ChatAI 101

An Introduction to ChatGPT, Google Gemini, Microsoft Copilot, and Claude2

Instructed by: Ken Bray

The information in this presentation is current as of February 2024 and should be considered valid for a period of 2 months. Be sure to check the revision history of AI Chatbots, as technology is always advancing.

Agenda

Part I: The Basics

- Objectives
- Introduction
- The Rise of Chatbots
- What's in AI?
- ELI5: How do they work?
- How can AI be used?

Part II: Putting it to Work

- Introductions to Chatbots
- Limitations
- Prompt Engineering
- Best Practices
- Ethical Considerations
- Discussion and Q&A

Prerequisites

To get the best experience from this session, you will need access to the following

- Computer
- ChatGPT, Free or Paid Account
- Google Gemini Account
- Microsoft Copilot Account (Benchmark Login)
- Anthropic Claude 2, Free or Paid Account

Objectives

- Walk away with a basic understanding of how AI Chatbots work.
- Understand the capabilities (and limitations) of each platform.
- Learn Prompt Engineering, and best practices.
- Learn tools you can put to immediate use.

Introduction

- Did you know Chatbots have been in existence since 1966?
- Some examples of “Chatbots” used in everyday life include predictive text feature on most modern smartphones, as well as “Hey Siri”, “Hey Google”, and “Alexa”.
- Due to the popularity and availability of ChatGPT, we’re seeing a rise in the number of Chatbots.

The Rise of (AI in) Chatbots

- OpenAI + ChatGPT = Catalyst in Technology.
- ChatGPT paved the way for Google, Microsoft, Claude, and others.
- AI Integration has revolutionized chatbots!
- Natural Language Processing, Machine Learning, and Deep Learning have rapidly advanced the technology at alarming rates.

What's in the AI of a Chatbot?

- Software == Car
- Model == Engine

Software

- The car you drive.
- ChatGPT, Google Gemini, Microsoft Copilot, and Claude are commonly referred to as Software.
- Software is what you are interacting with, User-interface.

Model

- The engine that powers the car you drive.
- The Architecture behind Chatbots are commonly referred to as Models (Language Models or Large Language Models).
- Models have names like GPT, PaLM, StableLM, LLaMa, MPT, etc.

ELI5: How do Chatbots Work?

- Think of a Chatbot like a friendly robot.
- Ask questions, just like you would ask a friend or a teacher.
- The Chatbot will give you an answer.
- You can keep the conversation going by asking more questions or talking about something else.

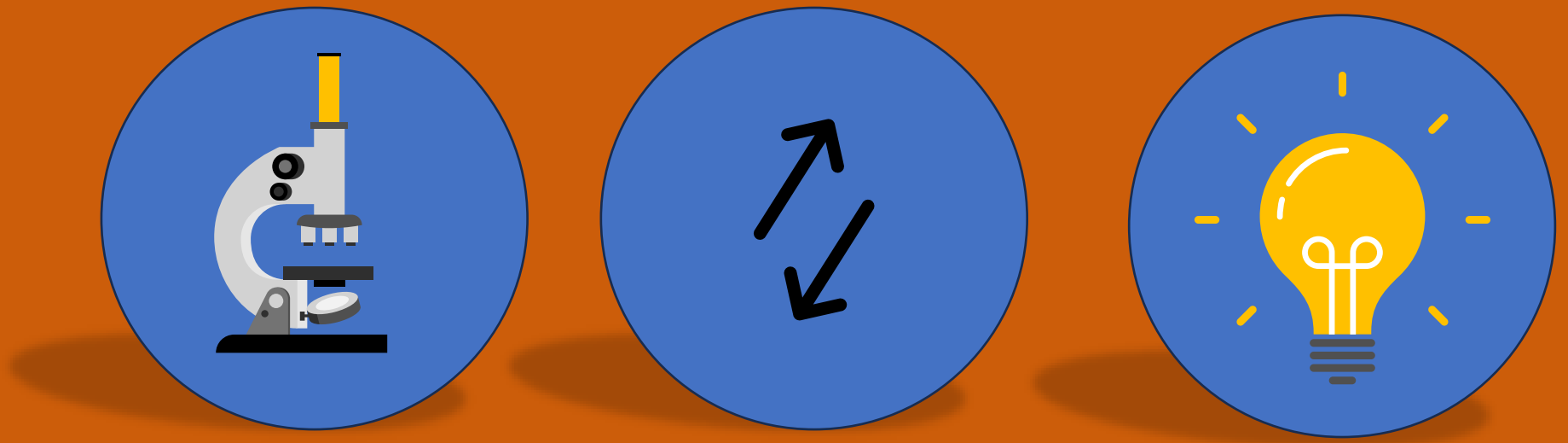
How does it <u>really</u> Work?

- Machine Learning – this is where the models are trained.
- Deep Learning – this is where the models are trained to grasp the nuances of our language and understand word associations.
- Algorithms – acts as a recipe to guide the models to appropriate answers based off calculated word associations.

How does it <u>really</u> Work? (Cont.)

- NLP – Natural Language Processing (this allows the model to communicate on a human-like level).
- NLP Challenges – knowing the different between an apple, and Apple; picking up on sarcasm, etc.
- NLP in Action – involves tokenization, contextual analysis, and output generation.

How can AI be used?



Find Organize Generate

Part 1: Takeaways

- AI Chatbots work of machine and deep learning, as well as natural language processing.
- Chatbots lack feeling and awareness, only mimic our human language.
- Operate on algorithms, which can be easily detected (ZeroGPT, GPTZero, etc.).
- AI chatbots fulfill three primary roles; find, organize, and generate.

Introduction to your Advanced AI Chatbots

- ChatGPT
- Microsoft Copilot
- Google Gemini
- Claude 2

ChatGPT

- ChatGPT, developed by OpenAI, is built on the Generative Pre-trained Transformer (GPT) models.
- Its development marked a significant leap in making AI conversational agents more powerful and versatile.
- ChatGPT stands out for its ability to understand context, generate human-like responses, and learn from interactions to improve its performance.

Copilot

- Developed by OpenAI for Microsoft, is built on the Generative Pre-trained Transformer (GPT) models.
- Microsoft Bing stands out for its combination of search engine functionality with advanced AI features, incentives for use, and unique integrations within the Microsoft ecosystem.
- Very similar (if not the same) as ChatGPT, with the added benefit of Company protection (if using your Benchmark login).

- Developed by Google, is built from the family of LaMDA, PaLM, and Gemini models.
- Its development is Google’s direct response to the release and popularity of Chat GPT.
- Gemini stands out for its deep integration into Google’s ecosystem, which provides it with a broad knowledge base, up-to-date information, and versatile response capabilities that stand out when compared to ChatGPT.

Claude 2

- Developed by Anthropic (by former OpenAI Architects), is built on similar Generative Pre-trained Transformer (GPT) models.
- Its development encompasses safe and ethnical AI practices, resulting in improved safety measures aimed at minimizing harmful or biased outputs.
- Claude stands out for advanced capabilities in creative writing, content creation, data analysis, and academic skills (math, coding, reasoning).

Limitations

ChatGPT

- Data Timeliness
- Understanding and Generating Complex Contexts (Hallucinations)

Microsoft Copilot

- Content Creation Limits
- Bias and Ethical Concerns

Google Gemini

- Accuracy and Reliability
- Contextual Depth

Claude 2

- User Intent
- Safety and Filtering

Operational Limitations

- There is a possibility that you may run into access limits* (fairly common with the FREE accounts for ChatGPT & Claude 2)
- Some of the exercises demonstrated later (ChatGPT) are only features associated with a paid account

*Switch to Google Gemini or Microsoft Copilot

Prompt Engineering

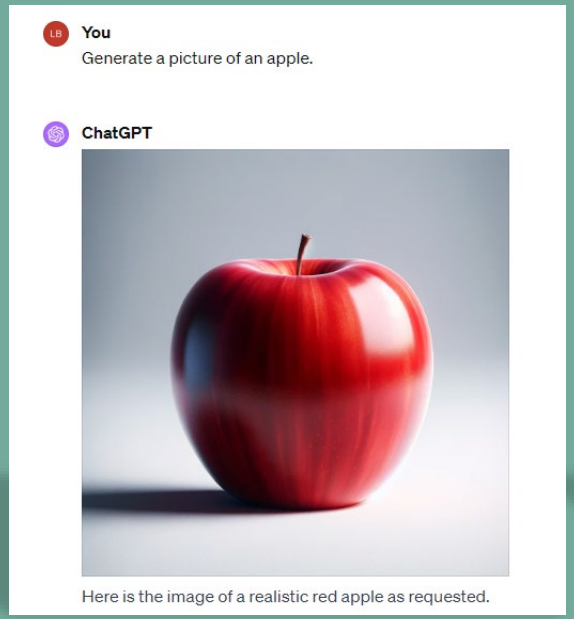
Science Behind the Prompt

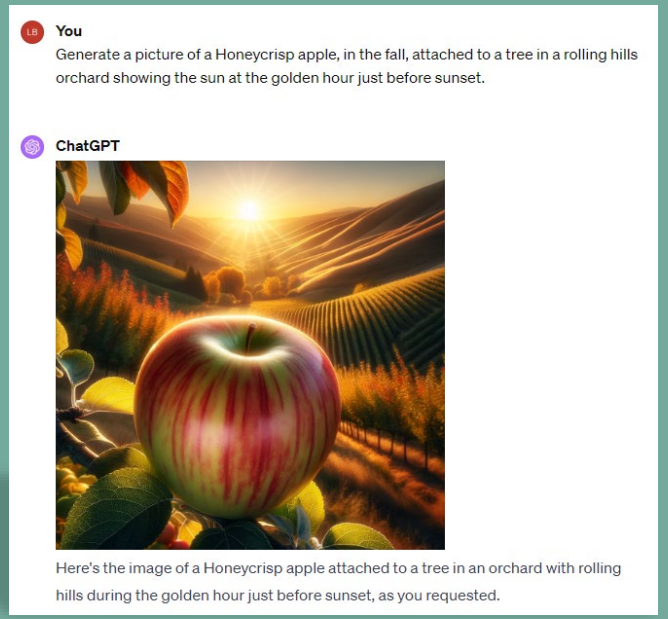
- Machine & Deep Learning gives AI the power to learn and understand our complex language.
- Each word is treated as a token and ran through a complex algorithm to formulate its response.
- AI looks at each token, what comes before and after, and can calculate a response.

Prompt Engineering

Feed the Algorithm!

- The more relevant words you use, the better your prompts will perform.





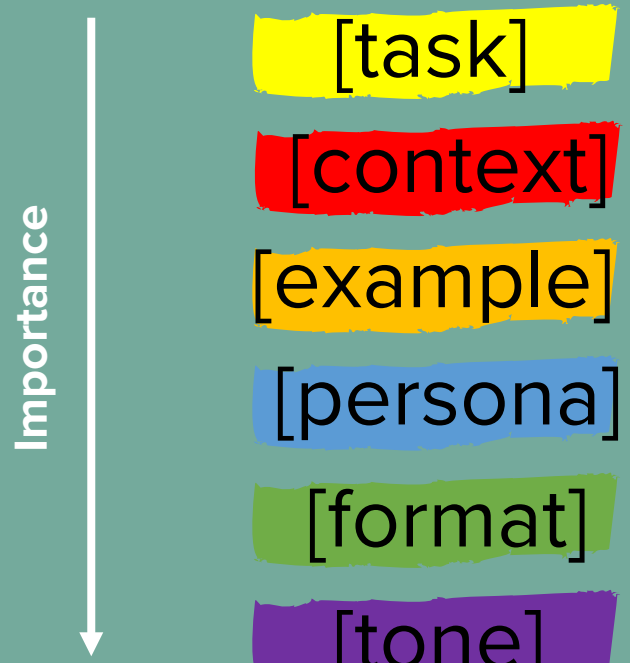
Prompt Engineering

Basic (aka “RACE” framework)

- Role – You are…
- Action – I need you to…
- Context – Here is some background info…
- Execute – Do the action this way…

Prompt Engineering

Advanced Framework



[task]	Ask, “I want you to [insert verb]...”
[context]	Provide background information...
[example]	Provide an example or structure...
[persona]	“You are [insert persona]...”
[format]	Tell AI what you want to see...
[tone]	[tone]

Prompt Engineering: Most Important

Pro tip: Always provide a task, with context.

[task]

[context]

“I am a content creator in the construction industry, please generate a weekly social media plan for the month of February.”

Prompt Engineering: Important

Pro tip: When available, provide an example to tailor your response.

[context]

[example]

“I need to write job descriptions for an open superintendent position for a construction company. Please draft the job description using the format of this existing job description below:”

Prompt Engineering: Good to Have

Who do you want AI to be?

[persona]

“You are a [hiring manager]…”
“You are a [social media expert]…”
“You are a [data analyst]…”

Prompt Engineering: Optimal Input

Pro tip: Adding a persona adds more context, increases word associations = better outcomes

[persona]

[context]

[task]

[example]

"You're a hiring manager responsible for writing job descriptions for an open superintendent position for a construction company. Please draft the job description using the format of this existing job description below:"

Prompt Engineering: Low Importance

Pro tip: Adding formatting can help organize outputs and data. Also, when including [format] in a prompt, including an [example] is unnecessary.

[format]

"Generate a social media content calendar for Instagram, highlighting women in construction, for each week in the month of March. Output the results in a table format with column headers: Week, Theme, Content."

Prompt Engineering: Low Importance

[format]

“Generate a social media post about a recent groundbreaking ceremony at the ‘Highlands at Wyomissing’. Make sure the post is written in the active voice and follows the AP style guide.”

Prompt Engineering: Low Importance

Pro tip: The average America literacy range falls around a 7th to 8th-grade reading level.

[format]

“Summarize the following article so that it is on an eighth-grade reading level.”

Prompt Engineering: Nice to Have

[tone]

“Make sure the results sound professional…”
“Use a casual tone…”
“Be sure to include a witty response”
“Sound pessimistic…”
“Show enthusiasm…”

Prompt Engineering Example

“You are a marketing specialist and content creator in the construction industry. For the month of March, you're going to focus on women in construction. Please generate a social media content calendar for the month. Be sure to output the results in a table format with the following headers: Week, Theme, Content. Make sure the theme tone is professional, encouraging, and supportive.”

Prompt Engineering Takeaways

- Provide as much information as you can.
- The more word associations that can be made by the software models, the better.
- At a minimum always supply an Action (or Task) with Context.
- Applying some form of framework to prompts will yield better results.
- Experiment between platforms.

Best Use Application

For reading and writing text under 8,000 words	ChatGPT or ChatGPT+
For reading and writing text over 8,000 words	Claude 2
For creating images	Microsoft Copilot or ChatGPT+
For analyzing images	Google Gemini or ChatGPT+ (Add-ons)
For real-time information	Microsoft Copilot or Google Gemini
For analyzing data	ChatGPT+
For writing code	ChatGPT or ChatGPT+

Best Practices

- Avoid using sensitive data (personal and or financial information).
- Avoiding asking AI to perform tasks above your level of understanding.
- Use clear and specific questions; Garbage in = Garbage out.
- Be aware of the tool’s limits; check your work continue to refine responses by asking more questions!

Ethical Consideration

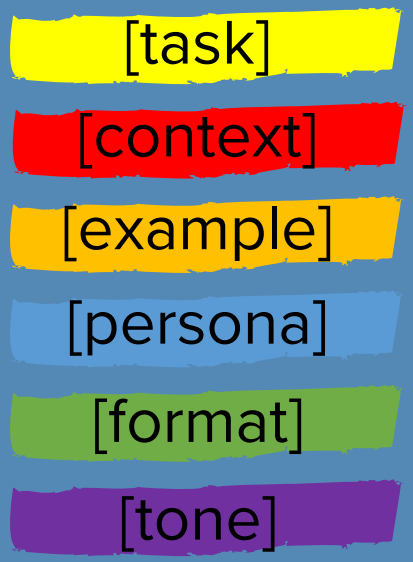
- Respect Copyright & Intellectual Property
- Respect other’s rights to privacy, safety, and well being.
- Cultural Bias - Learns from Humans, Humans suck.
- Should we be respectful in our interactions with AI?

Discussion and Q&A

Prompt Exercise 1: Rewriting

- Write a prompt that explains a difficult concept in our industry using the “RACE” or “Advanced” framework.

- Write a prompt that changes your tone in a recent email using the “RACE” or “Advanced” framework.



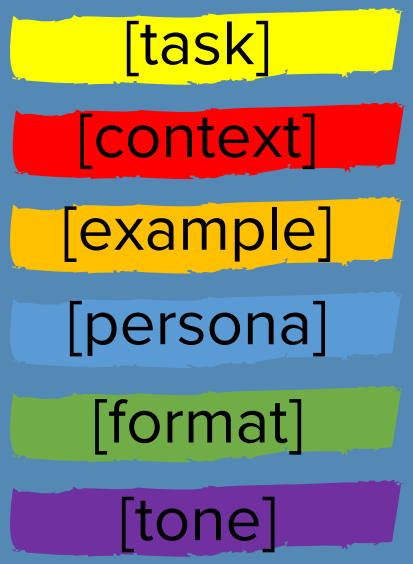
Prompt Exercise 2: Summarizing

- Find an article from the news (or industry website, etc.) and write a prompt to summarize the article using the “RACE” or “Advanced” framework.

Prompt Exercise 3: Extraction

- Write a prompt to extract data from an article, email, or file using the “RACE” or “Advanced” framework.

- Write a prompt to extract data from a public LinkedIn profile, highlighting a persons abilities using the “RACE” or “Advanced” framework.



Prompt Exercise 4: Generating

- Write a prompt that will generate some form of content (social media post, LinkedIn post, blog post) using the “RACE” or “Advanced” framework.