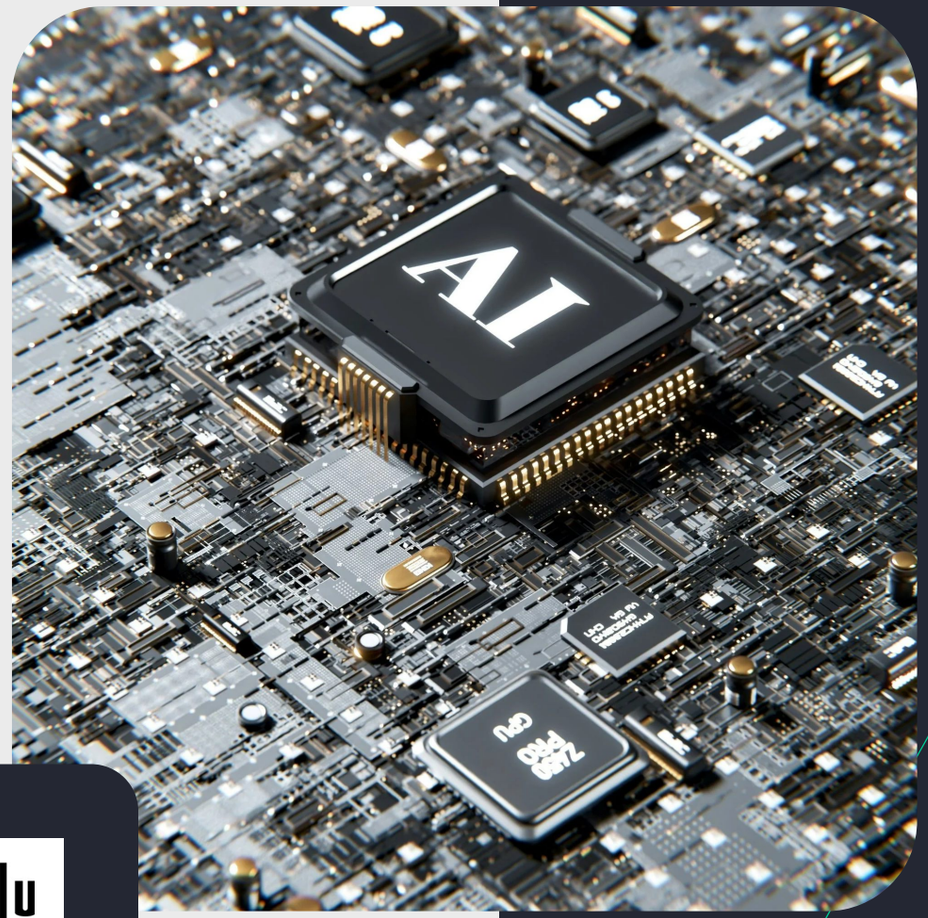
Simplifying Security with Local AI Agents

By Abhinandan Khurana

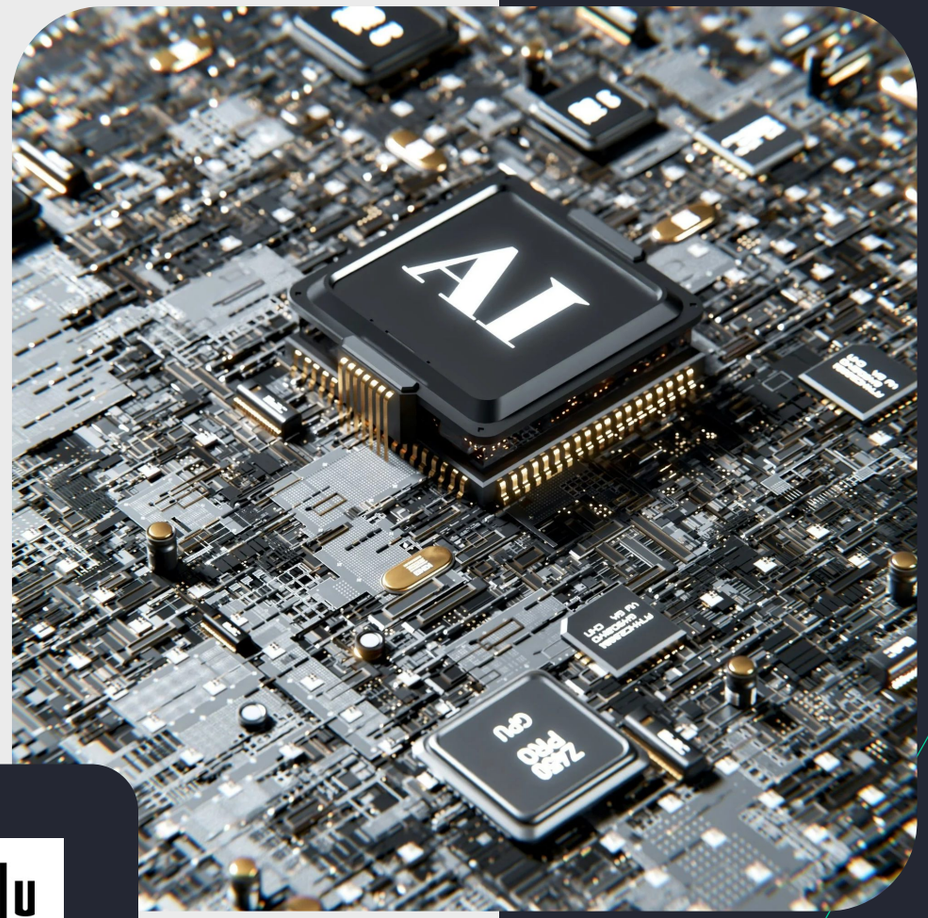
Agenda
TL;DR

- Help you understand and streamline your own research
- When, Why and How?
- Problems & possible Solutions



Disclaimer
What not to expect?

- Code Discussion
- "Perfect", no flaw PoC Demo
- this session is enough!



Audience POV?

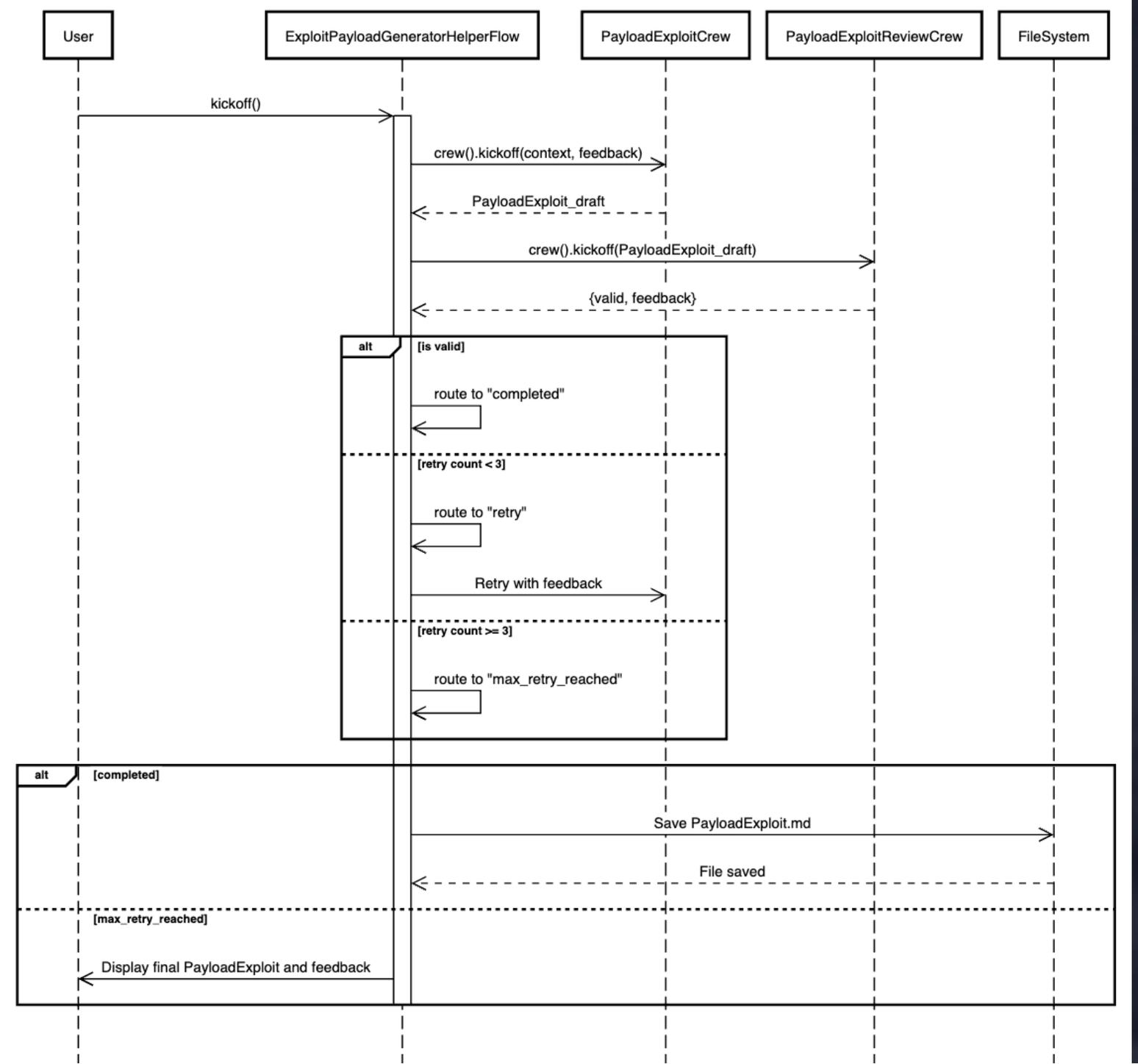
Problems?

Challenges?

Ideas?

Demo 1

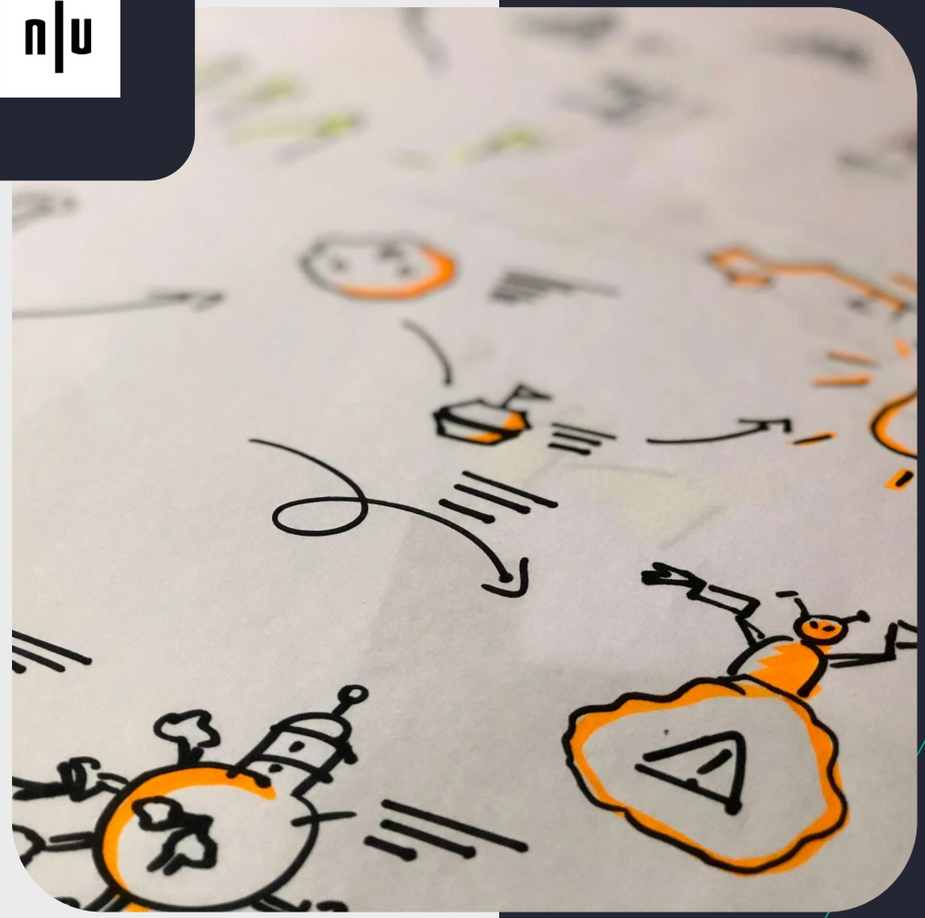
Exploit Payload Generator Agent



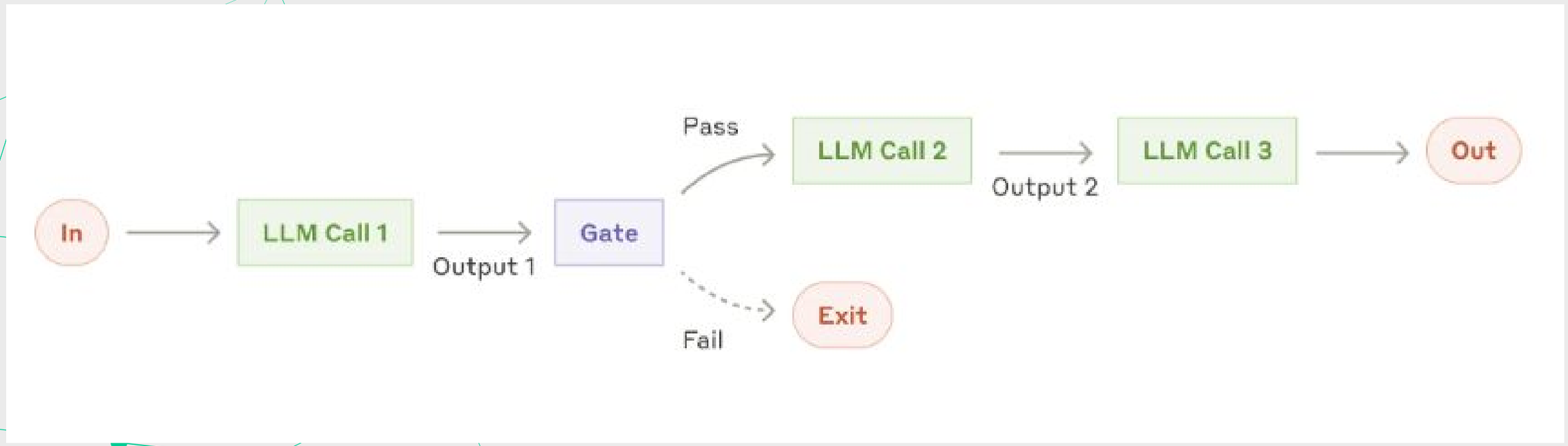
How does it work?

Introduction to AI Agents

1. Prompt Chaining
2. Routing
3. Parallelization
4. Orchestrator-Workers
5. Evaluator-Optimizer



Prompt Chaining

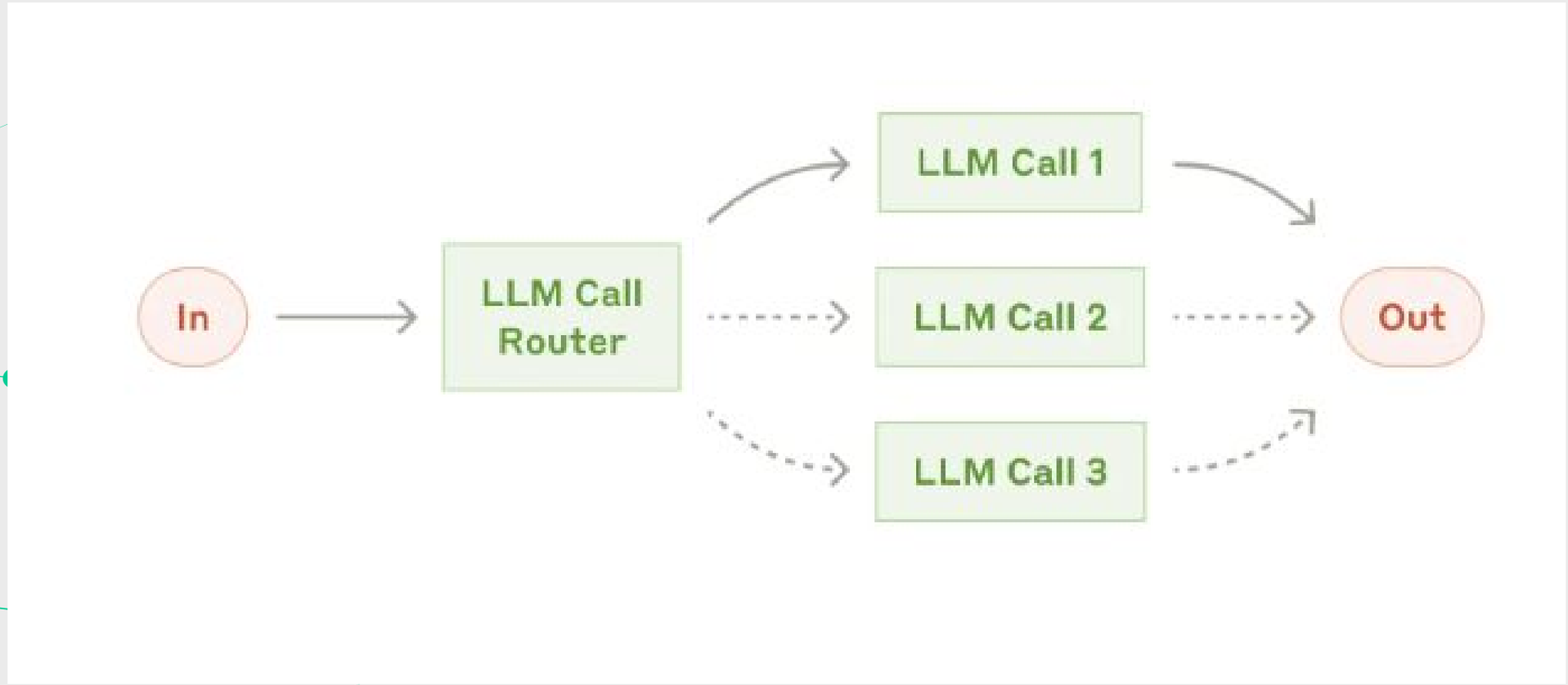


Workflow: Chaining

accuracy over latency

Example: document optimization

Routing

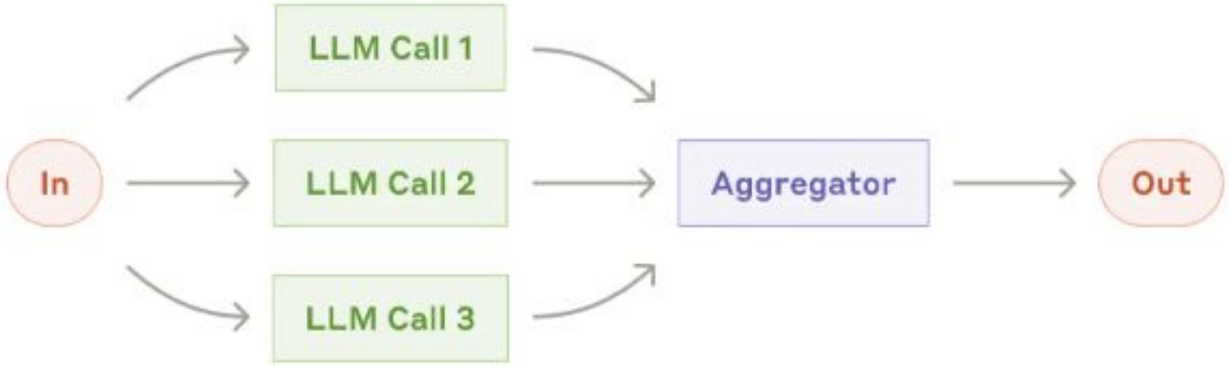


Workflow: Routing

complex to specific

Example: divide to efficient llm tasks

Parallelization



Workflow: parallel

Sectioning & Voting

Example: adversary network intrustion

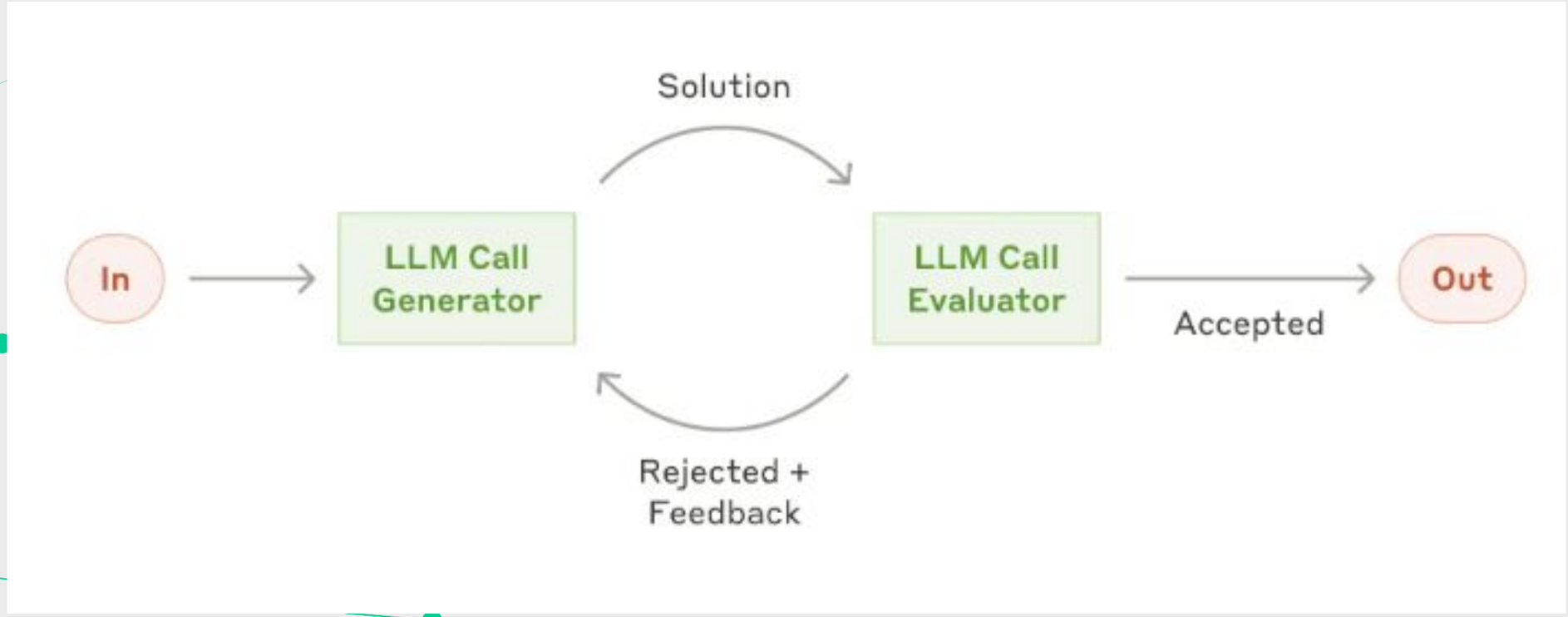
Orchestrator-Workers

Workflow: master-slave

dynamic process

Example: investigate and patch SCA issues

Evaluator-Optimizer



FeedBack Loop

iteration

Example: Demo 1

What local LLMs to use for cyber security usecase?

Hermes3

whiterabbitneo

deepseek

dolphin mixtral

qwen2.5

additional

https://llm.extractum.io/list/?uncensored

Search

Best Uncensored LLMs. Top-Ranked Uncensored Large Language Models. *

Was this list helpful?

Model Size: 0K – 999B

Model VRAM: 0 – 768

Columns ▼ | Quick Filters ▼ | × Reset Filters & Sorting

Model Name	Maintainer	Size	Score	VRAM (GB)	Quantized	License
Oxy 1 Small	oxyapi	14B	0.43	29.7	•	apache-2.0
EVA Qwen2.5 72B V0.2	EVA-UNIT-01	72B	0.41	146	•	other
Dobby Mini Unhinged Llama 3.1 8B	SentientAGI	8B	0.41	16.1	•	llama3.1
Hermes 3 Llama 3.2 3B	NousResearch	3B	0.39	6.5	•	llama3
Hermes 3 Llama 3.1 8B	NousResearch	8B	0.39	16.1	•	llama3
Violet Twilight V0.2	Epiculous	12B	0.38	24.5	•	apache-2.0
Wayfarer 12B	LatitudeGames	12B	0.38	24.5	•	apache-2.0
Dobby Mini Leashed Llama 3.1 8B	SentientAGI	8B	0.38	16.1	•	llama3.1
Hermes 3 Llama 3.1 70B	NousResearch	70B	0.37	141.9	•	llama3

What tech stack to use for building agents?

crewAI

Langchain

llamaindex

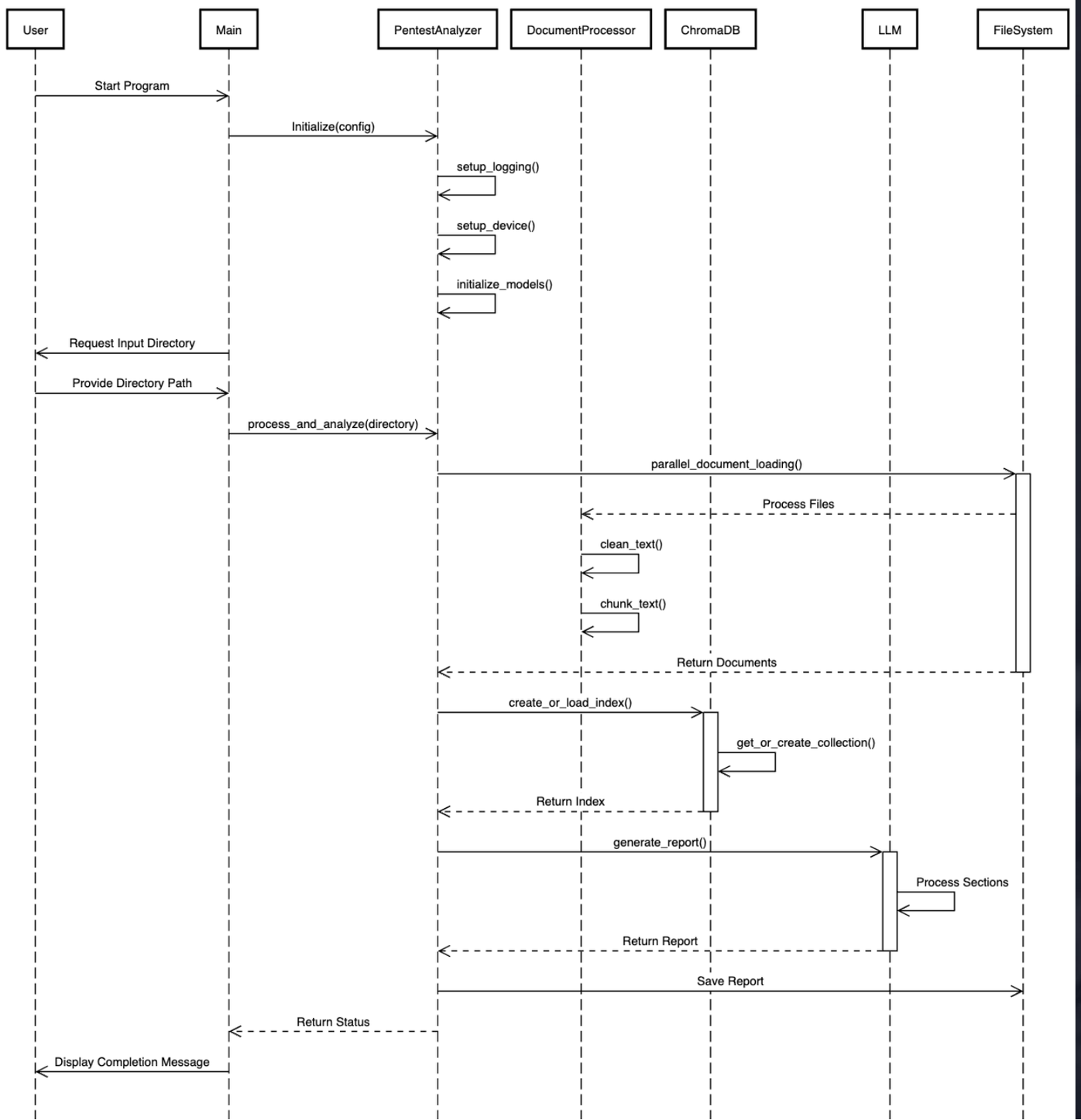
smolagents

langflow (no code)

n8n (no code)

Demo 2

RAG based Pentest report Generator Agent

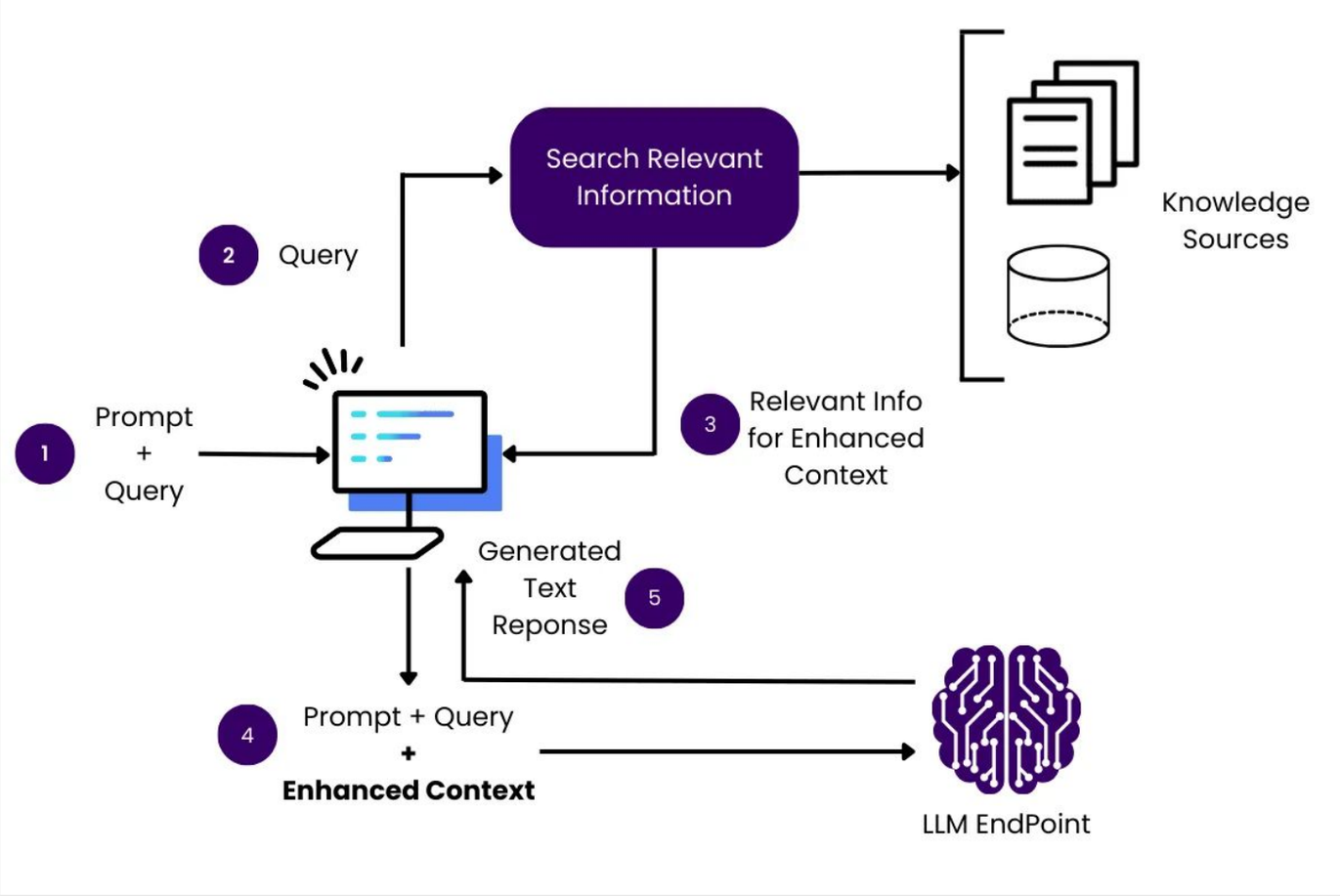


How does it work?

What is Retrieval Augmented Generation(RAG)?

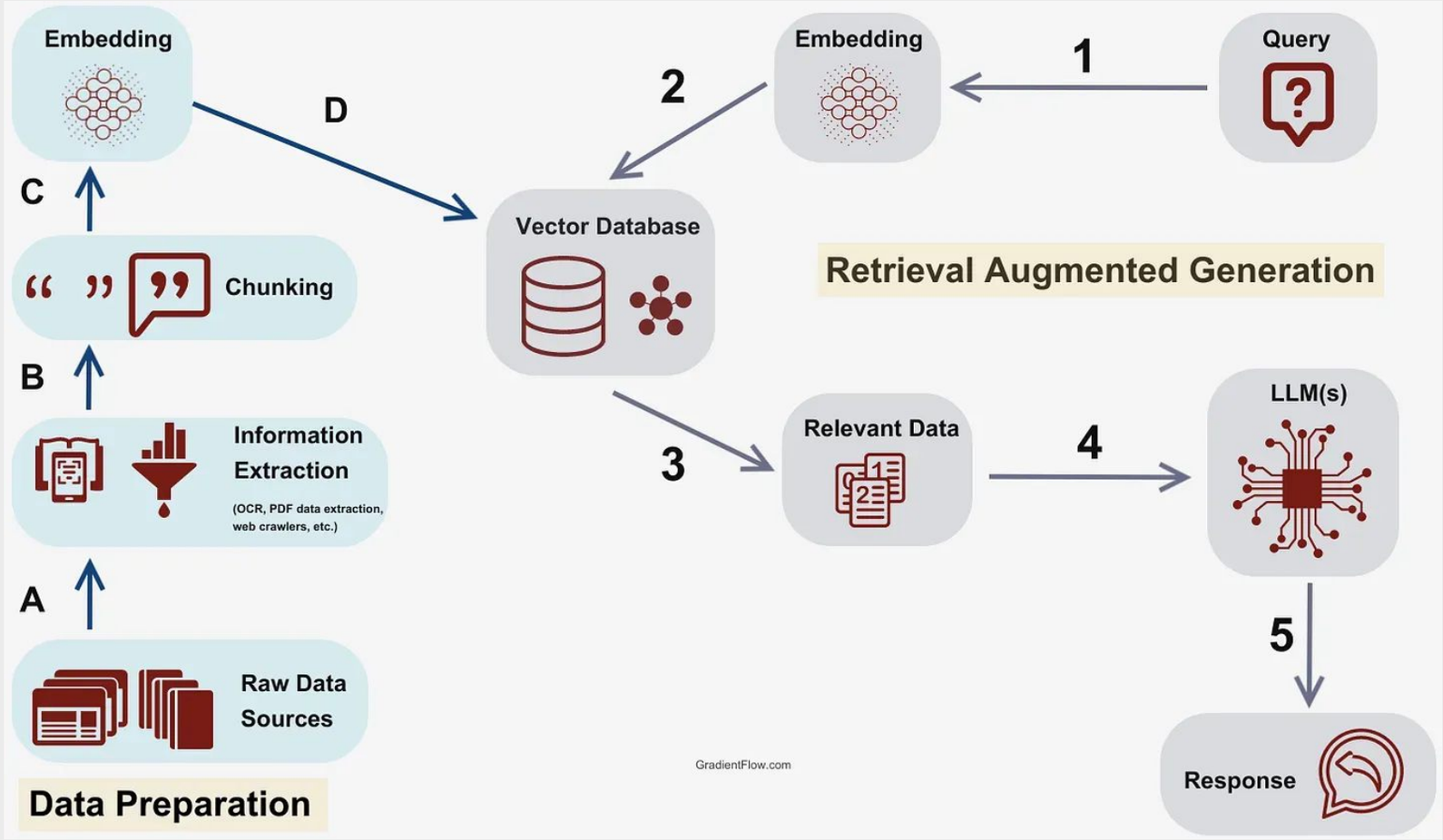
Retrieval Augmented Generation

What is RAG?



How does it work?

- Text chuncking
- Embedding
- Creating a vector Database



Any Questions?

Thank you!

GitHub: https://github.com/Abhinandan-Khurana

Topmate: https://topmate.io/abhinandan_khurana