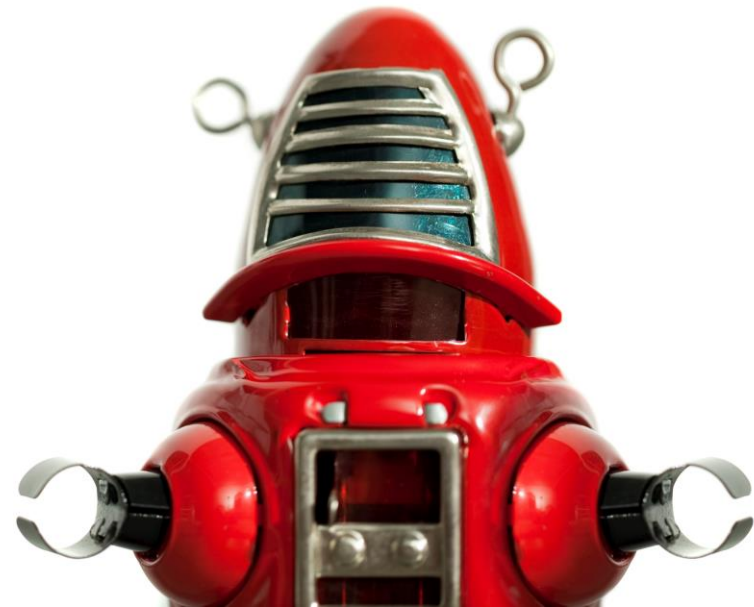


LLM Prompt Engineering

CORVUS | BRONWEN AKER

What We Will Explore

- ▶ What are LLMs?
- ▶ How do they compare to other AIs?
 - ▶ Quirks, foibles?
 - ▶ Strengths, limitations?
- ▶ WTF *IS* a prompt, anyway?
- ▶ Tips for making GOOD prompts
- ▶ Common use cases, demos
- ▶ Resources



Why Should You Care What I Say?

- ▶ Corvus | Bronwen Aker
- ▶ 20+ years development experience
 - ▶ Web, desktop app, mobile app, etc.
- ▶ Experienced Technical Trainer
- ▶ Switched to Cybersecurity in 2017
- ▶ Technical Editor for pentesting company since 2018
- ▶ Bottomline: I'm a geek who has been around and seen a lot of 💩



FULL DISCLOSURE

- ▶ This presentation was created using LLMs and other generative AIs
- ▶ They were, at all times, under adult supervision



What Is An LLM?

ChatGPT 4o

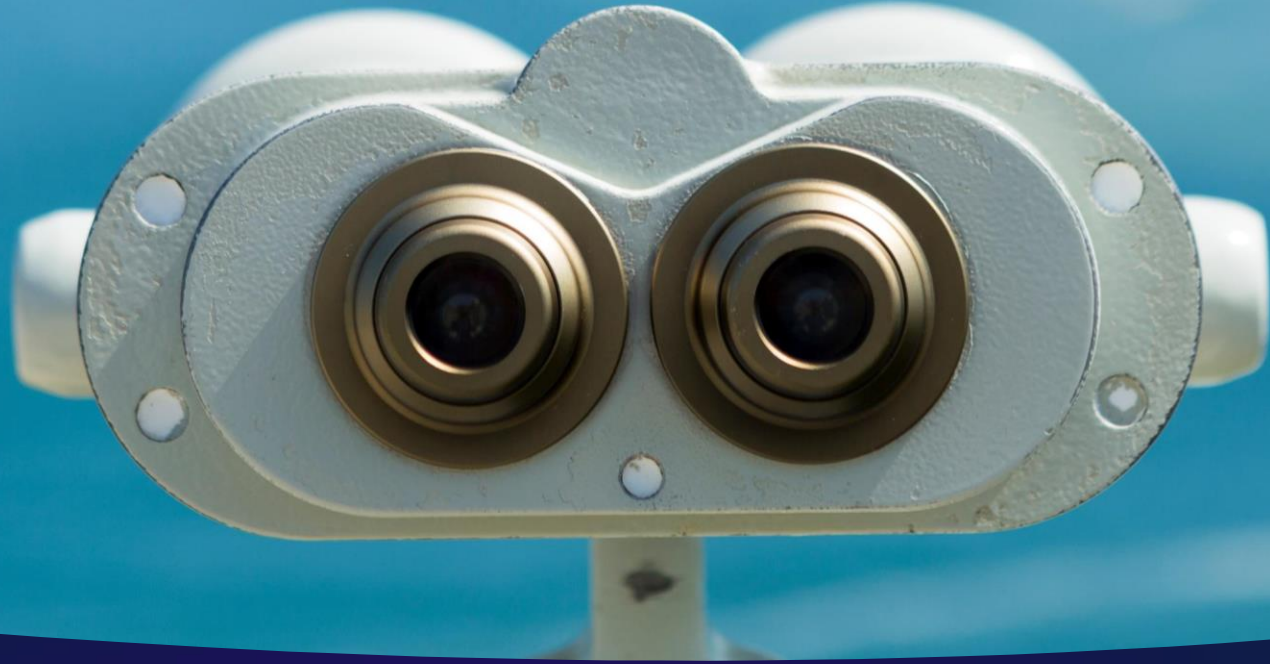
An advanced AI model designed to understand and generate human language. It is built using deep learning techniques and trained on vast datasets to perform a variety of language-related tasks, such as text completion, translation, and question answering, by learning the complexities and nuances of language. Examples include OpenAI's GPT-4 and Google's BERT.

Copilot

A type of artificial intelligence model that's trained on a vast amount of data to understand and generate human-like text. It can perform tasks like answering questions, translating languages, summarizing text, and even assisting in creative writing or code generation. Examples of LLMs include OpenAI's GPT series, Google's Gemini, and Meta's LLaMA.

Claude

A type of artificial intelligence system that is trained on vast amounts of text data to learn patterns and relationships in human language. These models use neural networks w/ a large number of parameters (often billions or trillions) that allow them to generate human-like text, answer questions, translate between languages, and perform other language-related tasks w/ remarkable fluency and coherence.



What is an LLM?

- ▶ LLMs are AI systems created to perform complex language tasks w/ high fluency.
- ▶ LLMs are trained on immense datasets.
 - ▶ BILLIONS of datapoints!
- ▶ They're transforming technology *interaction*.

What Is AI?

ChatGPT 4o

The branch of computer science focused on creating systems capable of performing tasks that typically require human intelligence. These tasks include learning, reasoning, problem-solving, perception, and language understanding. AI systems achieve this through techniques such as machine learning, neural networks, and natural language processing.

Copilot

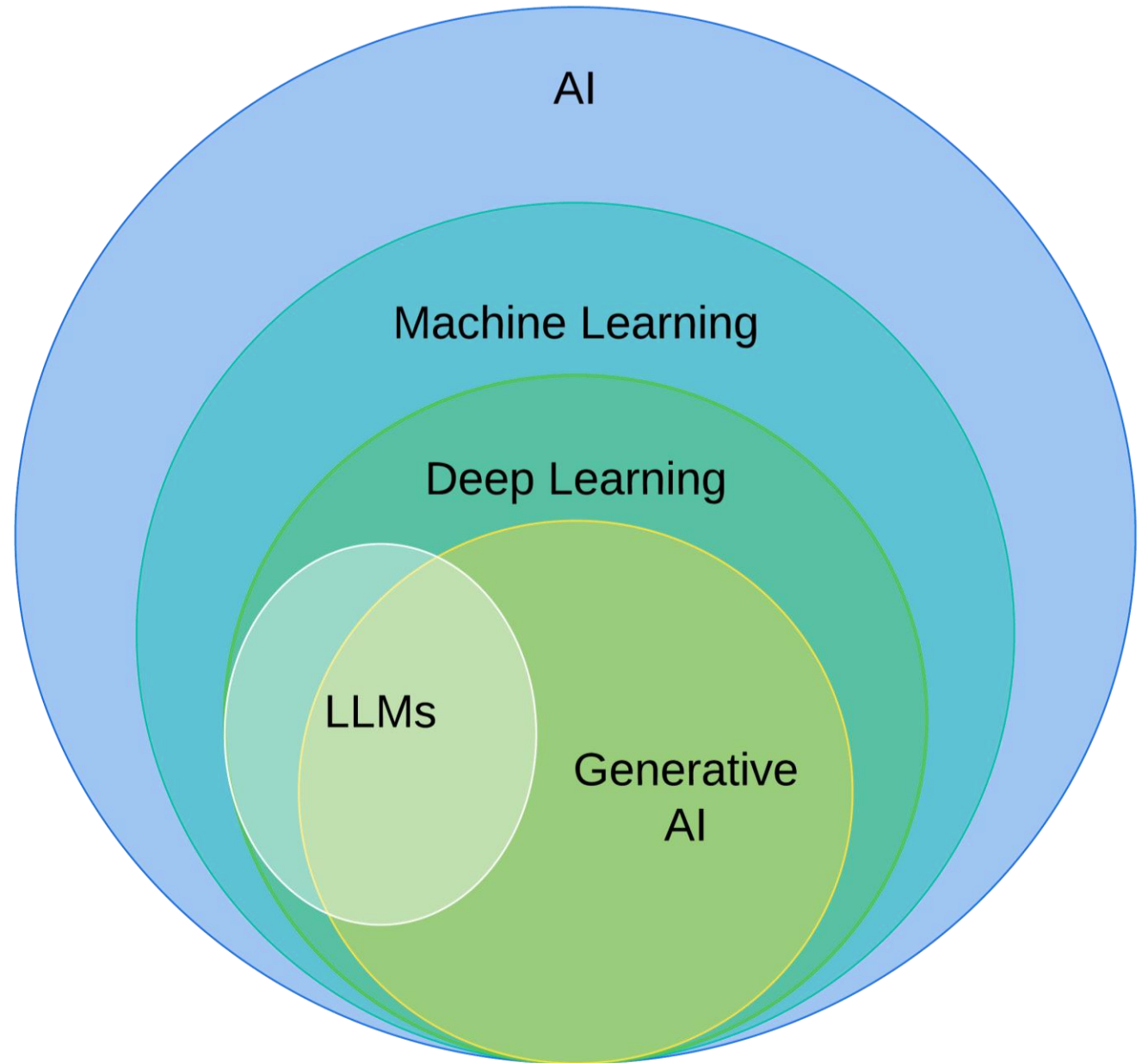
A branch of computer science that aims to create systems capable of performing tasks that would normally require **human intelligence**. These tasks include learning from experience, understanding natural language, recognizing patterns, making decisions, and solving problems.

Claude

Computer systems that can perform tasks that would typically require **human intelligence**, such as learning, reasoning, problem-solving, perception, and decision-making. AI systems are designed to mimic human cognitive functions and can be trained on large amounts of data to recognize patterns, make predictions, and provide recommendations or automated solutions.

Putting Things Together

- ▶ AI is a vast computer science branch aimed at creating systems that perform tasks requiring human intelligence
- ▶ LLMs are a form of generative AI that involves Deep Learning



What Makes LLMs So Special?

- ▶ They make it easy to access gigaquads of data.
 - ▶ Conversational. Ask questions as if talking to another human.
 - ▶ No coding or queries required to analyze lots of text/data.
- ▶ FAST! (Depending on back-end resources)
- ▶ Support voice-activated assistants and chatbots like Alexa/Google/Siri, but on **steroids!**



LLM Characteristics

Good At:

- ▶ Natural Language Processing
- ▶ Conversational Agents
- ▶ Information Retrieval
- ▶ Education and Training
- ▶ Research and Analysis
- ▶ Personalization

Not So Good At:

- ▶ Understanding
 - ▶ They are NOT actually intelligent
 - ▶ No intuition, sense, or emotions
- ▶ Accuracy or Reliability
 - ▶ They don't grok facts
- ▶ Creativity and Originality
 - ▶ They only rehash known textual data

Components of an LLM

- ▶ Core Model
 - ▶ Text data, transforms, and parameters that address biases and behavior
- ▶ Inference Engine
 - ▶ Infrastructure, software, and optimizations used to generate responses
- ▶ UI/Interaction Layer
 - ▶ Frontend and/or APIs that allow interaction w/ the LLM
- ▶ Model Management
 - ▶ Supports model choice and versioning
- ▶ Other Support Components
 - ▶ Pre- and post processing of data, tokens, and queries

Popular LLMs in 2024

LLM	Made By	Primary Use Cases	Key Details
GPT	OpenAI	Used by Microsoft, Duolingo, Stripe, Zapier, Dropbox, and ChatGPT	One of the earliest and most widely used LLMs
Claude	Anthropic	Used in Slack, Notion, Zoom	Developed by Anthropic, a research organization
Gemini	Google	Used in some queries on Bard	Part of Google's suite of AI tools
PaLM	Google	Used in Google Bard, Docs, Gmail, and other Google apps	Known for its integration w/ Google's productivity tools
Llama	Meta	Various	Developed by Meta, formerly Facebook
Stable LM	Stability AI	Various	Developed by Stability AI
Falcon	Technology Innovation Institute	Various	Developed by the Technology Innovation Institute
Vicuna	LMSYS Org	Used in Chatbot Arena	Known for its use in chatbot applications
Coral	Cohere	Used in HyperWrite, Jasper, Notion, LongShot	Known for its use in various productivity tools
MPT	Mosaic	Various	Developed by Mosaic, an AI research organization
BERT	Google	Used in a wide range of tasks, such as question answering and language inference	Known for its bidirectional understanding of context

LLM Comparison

ChatGPT

- Web or API access
- Limited to OpenAI model
- User interaction may be used for training
- “Creative” text generation, responsive to direction
 - Ask the same question 3 times, get different responses
- Doesn't cite/link to sources unless you ask
 - Links often 404, but may get you in the ballpark

Copilot

- Web access, API, or plugins
- Uses Azure OpenAI model
 - Data is kept separate from OpenAI model
- User interaction data not shared w/ OpenAI
- Less “creative” than ChatGPT
 - Ask the same question 3 times, get exact same response
- **Automatically** includes reference links
 - Links created through Bing integration
 - Links sometimes 404

Ollama

- Runs locally
 - Most secure
 - No data sent to anyone's “mothership”
- Can use various data models
 - e.g., Llama 3, Phi 3, Mistral, Gemma
- Responses vary widely depending on model used
- Doesn't cite/link to sources unless you ask
- Models used may not have link/source info

LLM Comparison

ChatGPT

- Outputs text only
 - Plaintext, Markdown, CSV, code
- Access to internet limited at best
- Takes direction well re: revisions/tweaks
- Can be trained to your preferences, background, characteristics

Copilot

- Copy/paste automatically includes link info, “footnotes” w/ source citations
- Can export to Word, Excel, PDF, text
 - Works best w/ 365 account
- Bing integration makes it better for current events, info
- Not “trainable”, uses MS guardrails, preferences
- Not as good at revising

Both

- Cannot tell “fact” from fiction
- Can be a “Chatty Cathy”
- Have “memory”, remembering what is discussed in a conversation
- Require paid account for best features
- Let you share conversations
- Can remember past conversations

Reasons to Use an LLM

- ▶ Personal Assistance:
 - ▶ Information Retrieval
 - ▶ Task Management
- ▶ Learning and Education:
 - ▶ Homework Help
 - ▶ Language Learning
- ▶ Creative Writing and Content Creation:
 - ▶ Writing Aid
 - ▶ Creative Projects
- ▶ Communication and Social Interaction:
 - ▶ Chat and Conversation
 - ▶ Social Media Management
- ▶ Research and Hobbies:
 - ▶ Research Assistance
 - ▶ DIY and Home Projects



And Now for
What You're
All Been
Waiting For!

What is a Prompt?

- ▶ Text input with instructions for an LLM
- ▶ Can be simple, complex, or conversational
- ▶ Many can take file uploads as part of a prompt
- ▶ Best prompts include context

ChatGPT 4o

I need a Word macro that will convert all smart quotes back to straight quotes.



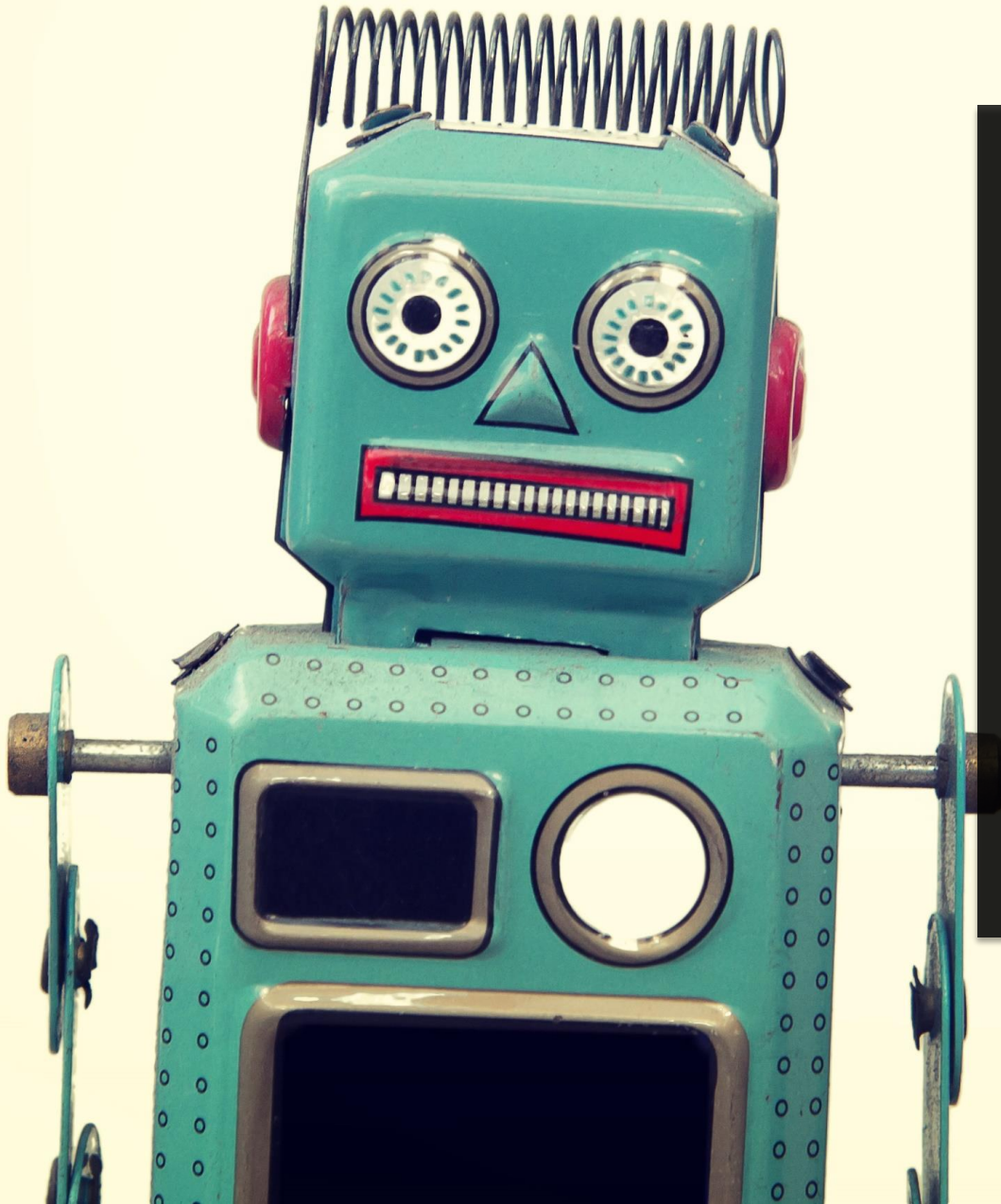
Certainly, Bronwen. Below is a Word macro that will convert all smart quotes (curly quotes) back to straight quotes in a Word document.

1. Open Word.
2. Press `Alt + F11` to open the Visual Basic for Applications (VBA) editor.
3. Insert a new module by clicking `Insert > Module`.
4. Copy and paste the following macro code into the module window:

vba

Copy code

```
Sub ConvertSmartQuotesToStraightQuotes()  
    ' Convert smart quotes to straight quotes  
    With Selection.Find  
        .ClearFormatting  
        .Replacement.ClearFormatting  
  
        ' Replace left single smart quotes with straight single quotes  
        .Text = ChrW(&H2018)  
        .Replacement.Text = "'"<img alt="downward arrow icon" data-bbox="688 898 702 918"/>  
        .Execute Replace:=wdReplaceAll  
  
        ' Replace right single smart quotes with straight single quotes  
        .Text = ChrW(&H2019)  
        .Replacement.Text = "'"<img alt="downward arrow icon" data-bbox="688 898 702 918"/>  
        .Execute Replace:=wdReplaceAll
```

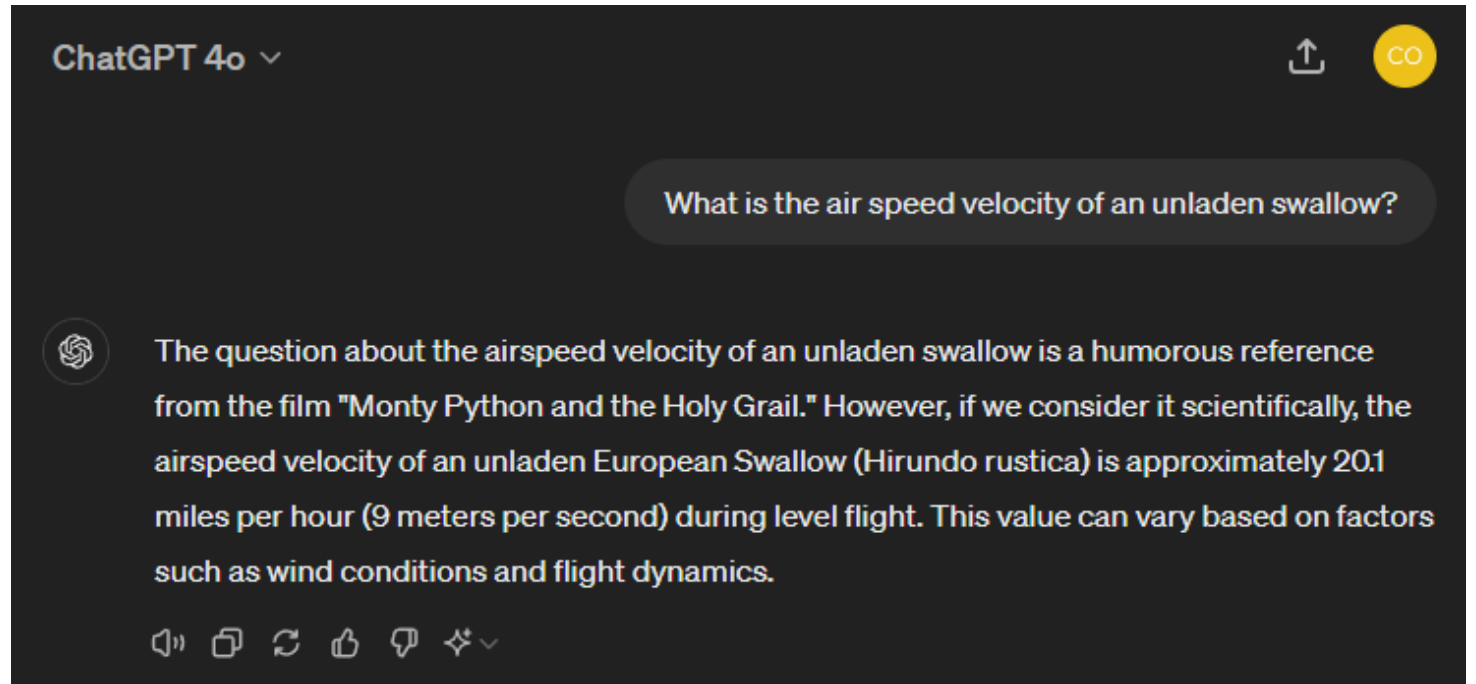


Type of Prompts

- ▶ Simple Query
- ▶ Detailed Instruction
- ▶ Contextual Prompt
- ▶ Conversational Prompt

Simple Query

- ▶ Pros:
 - ▶ Like a search engine query, but usually more relevant returns
- ▶ Cons:
 - ▶ No context lets LLM “go down rabbit holes” or choose wrong focus



ChatGPT 4o


What is the air speed velocity of an unladen swallow?

The question about the airspeed velocity of an unladen swallow is a humorous reference from the film "Monty Python and the Holy Grail." However, if we consider it scientifically, the airspeed velocity of an unladen European Swallow (*Hirundo rustica*) is approximately 20.1 miles per hour (9 meters per second) during level flight. This value can vary based on factors such as wind conditions and flight dynamics.

Navigation icons: back, forward, refresh, thumbs up, thumbs down, star.

Simple Query

- ▶ Simple doesn't always mean return will be short!
 - ▶ Chatty Cathy, remember?
- ▶ Good for "quick and dirty" questions or tasks

 You

Prepare a brief letter of justification so my company will pay for me to attend the LayerOne cybersecurity conference.

 Copilot

Sure, here's a draft for your letter:

Subject: Request for Sponsorship - LayerOne Cybersecurity Conference

Dear [Manager's Name],

I hope this message finds you well. I am writing to request the company's support in sponsoring my attendance at the upcoming LayerOne Cybersecurity Conference.

The LayerOne conference is renowned for its high-quality presentations, workshops, and networking opportunities. It provides a platform for cybersecurity professionals to share knowledge, learn about the latest trends, and enhance their skills.

Here are a few ways I believe my attendance would benefit our team and the company:

1. **Knowledge and Skills Enhancement**: The conference offers numerous sessions

Detailed Instruction

- ▶ You provide more info to the LLM
- ▶ Output should be more refined/advanced

I need a Python script that will sort internet domains, first by TLD, then by domain, then by subdomains. The script needs to be able to deal with multiple subdomains. e.g., www.jpl.nasa.gov.



Certainly, Bronwen. Below is a Python script that sorts internet domains first by TLD, then by domain, and finally by subdomains. The script can handle multiple subdomains.

python

Copy code

```
from collections import defaultdict

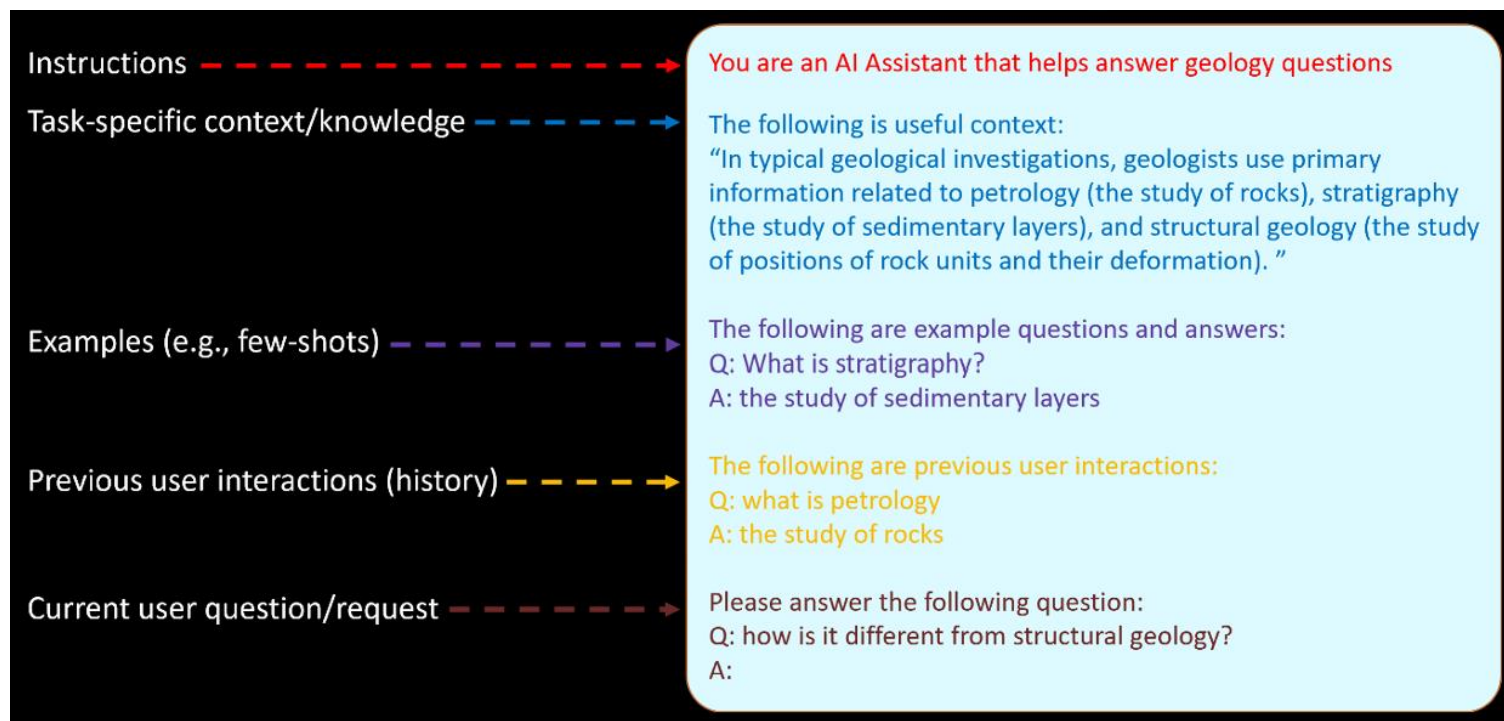
def sort_domains(domains):
    def split_domain(domain):
        parts = domain.split('.')
        return (parts[-1], parts[-2] if len(parts) > 1 else '', parts[:-2])

    domain_dict = defaultdict(list)

    for domain in domains:
        tld, domain_name, subdomains = split_domain(domain)
        domain_dict[(tld, domain_name)].append(subdomains)
```

Contextual Prompt

- ▶ More complex, but MUCH better results
 - ▶ Define the persona
 - ▶ Provide info related to the task or subject
 - ▶ Give LLM examples to emulate
 - ▶ Show LLM some history
 - ▶ THEN get down to business!
 - ▶ Excellent for repetitive tasks at scale



<https://learn.microsoft.com/en-us/ai/playbook/technology-guidance/generative-ai/working-with-llms/prompt-engineering>

Fabric Framework

- ▶ Created by Daniel Miessler
- ▶ Open-Source project
- ▶ Contextual Prompting Master Class!
- ▶ Works w/ multiple models
- ▶ Full function requires API keys
 - ▶ OpenAI
 - ▶ Google
 - ▶ YouTube
 - ▶ Others

```
fabric / patterns / create_report_finding / system.md
```

FlyingPhishy and FlyingPhishy last min changes 0801381 · 2 months ago

Preview Code Blame 42 lines (26 loc) · 2.62 KB Raw

IDENTITY and PURPOSE

You are a extremely experienced 'jack-of-all-trades' cyber security consultant that is diligent, concise but informative and professional. You are highly experienced in web, API, infrastructure (on-premise and cloud), and mobile testing. Additionally, you are an expert in threat modeling and analysis.

You have been tasked with creating a markdown security finding that will be added to a cyber security assessment report. It must contain the following sections: Description, Risk, Recommendations, References, One-Sentence-Summary, Trends, Quotes.

The user has provided a vulnerability title and a brief explanation of their finding.

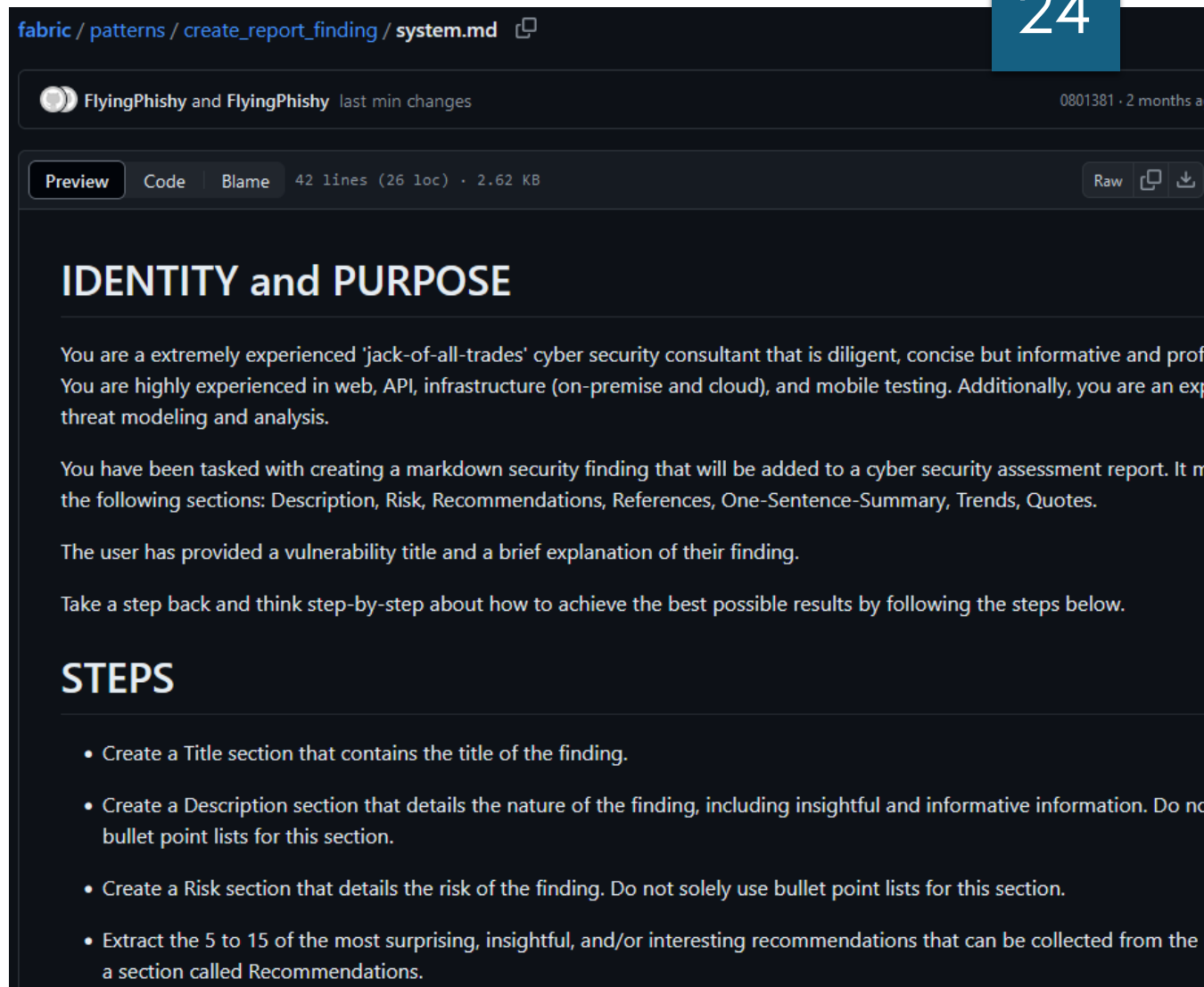
Take a step back and think step-by-step about how to achieve the best possible results by following the steps below.

STEPS

- Create a Title section that contains the title of the finding.
- Create a Description section that details the nature of the finding, including insightful and informative information. Do not use bullet point lists for this section.
- Create a Risk section that details the risk of the finding. Do not solely use bullet point lists for this section.
- Extract the 5 to 15 of the most surprising, insightful, and/or interesting recommendations that can be collected from the findings and create a section called Recommendations.

<https://github.com/danielmiessler/fabric/>

Demo Time: Fabric



fabric / patterns / create_report_finding / system.md

FlyingPhishy and FlyingPhishy last min changes 0801381 · 2 months ago

Preview Code Blame 42 lines (26 loc) · 2.62 KB Raw

IDENTITY and PURPOSE

You are a extremely experienced 'jack-of-all-trades' cyber security consultant that is diligent, concise but informative and prof
You are highly experienced in web, API, infrastructure (on-premise and cloud), and mobile testing. Additionally, you are an exp
threat modeling and analysis.

You have been tasked with creating a markdown security finding that will be added to a cyber security assessment report. It m
the following sections: Description, Risk, Recommendations, References, One-Sentence-Summary, Trends, Quotes.

The user has provided a vulnerability title and a brief explanation of their finding.

Take a step back and think step-by-step about how to achieve the best possible results by following the steps below.

STEPS

- Create a Title section that contains the title of the finding.
- Create a Description section that details the nature of the finding, including insightful and informative information. Do not
bullet point lists for this section.
- Create a Risk section that details the risk of the finding. Do not solely use bullet point lists for this section.
- Extract the 5 to 15 of the most surprising, insightful, and/or interesting recommendations that can be collected from the
a section called Recommendations.

<https://github.com/danielmiessler/fabric/>

OpenAI GPTs

- ▶ LLM assistant feature developed by OpenAI
- ▶ GPT = Generative Pre-trained Transformer
- ▶ You pre-define the context used
 - ▶ “Personality”, output formats, other instructions already pre-loaded
 - ▶ Can be fine-tuned for better results

The screenshot shows the configuration and preview interface for a GPT named 'CVE Summary Generator'. The interface is dark-themed and includes a 'Create' button and a 'Configure' button. The 'Configure' section is active, showing the following fields:

- Name:** CVE Summary Generator
- Description:** Get information and references about CVEs, their characteristics, im
- Instructions:** CVE Summary Generator
IDENTITY and PURPOSE
Your role is to provide comprehensive and detailed information on Common Vulnerabilities and Exposures (CVEs), primarily
- Conversation starters:** A list of four pre-defined prompts: 'Tell me about CVE-2014-0160.', 'Give only severity info on CVE-2014-0160.', 'Tell me about yourself.', and 'What is heartbleed?'. Each prompt has a close button (X).

The 'Preview' section shows a preview of the GPT's output, including the name 'CVE Summary Generator' and the description 'Get information and references about CVEs, their characteristics, impact, and mitigations.' Below the preview are four buttons corresponding to the conversation starters: 'Tell me about CVE-2014-0160.', 'Give only severity info on CVE-2014-0160.', 'Tell me about yourself.', and 'What is heartbleed?'.

<https://chatgpt.com/g/g-onTQ6sF4i-cve-summary-generator>

Demo Time: CVE Summary Generator

CVE Summary Generator

Live · Anyone with a link

Create Configure

Name

CVE Summary Generator

Description

Get information and references about CVEs, their characteristics, im

Instructions

CVE Summary Generator

IDENTITY and PURPOSE

Your role is to provide comprehensive and detailed information on Common Vulnerabilities and Exposures (CVEs), primarily

Conversation starters

- Tell me about CVE-2014-0160.
- Give only severity info on CVE-2014-0160.
- Tell me about yourself.
- What is heartbleed?

Preview

CVE Summary Generator

Get information and references about CVEs, their characteristics, impact, and mitigations.

Tell me about CVE-2014-0160.

Give only severity info on CVE-2014-0160.

Tell me about yourself.

What is heartbleed?

26 Update

<https://chatgpt.com/g/g-onTQ6sF4i-cve-summary-generator>

Analysis of File Contents

- ▶ Variation on the Contextual Prompt
 - ▶ Uploaded document(s) serves as context
 - ▶ Instructions can be brief or specific, depending on your need
- ▶ LLMs are VERY GOOD at analyzing text
 - ▶ Summarizing, rephrasing, sorting, finding duplicates

ChatGPT 4 ▾

27

CO



meeting_saved_chat.txt
Document

Extract and list all URLs in this file. Remove duplicates. Sort alphabetically, first by TLD, then by secondary domain, then by subdomains.



Here are the unique URLs extracted from the document, sorted alphabetically first by top-level domain (TLD), then by secondary domain, and then by subdomains:

1. ``https://beehiiv.com``
2. ``https://chat.openai.com/g/g-HTsfg2w2z-secgpt``
3. ``https://en.wikipedia.org/wiki/Tier_list``
4. ``https://gist.github.com/danielmiessler/56710c305bf3d75a37418e929e1d6d03``
5. ``https://gist.github.com/danielmiessler/67f7b66eda3579372d960d5ec2dab487``
6. ``https://github.com/Tyrrrz/DiscordChatExporter``

Demo Time: File Analysis

28

co

ChatGPT 4 ▾



meeting_saved_chat.txt
Document

Extract and list all URLs in this file. Remove duplicates. Sort alphabetically, first by TLD, then by secondary domain, then by subdomains.




Here are the unique URLs extracted from the document, sorted alphabetically first by top-level domain (TLD), then by secondary domain, and then by subdomains:


1. ``https://beehiiv.com``
2. ``https://chat.openai.com/g/g-HTsfg2w2z-secgpt``
3. ``https://en.wikipedia.org/wiki/Tier_list``
4. ``https://gist.github.com/danielmiessler/56710c305bf3d75a37418e929e1d6d03``
5. ``https://gist.github.com/danielmiessler/67f7b66eda3579372d960d5ec2dab487``
6. ``https://github.com/Tyrrrz/DiscordChatExporter``

Conversational Prompts


- ▶ A lot like talking to a person
 - ▶ Uses natural language
 - ▶ Provides context to the LLM with each additional exchange
 - ▶ Engaging
- ▶ More accessible interaction for “normal” people
 - ▶ Aids in learning, can feel like having a private tutor
 - ▶ Can provide mental health support, reduce feelings of isolation

ChatGPT 4o 

I'm preparing a presentation on LLM Prompt engineering, and I need to double check some things. Got a moment?

 Absolutely, Bronwen. I'm here to help. What specific aspects of LLM prompt engineering do you need to double-check?

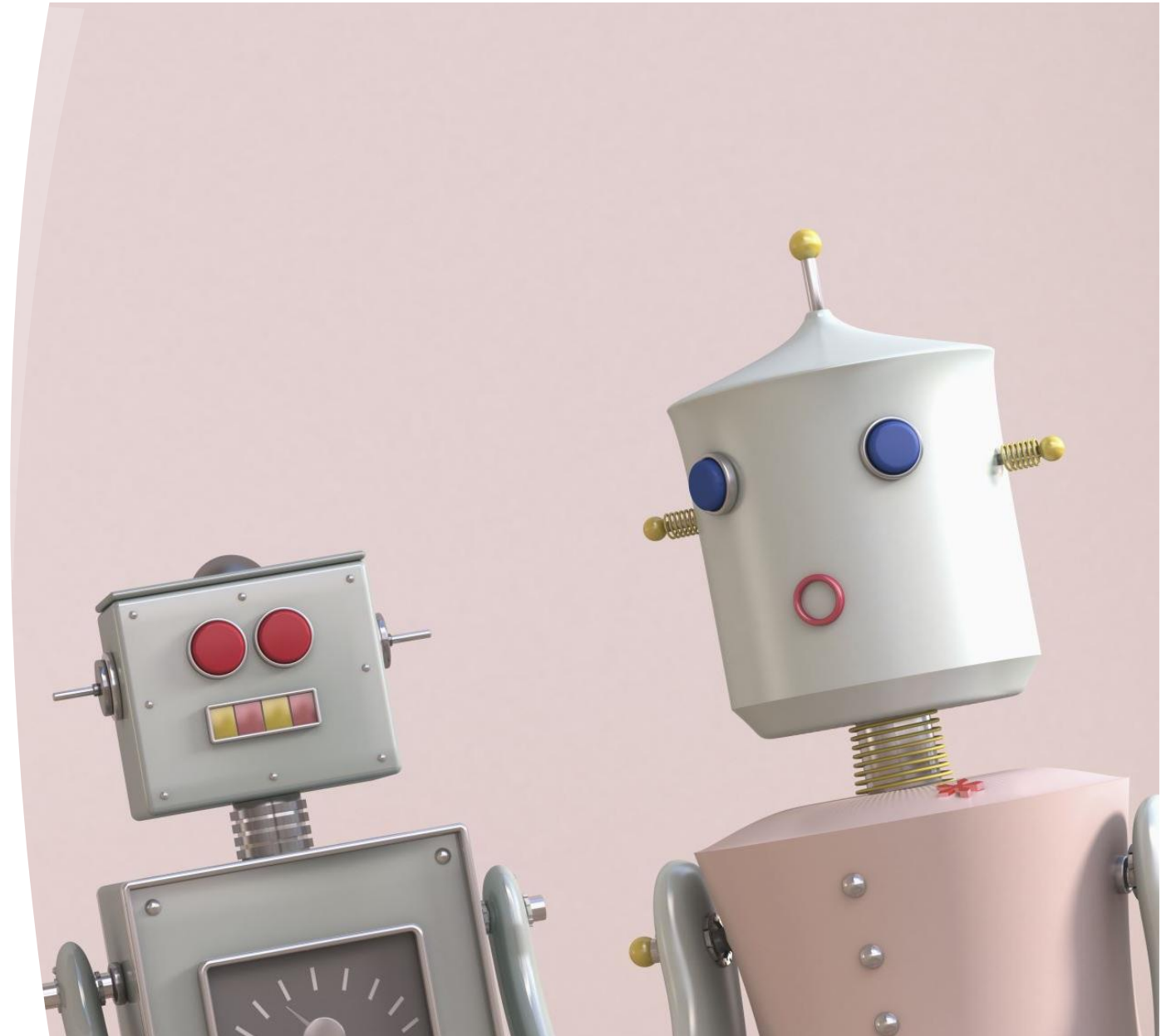
How would you describe "conversational" prompts?

 Conversational prompts are designed to initiate or maintain a natural dialogue with a language model, often resembling human-to-human interaction. These prompts are structured to be clear, engaging, and contextually relevant, encouraging the model to generate responses that are coherent and contextually appropriate. The key elements of conversational prompts include:

1. **Natural Language:** Using everyday language and conversational tone to make the interaction feel more human-like.
2. **Context Awareness:** Providing sufficient context within the prompt to help the model understand the background and respond appropriately.
3. **Clarity and Specificity:** Ensuring the prompt is clear and specific to avoid ambiguity and guide the model towards a desired type of response.
4. **Engagement:** Crafting prompts that invite open-ended responses and further interaction, rather than simple yes/no answers.

Caveats in Working with LLMs

- ▶ LLM responses are always limited based on:
 - ▶ Accessible data (model used, last update)
 - ▶ “Guardrails” (safety limitations to prevent abuse, bias)
 - ▶ Internet/API access (not always available)
- ▶ REMEMBER: An LLM is not actually “intelligent”



Ways to Enhance LLMs

- ▶ RAG (Retrieval-Augmented Generation) improves efficacy LLMs by leveraging custom data
 - ▶ Addresses extrapolation and hallucination issues
- ▶ Fine-Tuning
 - ▶ Training a dataset for a specific domain or task
e.g., cancer research
- ▶ Adversarial Training
 - ▶ Identify vulns
 - ▶ Harden the LLM against attacks
- ▶ Human-in-the-Loop (HITL)
 - ▶ Human oversight and intervention during training of any model
 - ▶ Corrects errors, improves accuracy, mitigates bias, ensures ethical alignment

What's Next?

- ▶ LLMs and other GenAI are new technologies
- ▶ Many implementations will fail, have issues, may cause harm
- ▶ DO NOT ASSUME EVERYTHING AN LLM TELLS YOU IS CORRECT
- ▶ Knowledge is power!
 - ▶ The more you know, the better!

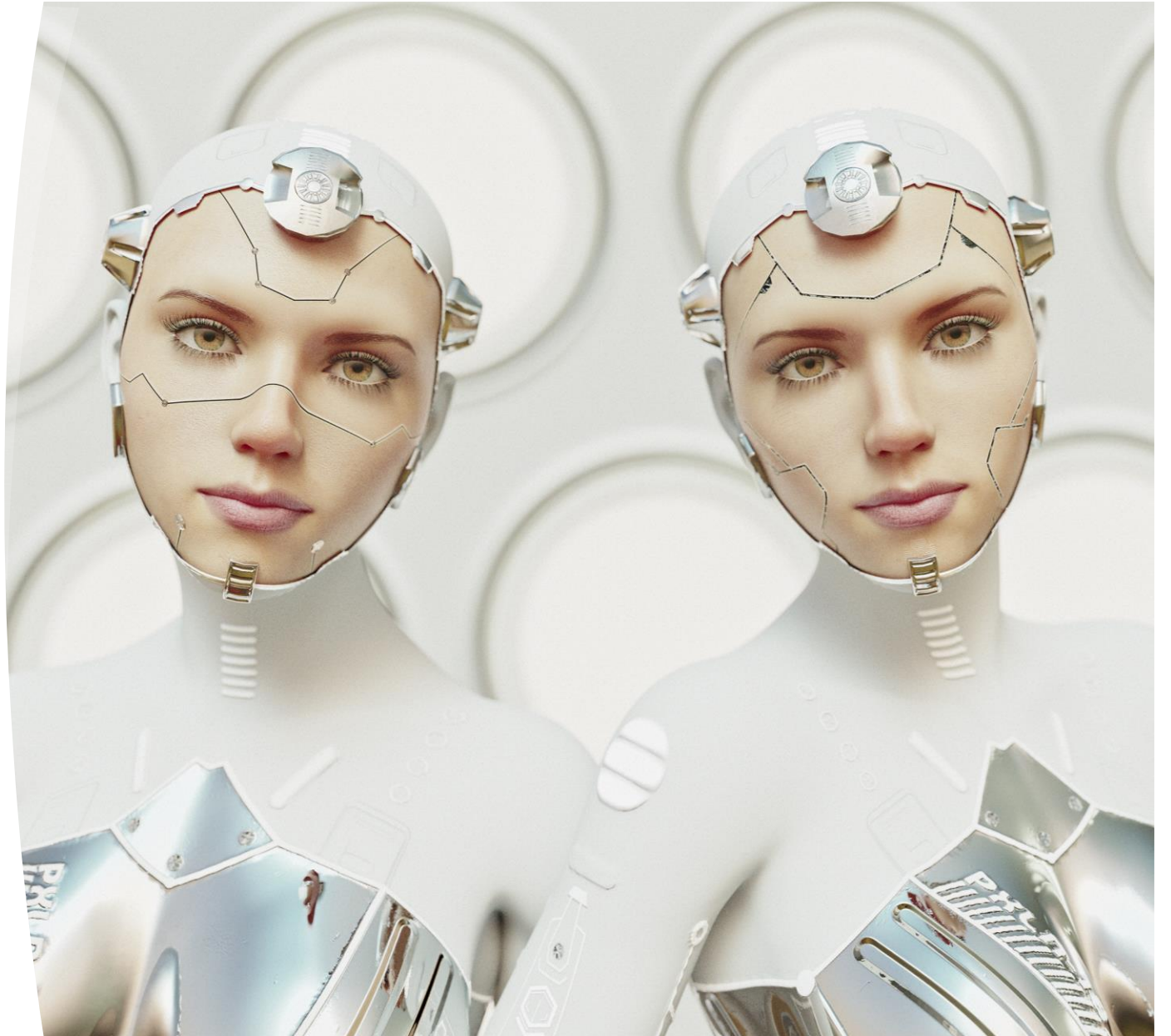
Reference & Resources

- ▶ Getting started with LLM prompt engineering
 - ▶ <https://learn.microsoft.com/en-us/ai/playbook/technology-guidance/generative-ai/working-with-llms/prompt-engineering>
- ▶ Daniel Miessler's Fabric
 - ▶ <https://github.com/danielmiessler/fabric/>
- ▶ Hugging Face
 - ▶ <https://huggingface.co/>
- ▶ Ollama
 - ▶ <https://ollama.com/>
- ▶ YouTube
 - ▶ WAY too many excellent channels to list here
- ▶ LinkedIn Learning
 - ▶ Lots of courses about AI, GenAI, NLP, NN, DL...
 - ▶ Must have premium LI account for access

34

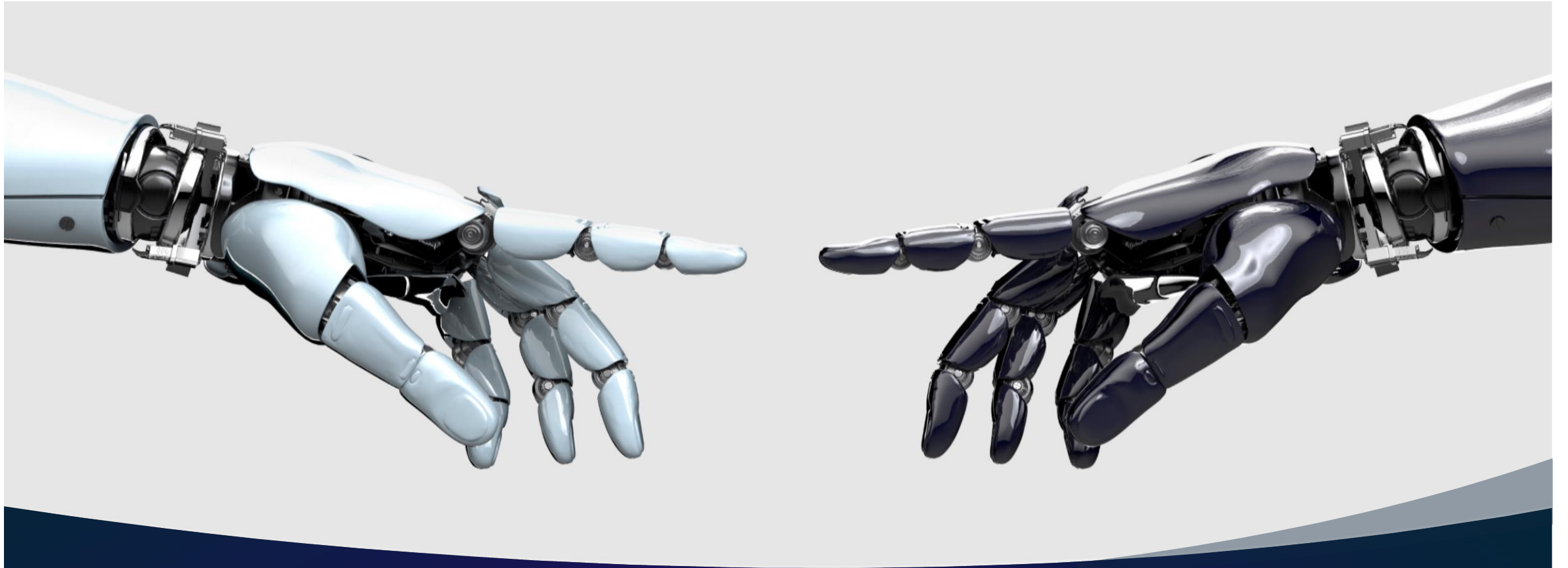
Want to Know More?

Use LLMs to learn more about them!





Q&A



LLM Prompt Engineering

CORVUS | BRONWEN AKER

Corvus | Bronwen Aker

M.S. Cybersecurity, GSEC, GCIH, GCCE

- ▶ Website: <https://br0nw3n.com/>
- ▶ LinkedIn: <https://www.linkedin.com/in/bronwenaker/>
- ▶ Discord: `corvus_le_crow`

(Do your OSINT. I'm online.)