Prompt Engineering to Support AI Enabled Research

Claire Kelley

Child Trends

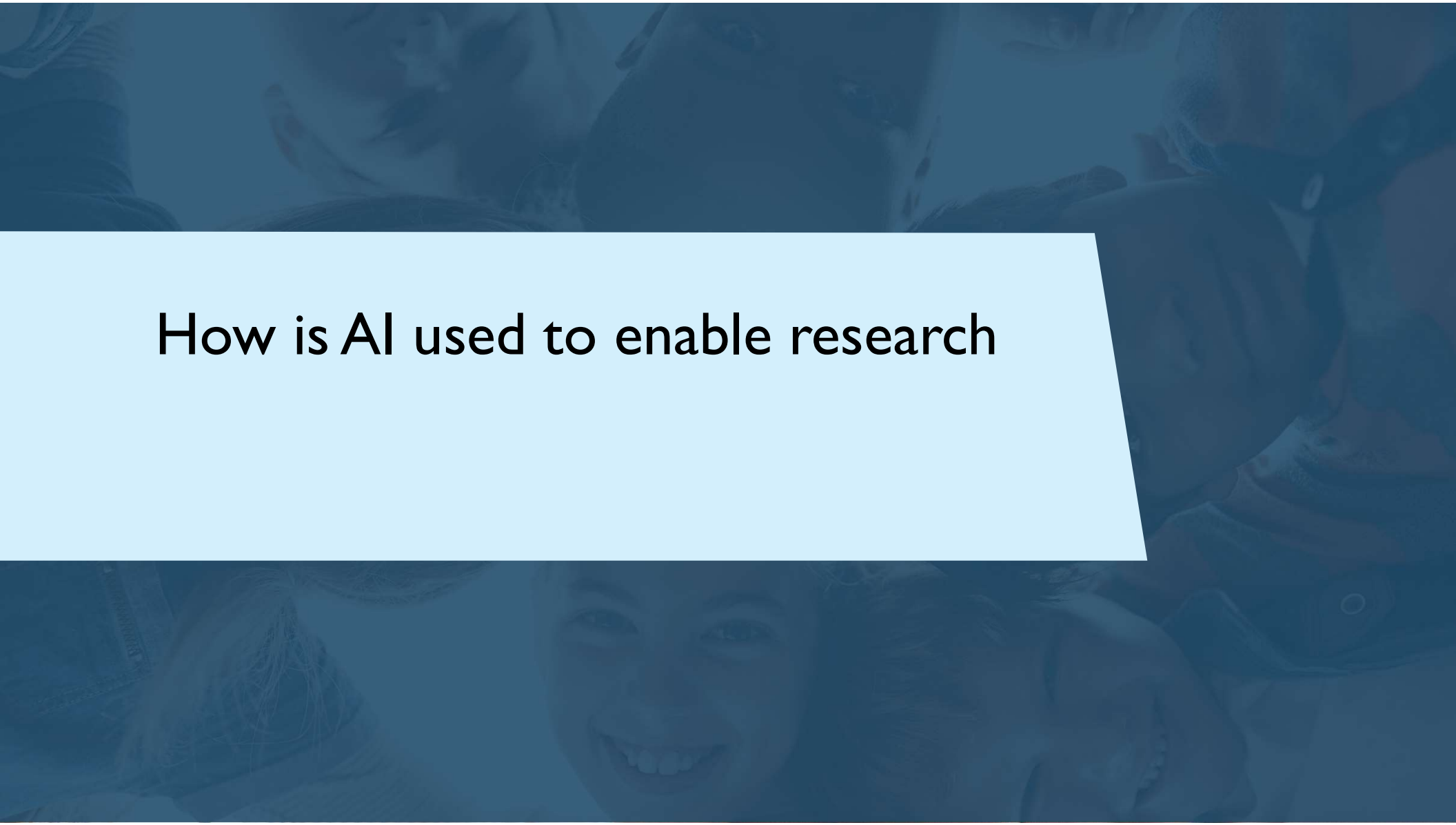
Introductions

Claire Kelley

Senior Data Scientist and Co-Director of Data Science

AGENDA

- How is AI used to enable research
- Introduction to Prompt Engineering
- Prompt Engineering Frameworks
- Testing and validating LLM results
- Case Study: Developing a research agenda for adolescent health with AI support



How is AI used to enable research

What do we mean by AI in research?

Development of AI	AI as a method	AI as a tool	AI as a research subject
• Development of new AI algorithms/methods • Validation of existing AI algorithms • E.g. developing a new neural network architecture to improve computer vision	• AI as a tool to conduct substantive research • E.g. Using machine learning to identify predictors of graduation	• AI is a non central part of the project, but is used to support research operations • E.g using AI to produce code to analyze data	• Substantive research on the impact of AI on areas of interest • e.g. how does AI use alter cognitive development

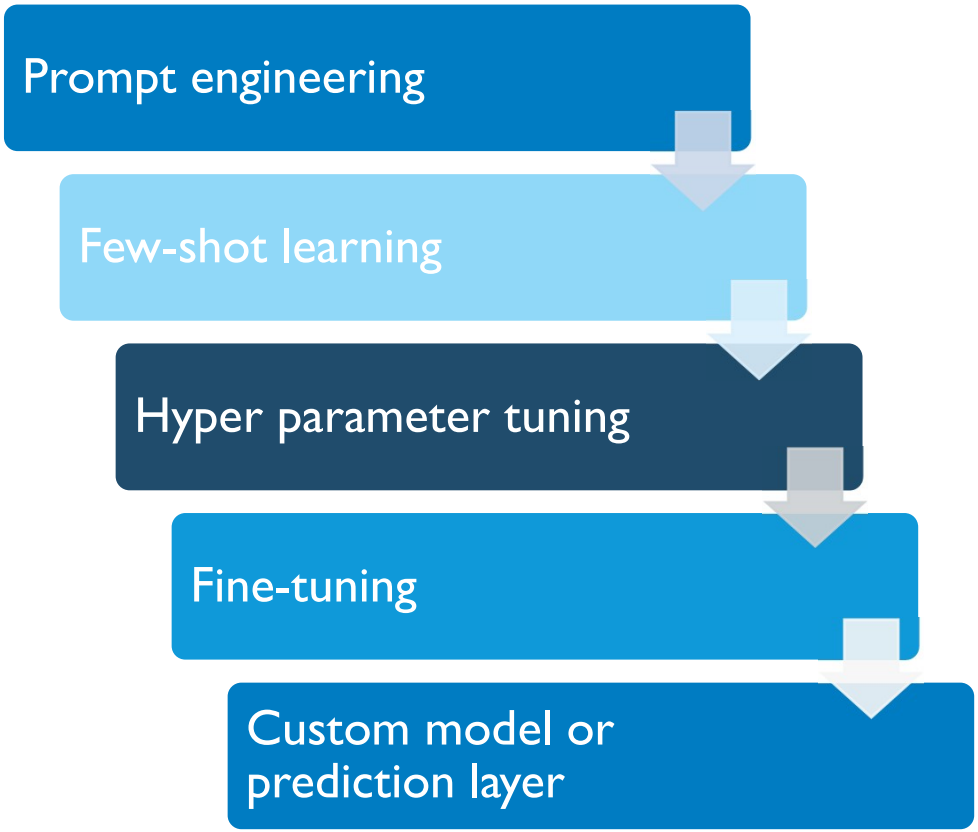
Today's focus

How is AI used in the research pipeline?

- Idea generation
- Data extraction
- Data analysis
- Literature review
- Coding support
- Data visualization
- Scientific writing
- Editing and Review

Introduction to Prompt Engineering

Improving LLM predictions: Order of Operations



Prompt Engineering

- The art of asking good questions
- Easiest, fastest way to improve your model
- Consider model selection and temperature (hyperparameters)
- Most models work better with one question, consider splitting multipart questions into multiple points

Prompt engineering basics

Contextualize	Provide context
Specify	Specify a persona
Delimit	Delimit text, triple quotes are often used
Instruct	Provide more detailed instructions about the results you want
Give Examples	Provide examples of desired output, including substantive output and formatting
Frame positively	Say what you DO want
Specify	Specify steps the model should take (if necessary)

Context Setting

ChatGPT

What book could I use to help 8th grade students understand the American revolution?

Memory updated

For 8th graders, a great book to help them understand the American Revolution is *Chains* by Laurie Halse Anderson. It's a historical novel that follows a young enslaved girl named Isabel, who is caught between the British and Patriots in New York City. The story brings the complexities of the Revolution to life from a personal perspective, helping students understand the broader themes of freedom, loyalty, and resilience.

Another good option is *The American Revolution: A Nonfiction Companion to Revolutionary War on Wednesday* by Mary Pope Osborne and Natalie Pope Boyce.

Context Setting

Remember that "current" is 2021

You are helpful teaching assistant who is helping to provide feedback on how to link social studies lessons to literature that is currently popular with young adults.

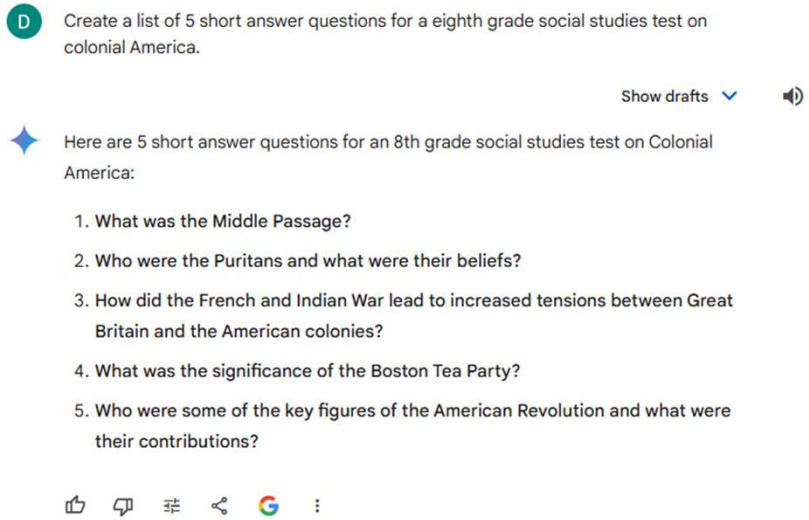
What book could I use to help 8th grade students understand the American revolution?

Memory updated

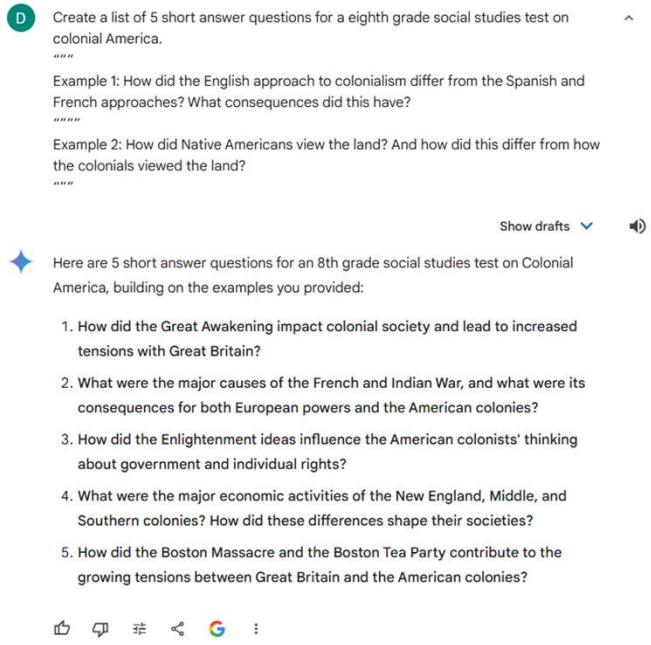
To link the American Revolution with literature that resonates with young adult readers, consider these options:

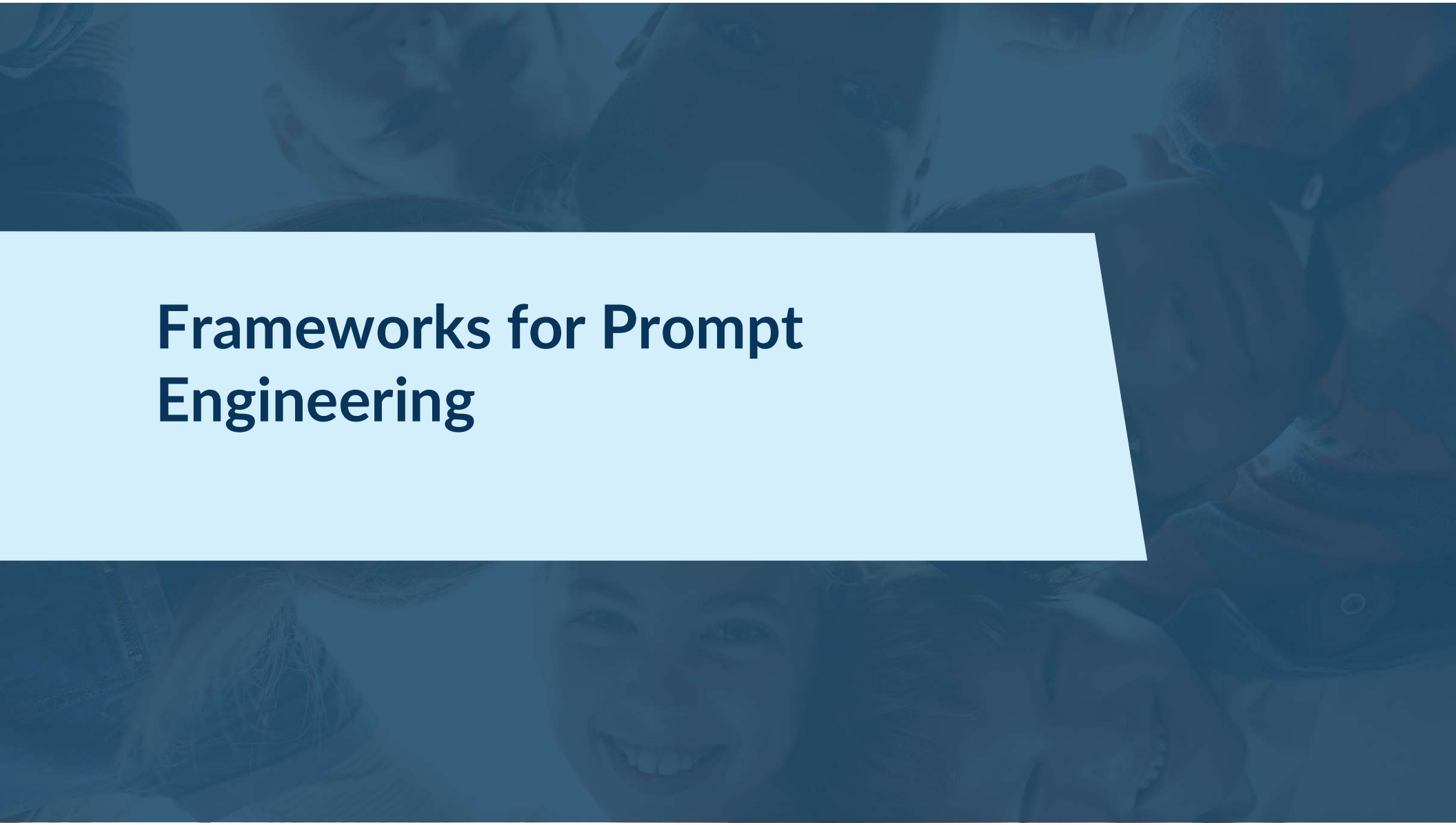
1. **Hamilton and Peggy!: A Revolutionary Friendship** by L. M. Elliott
   This book taps into the popularity of *Hamilton: An American Musical* by focusing on Peggy Schuyler, the lesser-known but sharp and witty sister of Angelica and Eliza. It gives students a way to understand the Revolution through the eyes of a young woman involved in the movement while engaging with a character they may recognize from the musical.
2. **My Brother Sam Is Dead** by James Lincoln Collier and Christopher Collier
   This classic novel gives a gritty, realistic portrayal of the Revolution through the story of Tim Meeker, whose family is divided between the Loyalists and the Patriots. The themes of family loyalty and moral ambiguity resonate with young adult readers who appreciate complex narratives.
3. **Revolutionary Friends: General George Washington and the Marquis de Lafayette** by Selene Castrovilla
   This picture book is accessible yet meaningful for 8th graders. It explores the friendship between Washington and Lafayette, a relationship that has fascinated many and is often referenced in popular history materials, including *Hamilton*.

Providing Examples

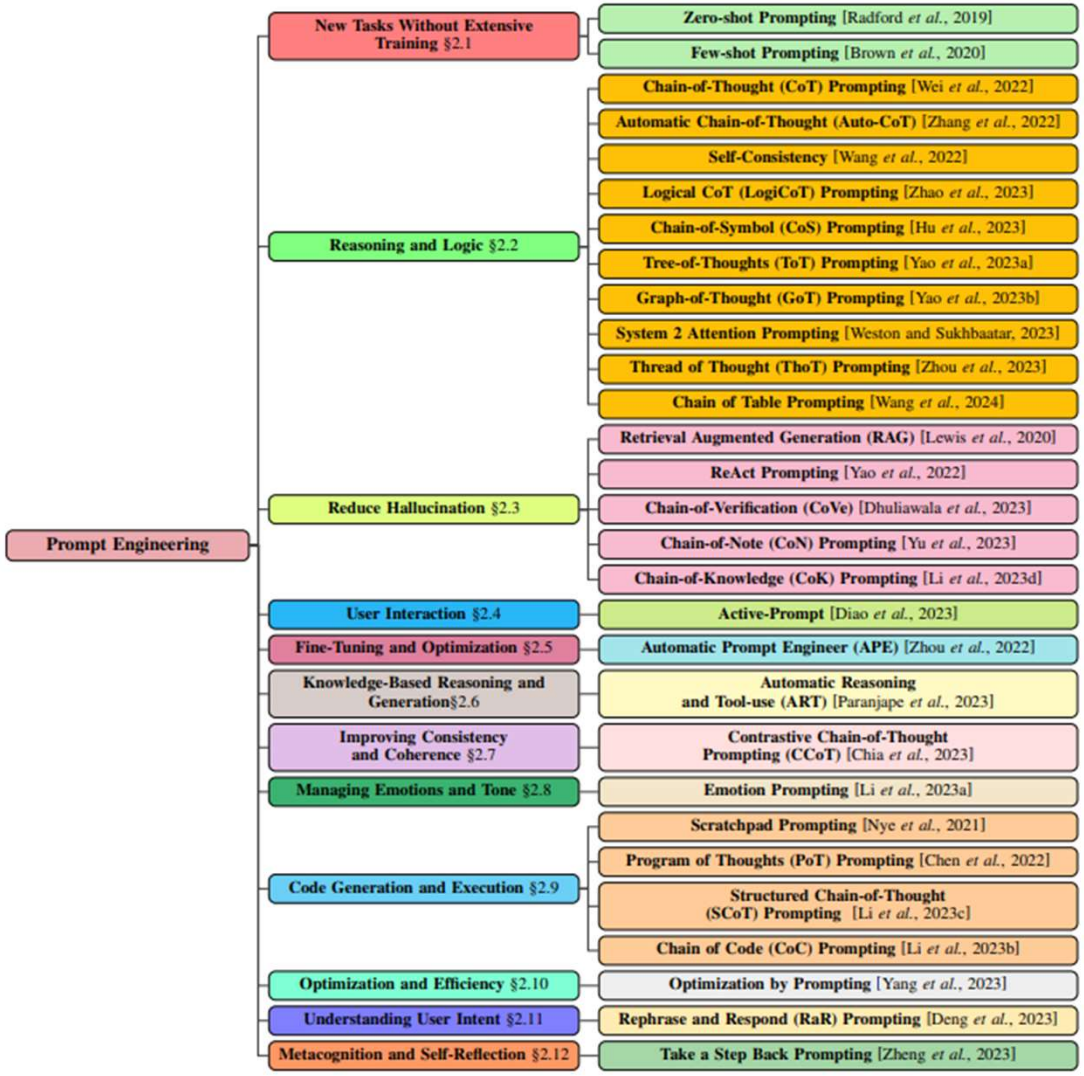


Providing Examples





Frameworks for Prompt Engineering



Research on Prompt Engineering

- Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. *arXiv preprint arXiv:2402.07927*.

Some top frameworks for research

- ReAct prompting (Yao et al,. 2022)
- Decomposed prompting (Khot et al., 2022)
- Chain of Thought prompting (Wei et al., 2022)
- Choice of prompting framework should be driven by the type of task; compare to published benchmarks where possible

ReAct framework

- Interweaves (1) reasoning and (2) task specific actions
- Good results for question answering, and ability for researcher/user to intervene
- Good ReAct prompts guide the model to think through the specific steps needed to answer
- Questions are answered step by step

ReAct framework: Example

Claim	Stranger Things is set in Bloomington, Indiana.
Thought 1	I should search for Stranger Things, and see if it is set in Bloomington, Indiana.
Action 1	Search[Stranger Things]
Observation 1	Stranger Things is an American science fiction horror drama television series created by the Duffer Brothers. Set in the 1980s, primarily in the fictional town of Hawkins, Indiana, the series centers on a number of mysteries and supernatural events occurring around the town and their impact on an ensemble of child and adult characters.
Thought 2	The observation says that it is set in a "fictional town of Hawkins, Indiana", so it is not set in Bloomington.
Action 2	Finish[REFUTES]

Yao et al., 2023

Decomposed Prompting

- Divide and conquer strategy
- Initial prompt decomposes the request into multiple sub-tasks which can be farmed out to sub-prompts, functions or other handlers
- Allows optimization and prompt engineering of all sub-components of the system (not all sub systems need to be llm prompts)
- Designed to handle complex and multi step tasks (as are commonly seen in research uses of AI)

Decomposed Prompting: Example

Note signposting

Decomposer prompt:

QC: Concatenate the first letter of every word in “Social Data Science" using spaces Q1: [split] What are the words in " Social Data Science "? #1: [“social”,”data”,”science”] Q2: (foreach) [str_pos] What is the first letter of #1? #2 [“s”, “d“,”s”] Q3: [merge] Concatenate #2 with spaces #3: “s d s" Q4: [EOQ]

split prompt

Split the provided sentences by spaces and return as a comma separated list. E.g. “Prompt: split the phrase “test train validate”; Response: [‘test’,’train’,’validate’]

…

Chain of Thought

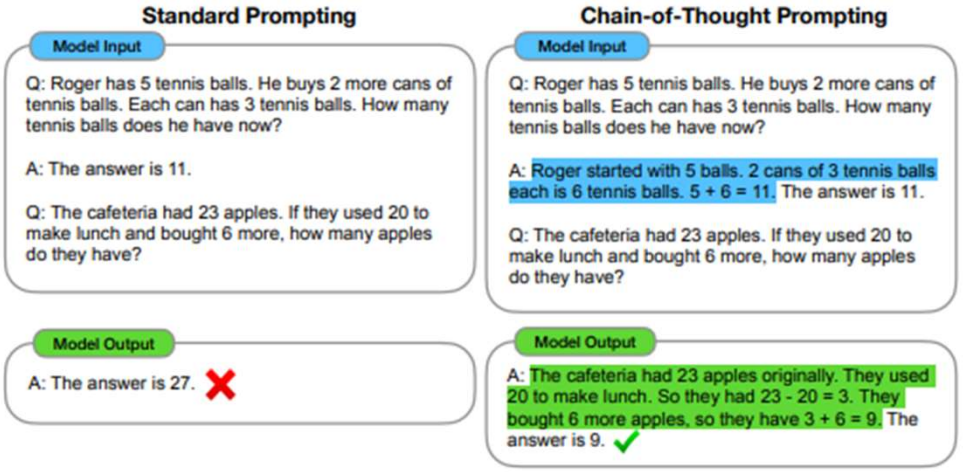


Figure 1: Chain-of-thought prompting enables large language models to tackle complex arithmetic, commonsense, and symbolic reasoning tasks. Chain-of-thought reasoning processes are highlighted.

Image from Wei et al., 2022

Why use chain of thought?

- Measurable improvements in performance for reasoning based tasks
- Provides transparency (and thus the ability to validate) into steps taken
- Extends the ability of LLMs in math and reasoning tasks
- Readily applicable to off the shelf large language models

How to use frameworks

Custom written prompts (with examples)

Can include specifications to look for information in specific sources

Wrappers

Such as langchain

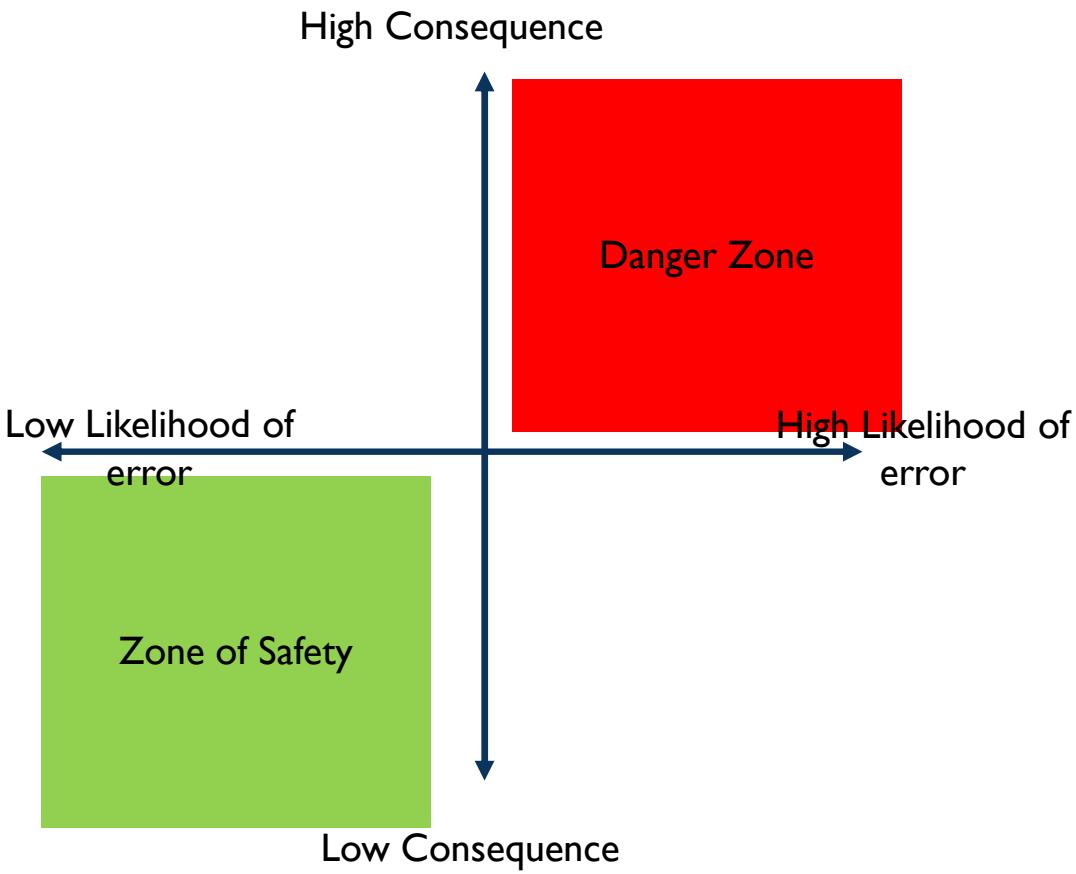
Testing and Validating LLM results

How do we validate AI work

- Manual Review
  - Best for project that produce one off products, or where incorporation in a human workflow is efficient
- Validation against manually coded data
  - Best for pipelines that need to be used without human reviews
  - Requires consensus among manual coders, and consensus on acceptable accuracy
  - Split your data: Test/Train
- Repeated/multiple predictions

How to think about risk level?

- Is the system providing ***live responses?***
  - Who is interacting with the system?
- How consequential would getting it wrong be?
- Are there opportunities for human oversight?
- How public is this?
- What are the ***benefits*** *of this?*



What AI project would be in each quadrant?

In sum:

- Never use AI without incorporating a validation element
- When validating against manually coded data, validation data can be used only once.
- Consider the risk level of an application (both likelihood and consequence)
- Communicate clearly with project teams

Two Case Studies: AI to support research

How did AI review 26,000 papers

- Case Study 1: Developing a research agenda for adolescent health (forthcoming, OPRE)
- Case Study 2: Streamlining a systematic review of healthy marriage and relationship education program

Questions?

Claire Kelley - Ckelley@childtrends.org

Child Trends®

References

- Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A systematic survey of prompt engineering in large language models: Techniques and applications. *arXiv preprint arXiv:2402.07927*.
- Khot, T., Trivedi, H., Finlayson, M., Fu, Y., Richardson, K., Clark, P., & Sabharwal, A. (2022). Decomposed prompting: A modular approach for solving complex tasks. *arXiv preprint arXiv:2210.02406*.
- Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship*, *49*(4), 102720.
- Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023, January). React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR)*.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, *35*, 24824-24837.