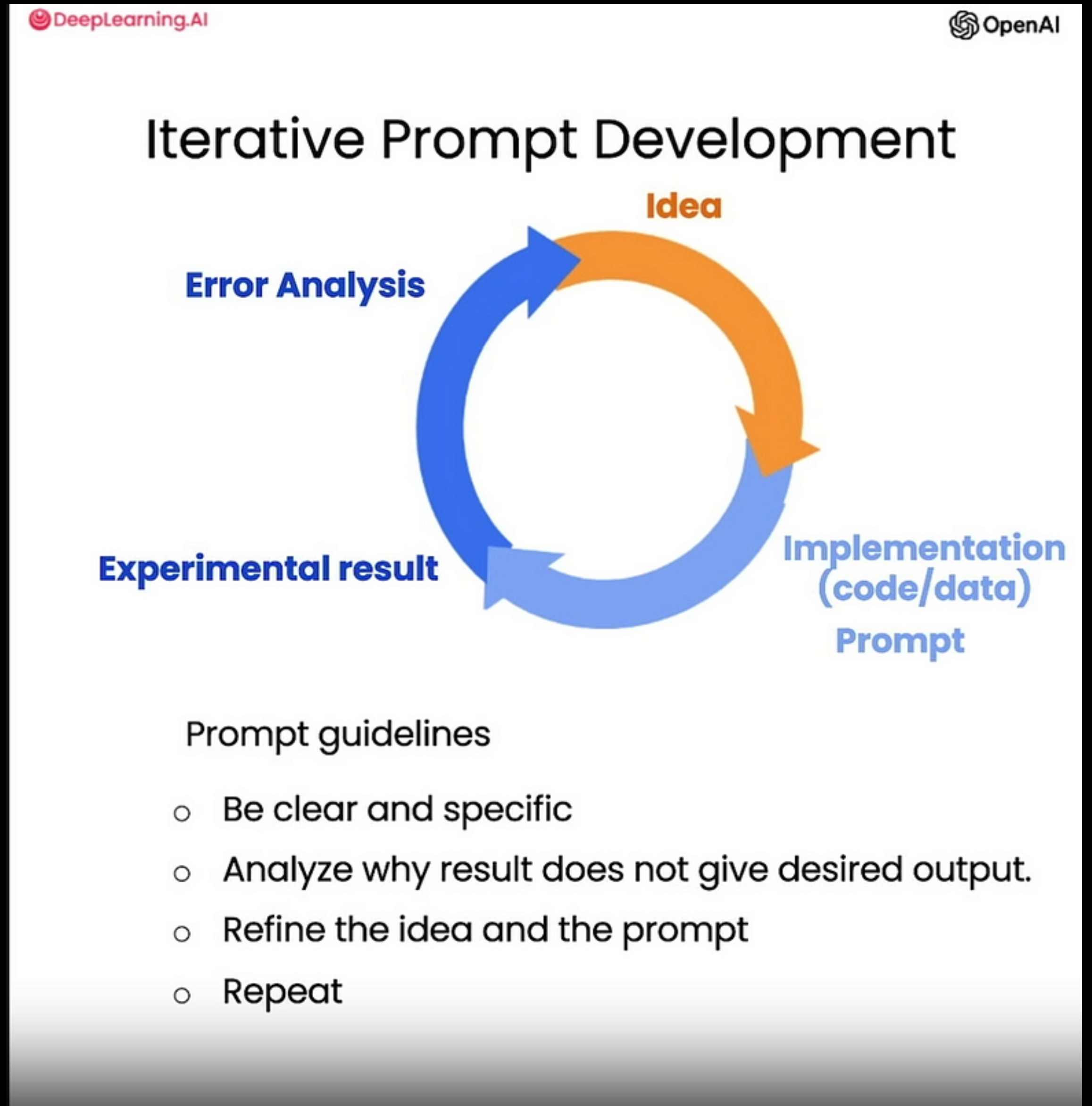
- Principle 1: Write clear and specific instructions
  - Use delimiters: Angle brackets: `<>`, XML tags: `<tag> </tag>`
  - Ask for structured output: HTML, JSON
  - Few-shot prompting: Give successful examples of completing tasks. Then ask model to perform the task
- Principle 2: Give the model time to think
  - Specify the steps to complete a task
  - Instruct the model to work out its own solution before rushing to a conclusion

# Model Limitations

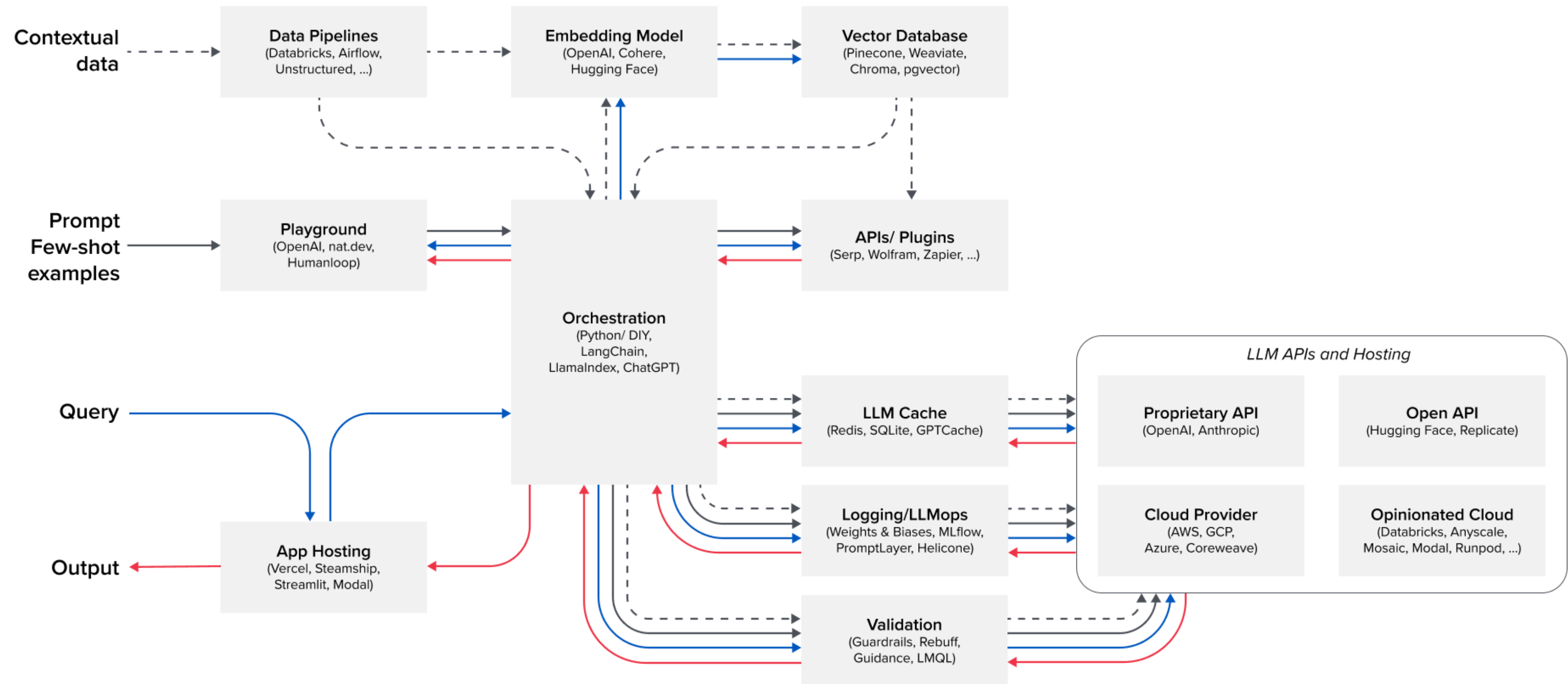
- Hallucination: Makes statements that sound plausible but are not true
- Reducing hallucinations: Instruct the model to first find relevant information, then answer the question based on the relevant information. Restrict source and request citation.

# Iterative Prompt Development

- Try something
- Analyze where the result does not give what you want
- Clarify instructions, give more time to think
- Refine prompts with a batch of examples



# Emerging LLM App Stack



## LEGEND

- Gray boxes show key components of the stack, with leading tools/systems listed
- Arrows show the flow of data through the stack
- - - -> Contextual data provided by app developers to condition LLM outputs
- > Prompts and few-shot examples that are sent to the LLM
- > Queries submitted by users
- > Output returned to users

# References

- [DeepLearning.AI Principles](#)
- [OpenAI ChatGPT API](#)
- [LangChain Framework](#)
- [GPT-Engineer](#)
- [A16z](#)



<http://bit.ly/MLW62023>