RAG(ing) The Android Malware

Conceptualizing AI/ML for benefits of Reverse Engineer

Shivang Desai

“Everything published, talked and/or discussed in this conference is solely based on my personal point of view, and does not represent my current, or past employers.”

Who Am I !

- Shivang Desai (Shiv)
- Sr. Security Researcher @ Microsoft
- 10+ years of experience
- Mobile Security and Malware Reverse Engineering.
- Loves blogging.
    - Blogs referenced in international media like The Forbes, The Register, Bleeping Computer, The Fox News, Mobile Hackers Handbook, MITRE etc

linkedin.com/in/desaishivang/

x.com/5h1vang

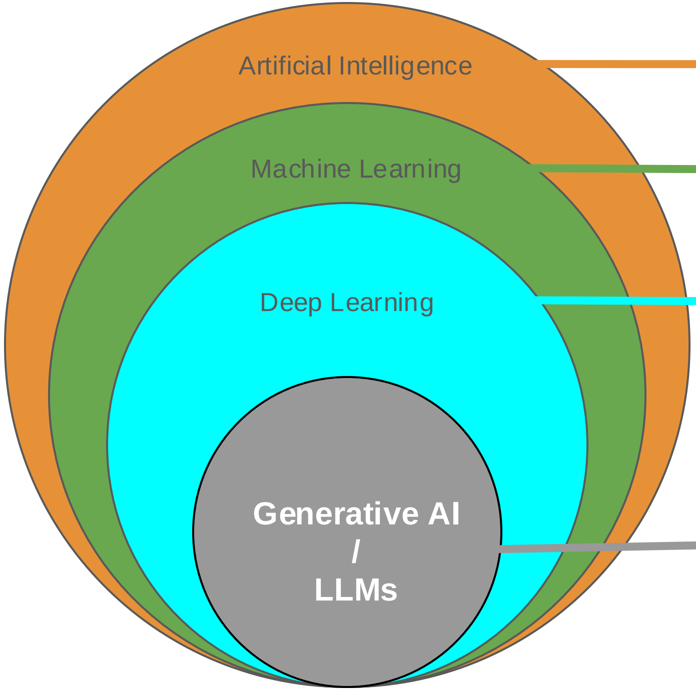
Next 40 mins !

- What’s That !
- Decoding the jargons !
- What are LLMs ?
- Sneak-Peak into RAG !
- Combining Android Malware + RAG
- Demos
  - Local Setup on laptop
  - Alternative of local setup
- Q&A

What’s That ?

- AI Chatbot developed by *OpenAI*
- Understands & Generates human-like text, based on the input it receives
- Based on GPT (Generative Pre-Trained Transformer) architecture.

Decoding the jargons.



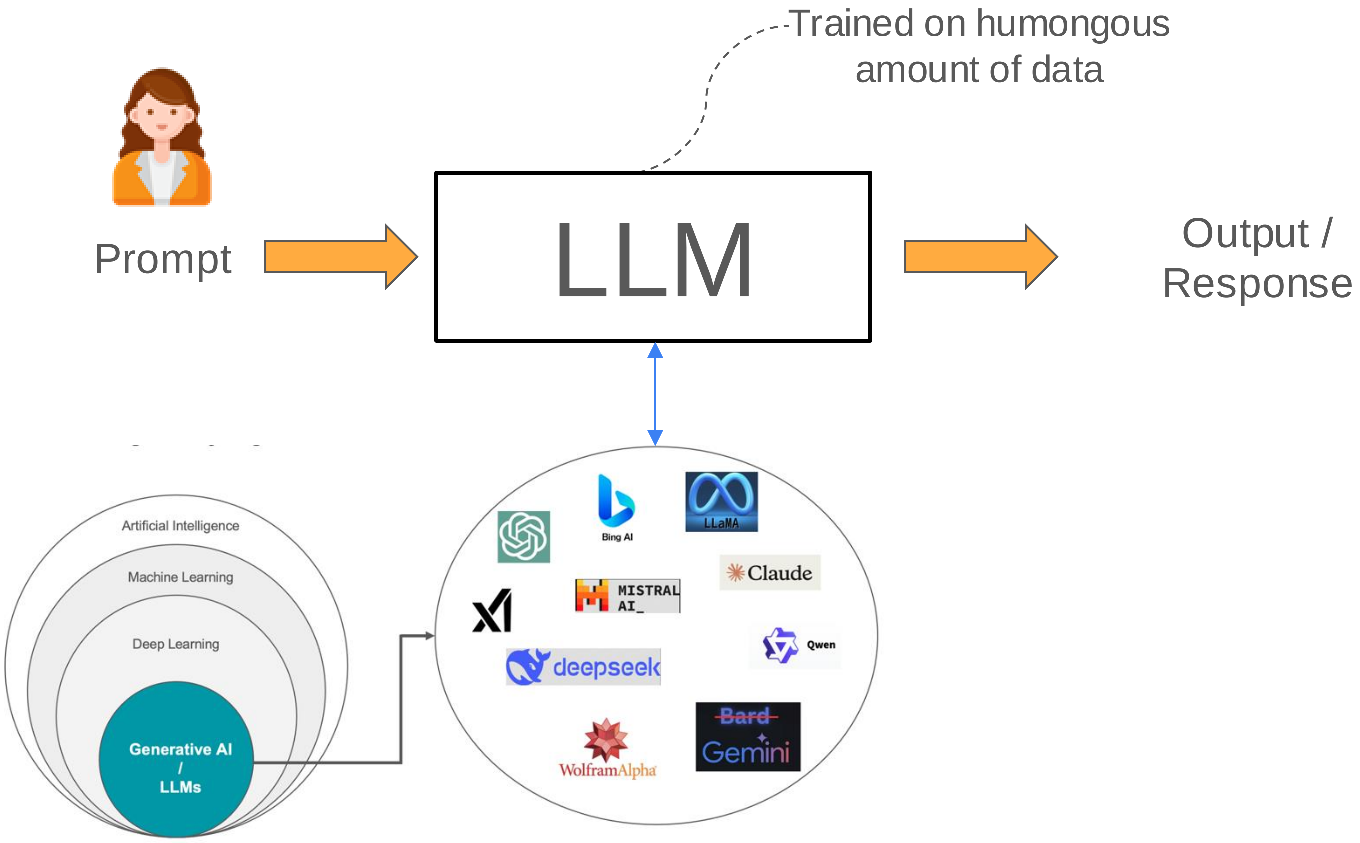
Involves techniques that equip computers to emulate human behaviour and intelligence.

Subset of Artificial Intelligence, which uses advanced algorithms to detect patterns in large data sets, allowing machines to learn and adapt.

Subset of Machine Learning, which uses neural networks for in-depth data processing and analytical tasks.

Subset of Deep Learning, that generates content like text, images, or code based on provided inputs. Trained on vast(Large) data sets, these models detect patterns and create outputs without explicit instructions.

ChatBots and LLMs !



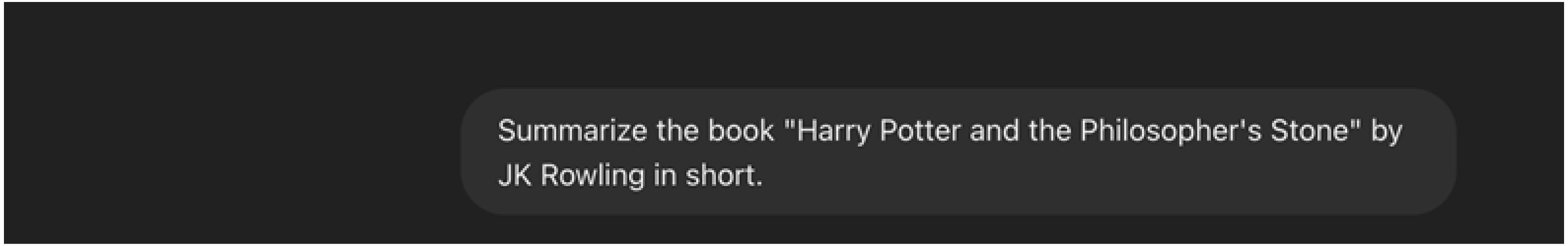
Trained on humongous amount of data

Prompt

LLM

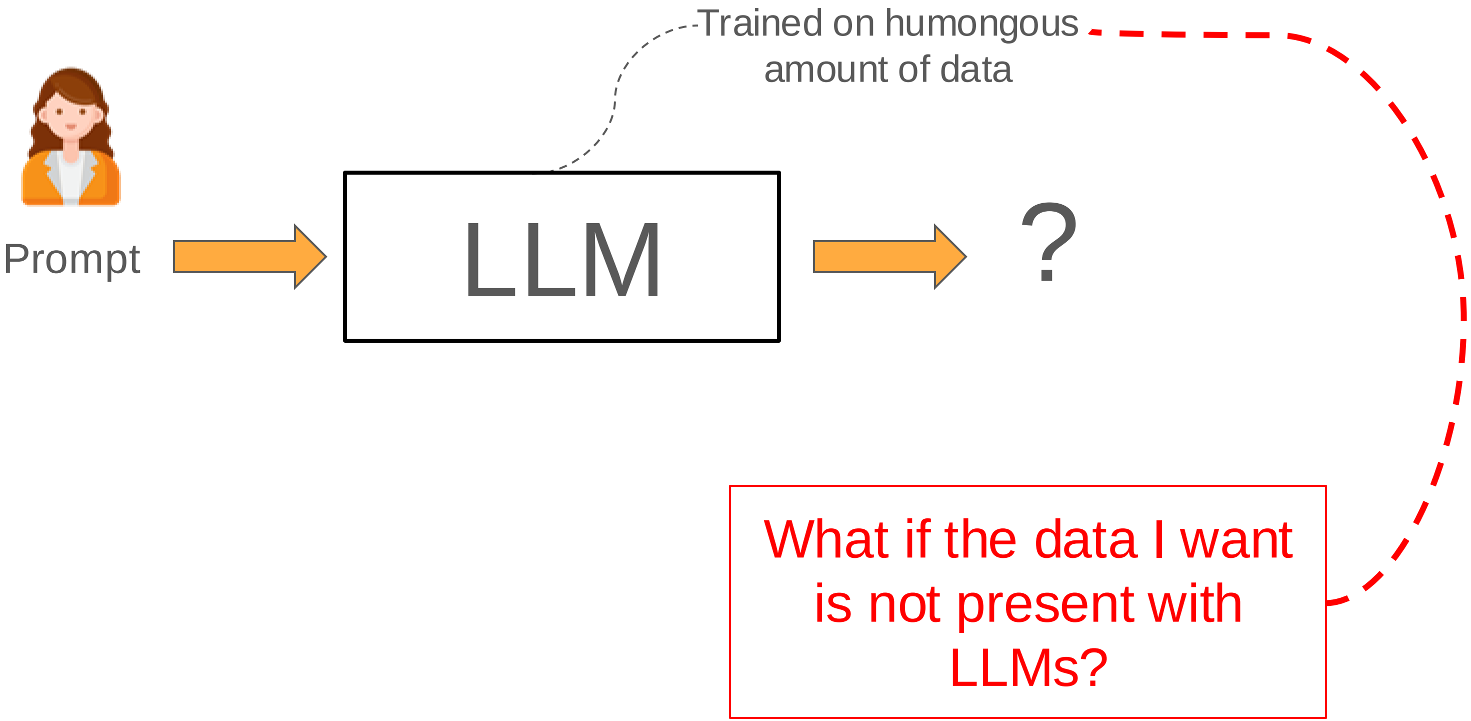
Output / Response

ChatBots !



Summarize the book "Harry Potter and the Philosopher's Stone" by JK Rowling in short.

ChatBots (Limitations)



Trained on humongous amount of data

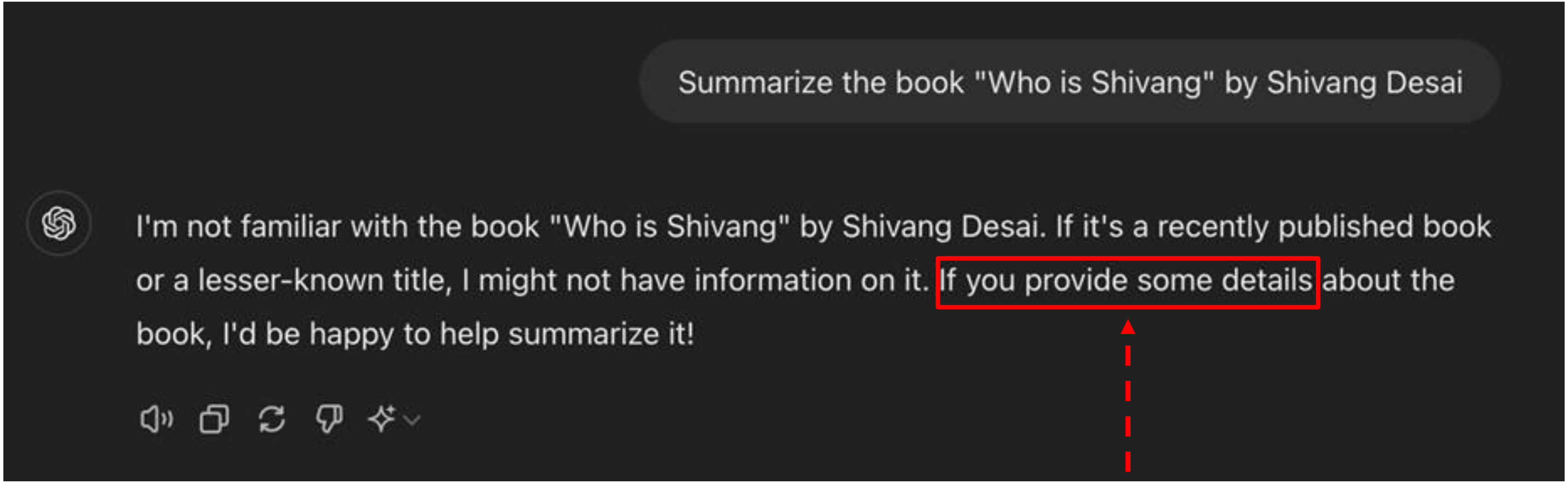
Prompt

LLM

?

What if the data I want is not present with LLMs?

ChatBots (Limitations)

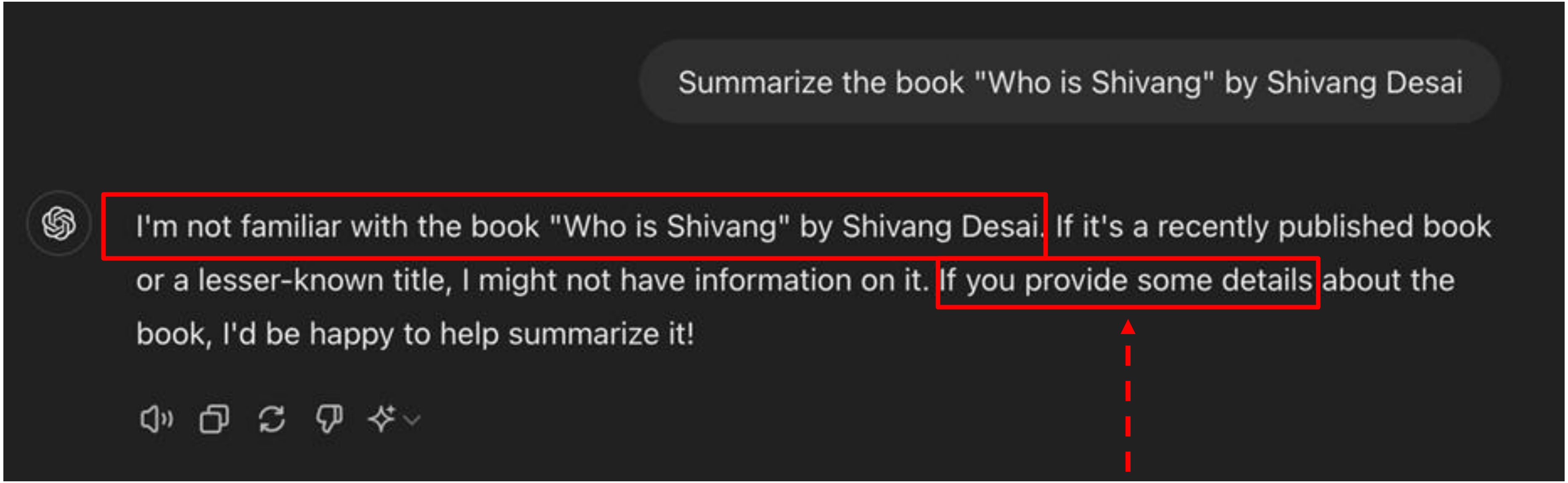


Summarize the book "Who is Shivang" by Shivang Desai

I'm not familiar with the book "Who is Shivang" by Shivang Desai. If it's a recently published book or a lesser-known title, I might not have information on it. If you provide some details about the book, I'd be happy to help summarize it!

External Data/Knowledge Base

ChatBots (Limitations)

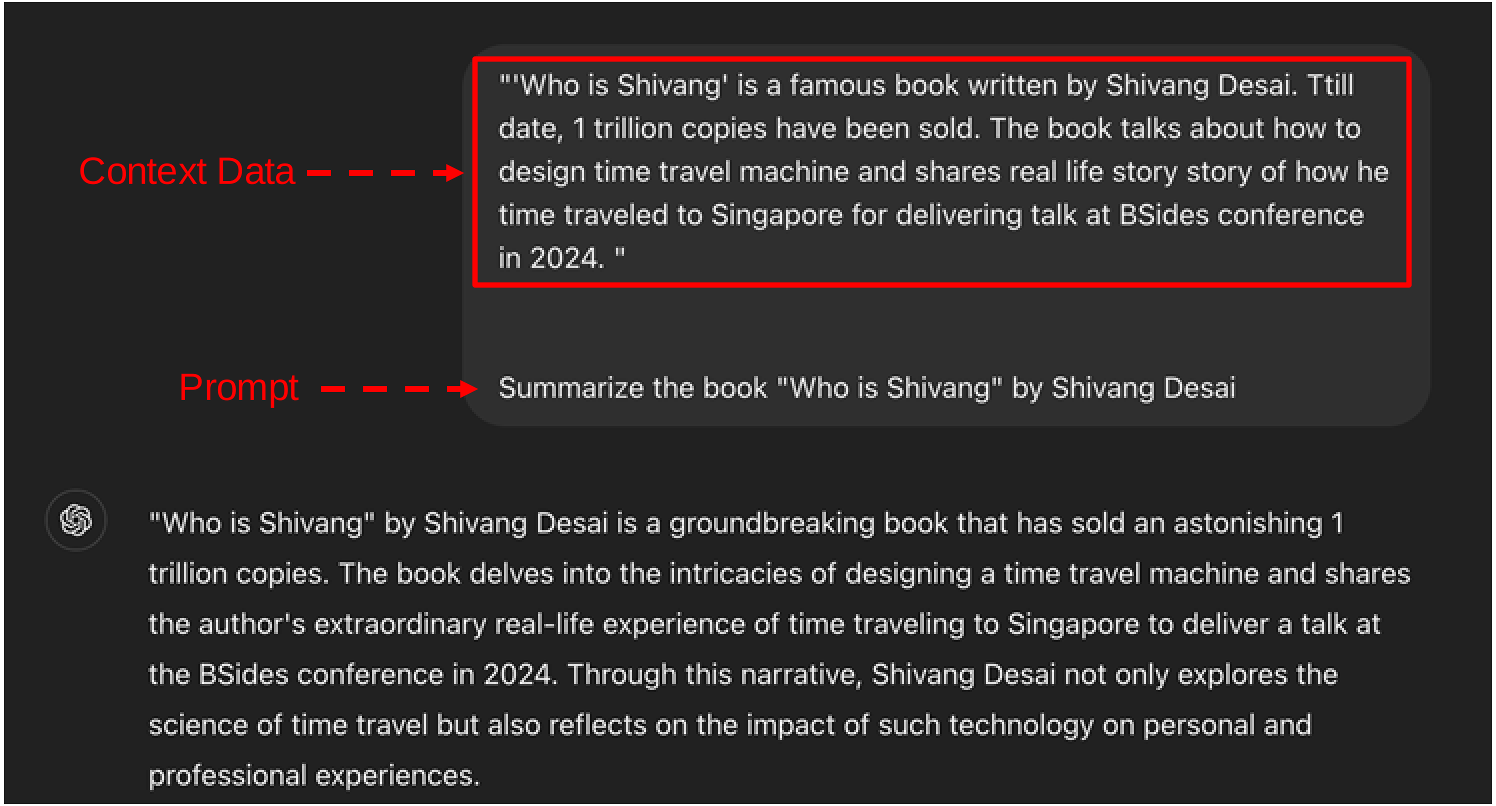


Summarize the book "Who is Shivang" by Shivang Desai

I'm not familiar with the book "Who is Shivang" by Shivang Desai. If it's a recently published book or a lesser-known title, I might not have information on it. If you provide some details about the book, I'd be happy to help summarize it!

External Data/Knowledge Base

Providing the Knowledge (Context)



Context Data

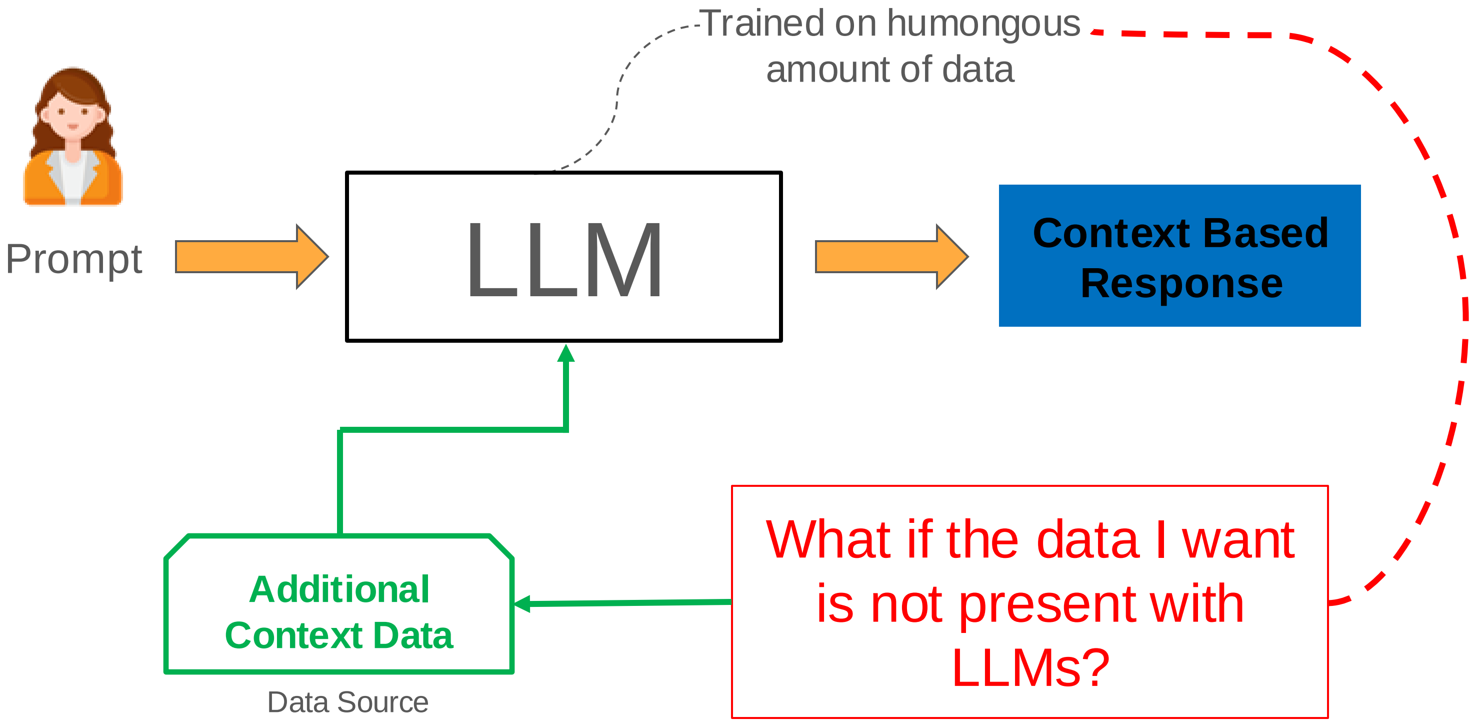
"'Who is Shivang' is a famous book written by Shivang Desai. Ttill date, 1 trillion copies have been sold. The book talks about how to design time travel machine and shares real life story story of how he time traveled to Singapore for delivering talk at BSides conference in 2024. "

Prompt

Summarize the book "Who is Shivang" by Shivang Desai

"Who is Shivang" by Shivang Desai is a groundbreaking book that has sold an astonishing 1 trillion copies. The book delves into the intricacies of designing a time travel machine and shares the author's extraordinary real-life experience of time traveling to Singapore to deliver a talk at the BSides conference in 2024. Through this narrative, Shivang Desai not only explores the science of time travel but also reflects on the impact of such technology on personal and professional experiences.

ChatBots (Overcoming The Limitations)



Trained on humongous amount of data

Prompt

LLM

Context Based Response

Additional Context Data

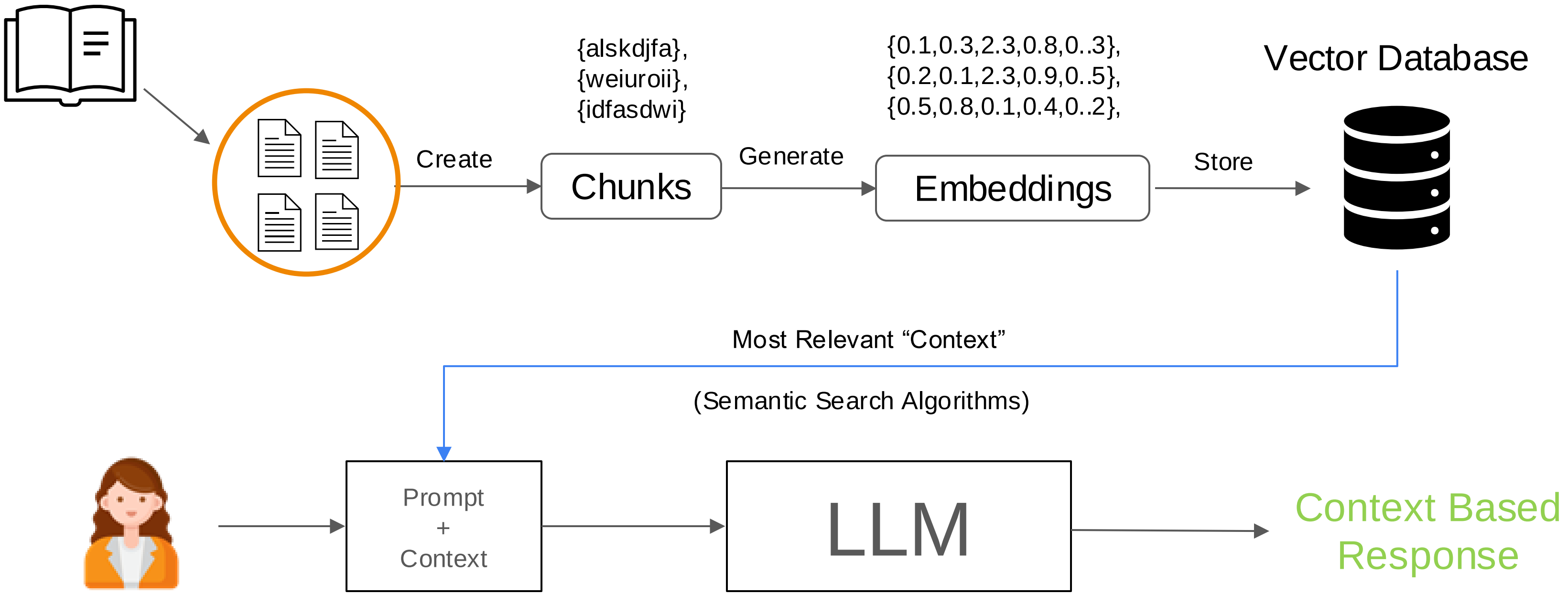
Data Source

What if the data I want is not present with LLMs?

Retrieval Augmented Generation (RAG)

“ RAG is an AI framework that retrieves data from an external knowledge base and helps pre-trained LLMs to generate accurate, up-to-date information.”

How to achieve RAG ?



{alskdjfa},
{weiuroii},
{idfasdwi}

{0.1,0.3,2.3,0.8,0..3},
{0.2,0.1,2.3,0.9,0..5},
{0.5,0.8,0.1,0.4,0..2},

Vector Database

Create

Chunks

Generate

Embeddings

Store

Most Relevant “Context”

(Semantic Search Algorithms)

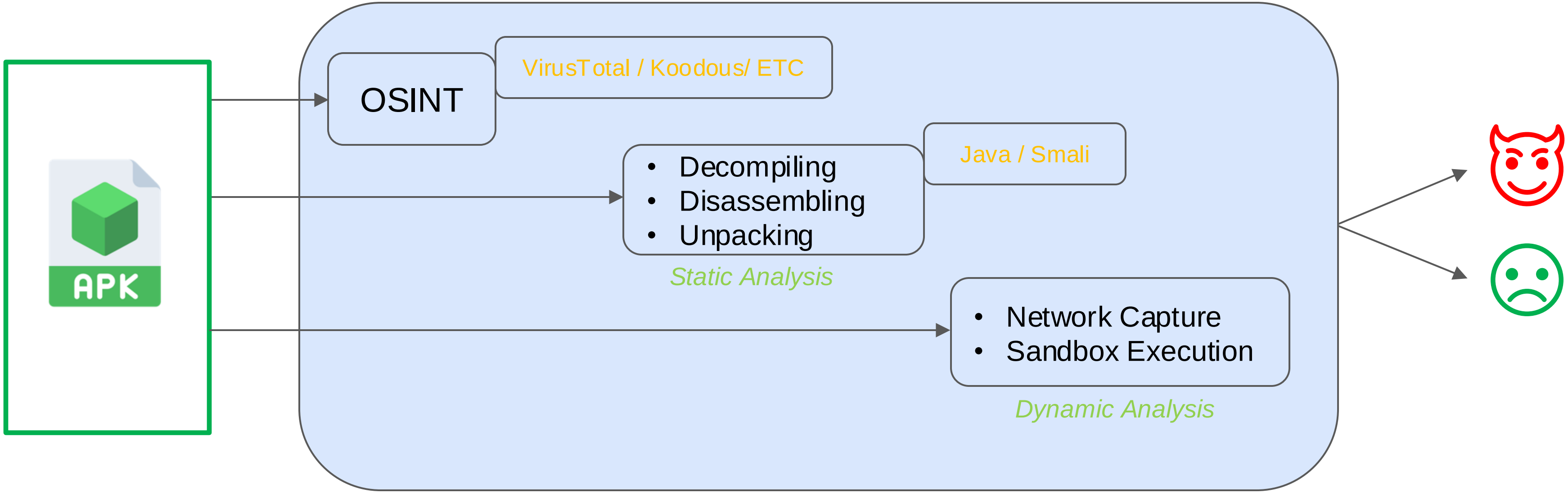
Prompt
+
Context

LLM

Context Based Response

RAG vs Android Malware ?

Traditional Android Malware Analysis



OSINT

VirusTotal / Koodous/ ETC

- Decompiling
- Disassembling
- Unpacking

Java / Smali

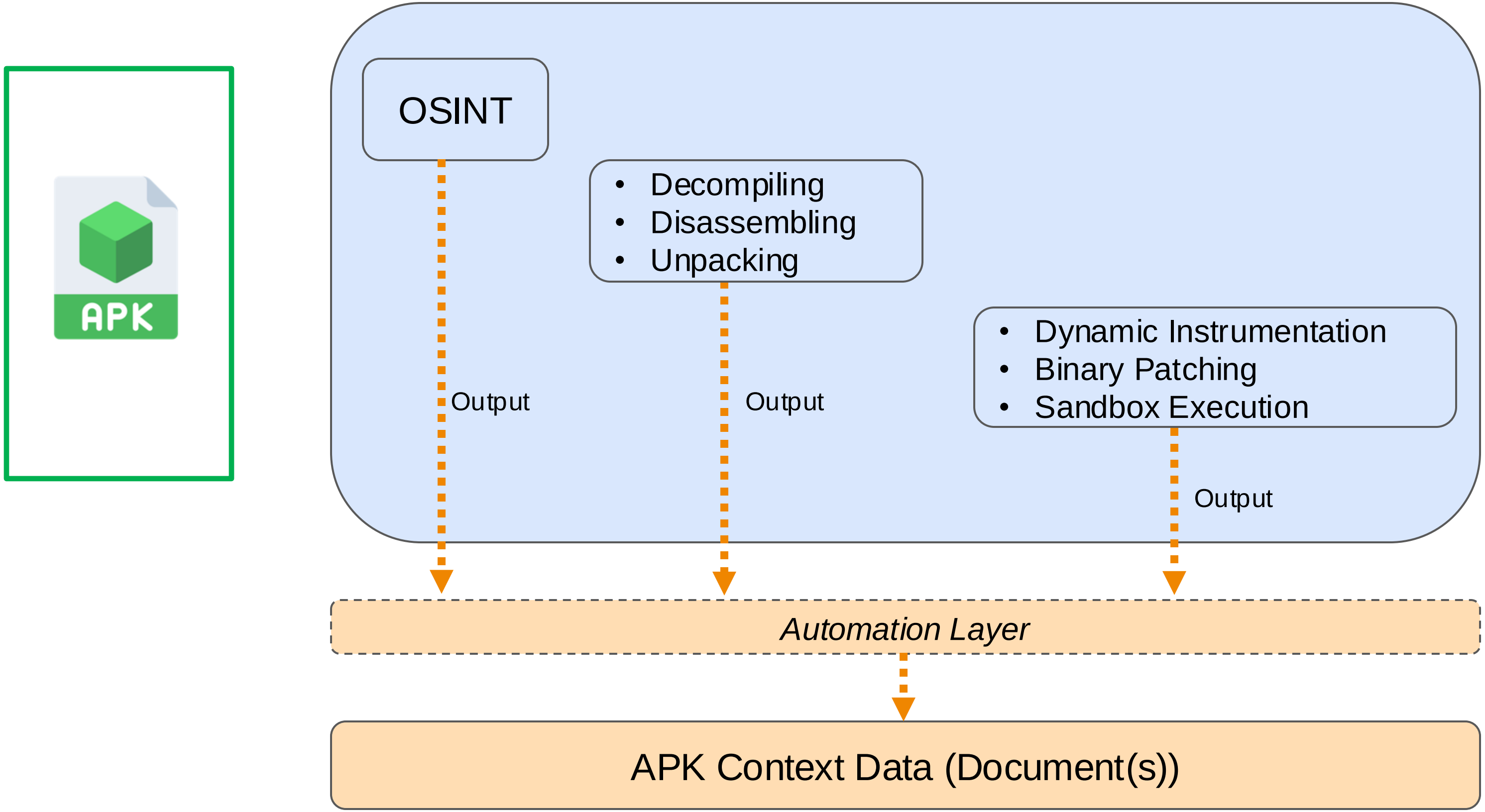
Static Analysis

- Network Capture
- Sandbox Execution

Dynamic Analysis

Malware Analysis Env.

Preps for RAG



OSINT

- Decompiling
- Disassembling
- Unpacking

- Dynamic Instrumentation
- Binary Patching
- Sandbox Execution

Output

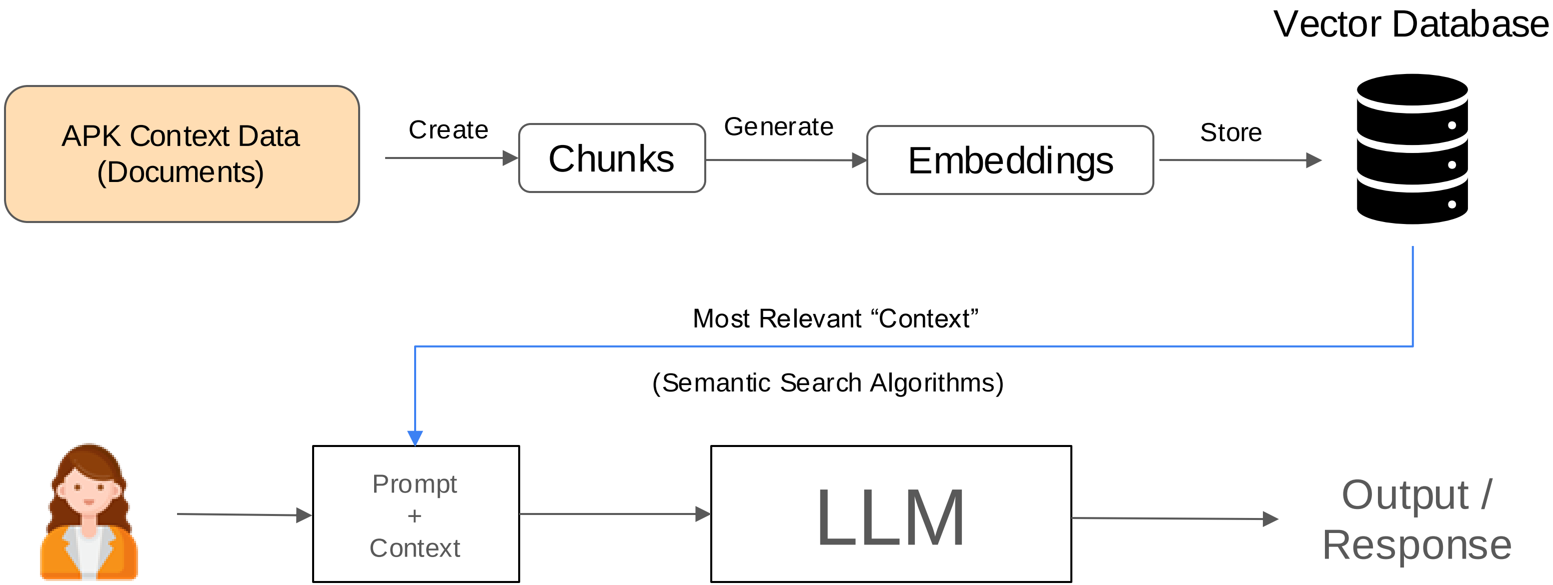
Output

Output

Automation Layer

APK Context Data (Document(s))

RAG(ging) Android Malware



Vector Database

APK Context Data
(Documents)

Create

Chunks

Generate

Embeddings

Store

Most Relevant “Context”

(Semantic Search Algorithms)

Prompt
+
Context

LLM

Output /
Response

Demo

1. Running LLMs on Macbook + demo app.
2. Alternative Setup with actual Android Spyware.

?

Demo 1 - Running LLM locally + demo app.

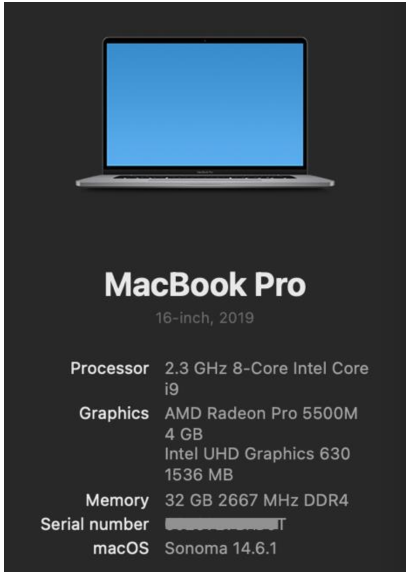
(Video Demo)

Source Code will be published on Github

github.com/shivang1989

1. Zer0 Cost Setup

ollama



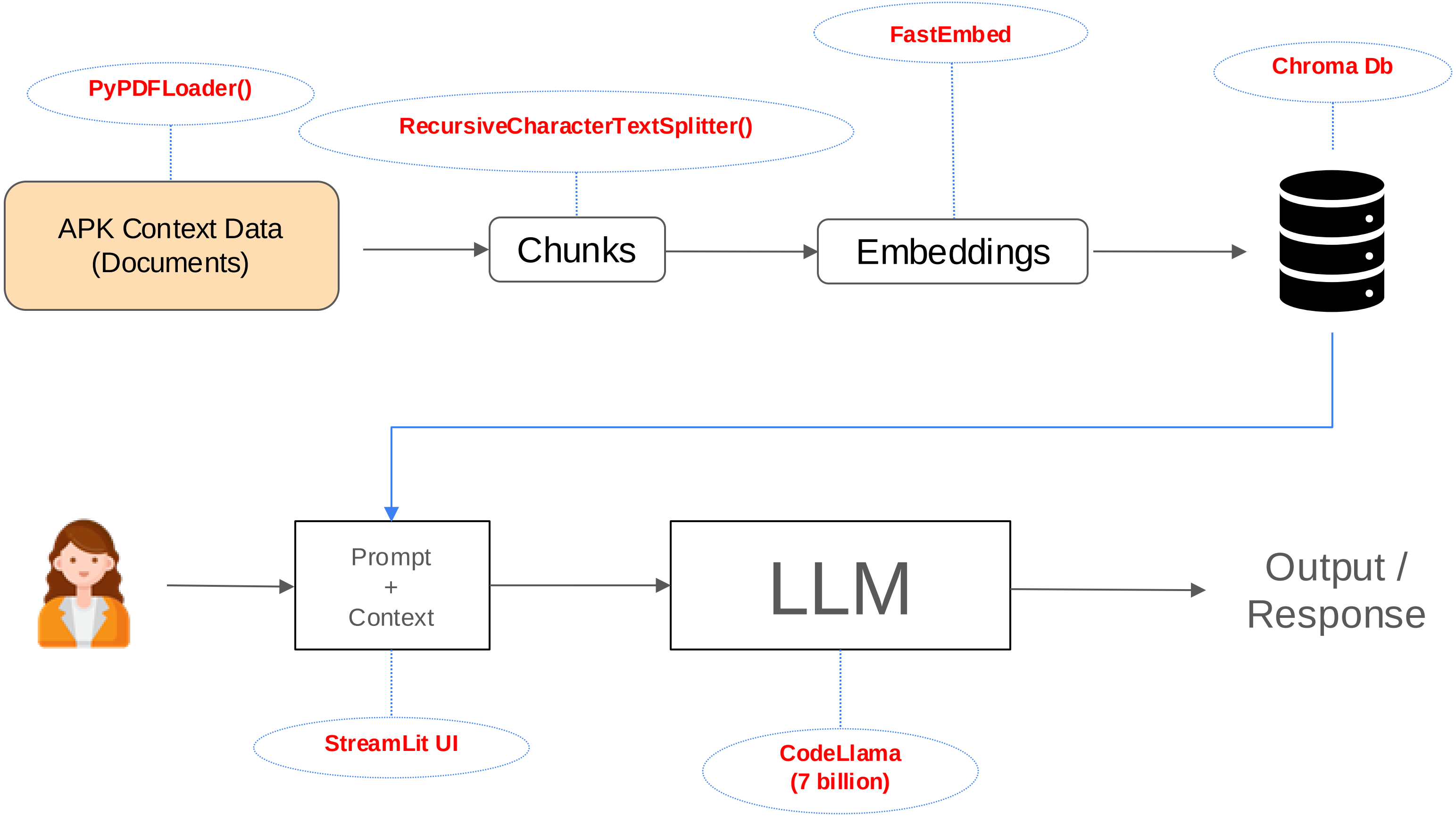
MacBook Pro

16-inch, 2019

Processor	2.3 GHz 8-Core Intel Core i9
Graphics	AMD Radeon Pro 5500M 4 GB Intel UHD Graphics 630 1536 MB
Memory	32 GB 2667 MHz DDR4
Serial number	[redacted]T
macOS	Sonoma 14.6.1

🦜🔗 LangChain

RAG Components



PyPDFLoader()

RecursiveCharacterTextSplitter()

FastEmbed

Chroma Db

APK Context Data
(Documents)

Chunks

Embeddings

Prompt
+
Context

LLM

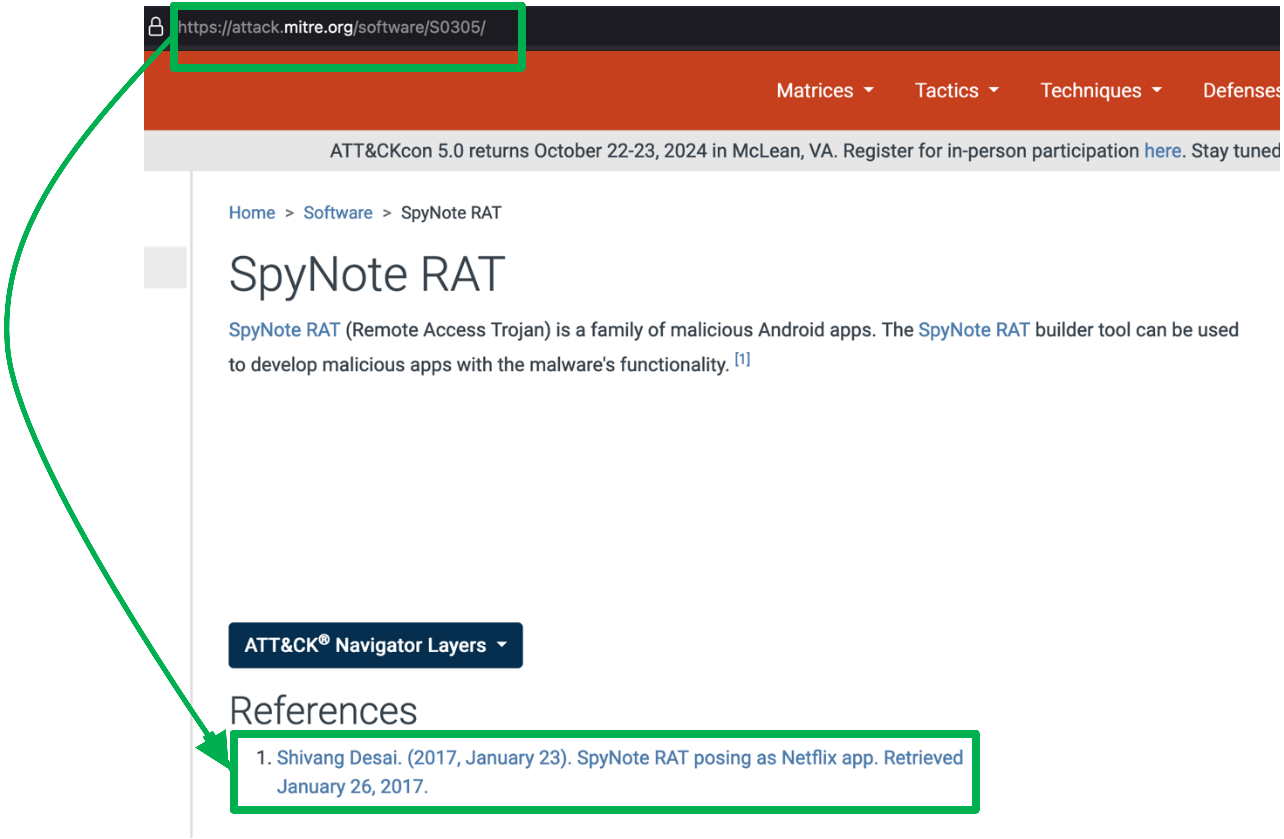
Output /
Response

StreamLit UI

CodeLlama
(7 billion)

Demo 2 - RAG with ChatGPT.

Demo 2 - RAG with ChatGPT.



ATT&CKcon 5.0 returns October 22-23, 2024 in McLean, VA. Register for in-person participation here. Stay tuned

Home > Software > SpyNote RAT

SpyNote RAT

SpyNote RAT (Remote Access Trojan) is a family of malicious Android apps. The SpyNote RAT builder tool can be used to develop malicious apps with the malware's functionality. [1]

References

1. Shivang Desai. (2017, January 23). SpyNote RAT posing as Netflix app. Retrieved January 26, 2017.

Limitations

- Local LLM setup is resource dependent
- Open Source LLMs are community/org dependent
- Obfuscation and Packed Malware
- Various types of context data
    - Smali / Java / Kotlin
    - Network Capture - .pcap

Thank You for being Fantastic Audience !

Q & A