



Large Language Models- Powered Defense Upgrade

Reshaping the Automated Path for APT Intelligence Operations and Live Drills

ZUOTONG FENG
fengzuotong.fzt@alibaba-inc.com

JIALIANG WU
wujiali.wjl@alibaba-inc.com



Zuotong Feng (@Imanfeng)

- Security Expert at Alibaba Cloud Inc.
- Offensive and defensive security, threat intelligence
- Speaker at POC 2018



Jialiang Wu (@KpLi0rn)

- Senior Security Engineer at Alibaba Cloud Inc.
- Building security capability assessment and measurement solutions

Team for ensure the security of Alibaba Cloud's infrastructure.

Agenda

- **Background & Motivation**
- **AI, Knowledge Graph, and the Future of BAS**
- **From Knowledge to Autonomous Defense Actions**
- **Demo**
- **Summary**



**Today, The Growing Asymmetry
Between Attack and Defense.**

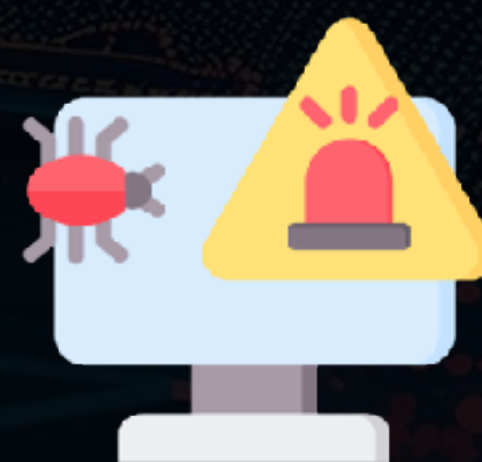
Calculate Cyber Risk Using Attack Paths

“ **Cyber Risk** = Digital Assets (**N**) × Attack Surfaces (**M**) × Vulnerabilities (**K**) ”



Digital Assets

Targets e.g., Cloud Services, Core DBs



Attack Surface

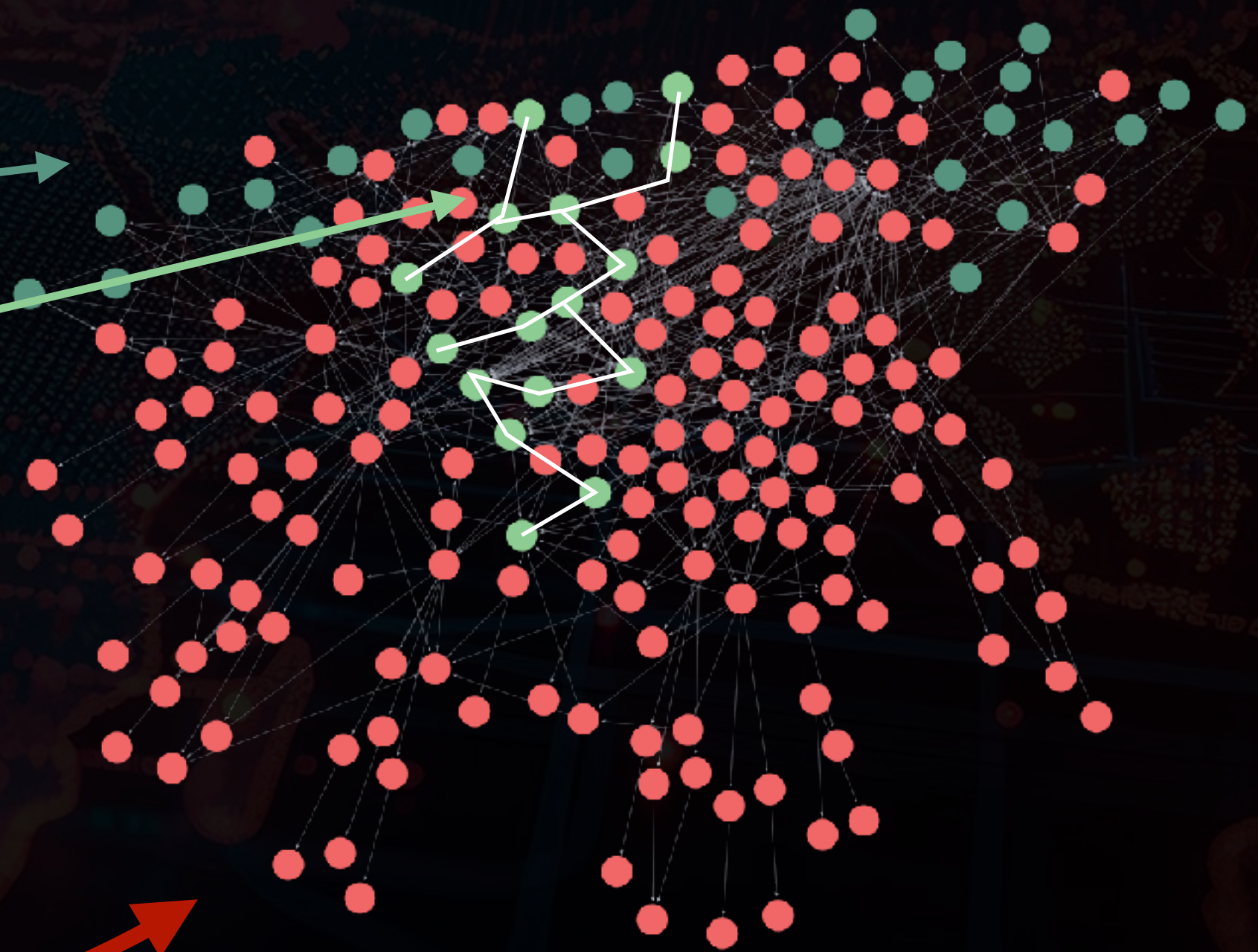
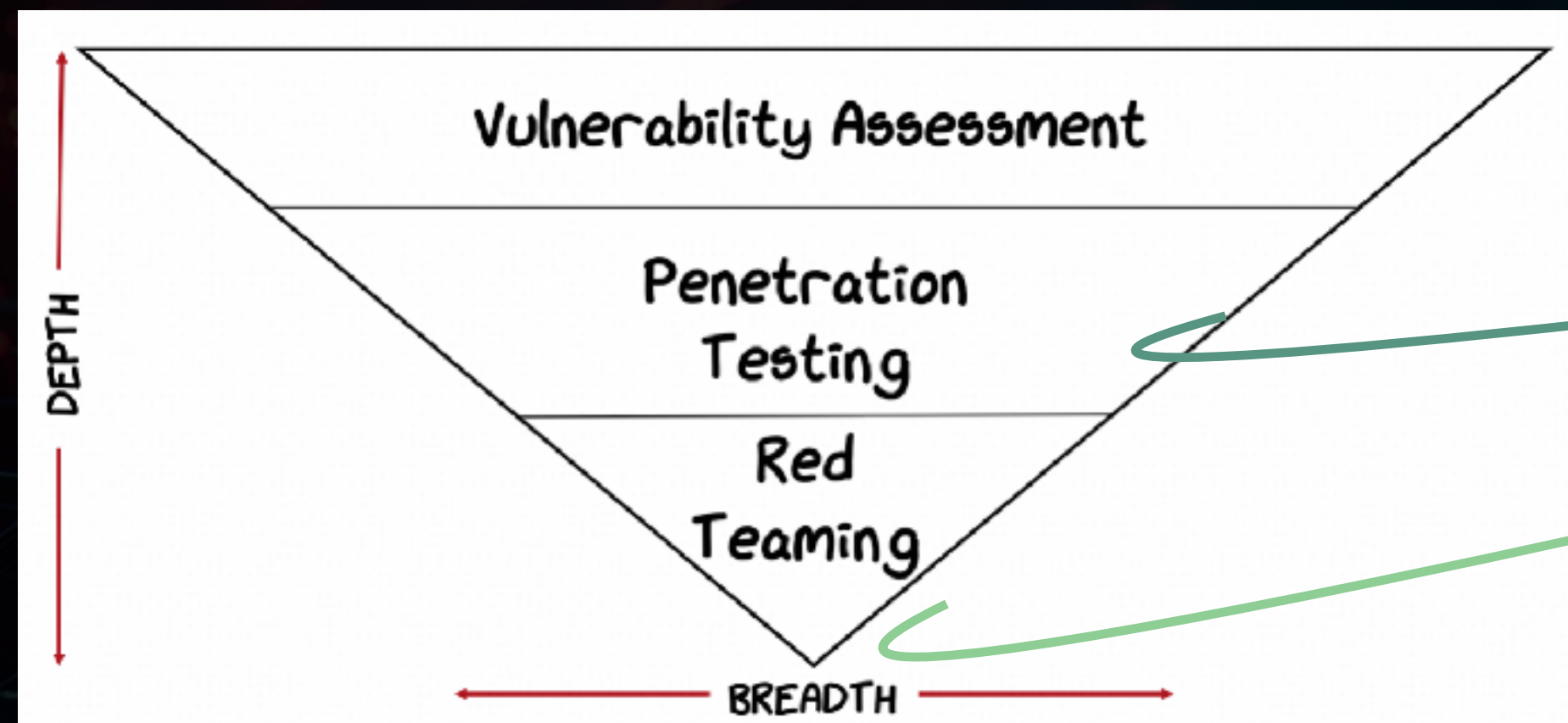
Exposure Points e.g., Open Ports



Vulnerabilities

Exploitable Flaws e.g., Critical CVEs

Pentesting & Red Teaming Are Not Enough: Coverage Gaps



Breach and Attack Simulation : Continuous, automated, scalable



→ BUT ←

- Traditional BAS relies on predefined scenarios
- Attack knowledge evolves faster than human updates

...

BAS with **Attack knowledge** – autonomously operating daily is our ideal path!

APT Intelligence

Red-Blue Confrontation Report

Externally Reported Vulnerabilities

...



However, this ideal path faces critical bottlenecks that demand solutions

The Bottleneck: Attack Knowledge is Fragmented & Unstructured



The Expert Bottleneck

Over-dependent on expert judgment to decode complex attack chains from like APT report, with teams facing 30-40 new APT reports each week.



Fragmented Intelligence

Critical intelligence is fragmented, forcing analysts to piece together across many reports.



Unstructured Text

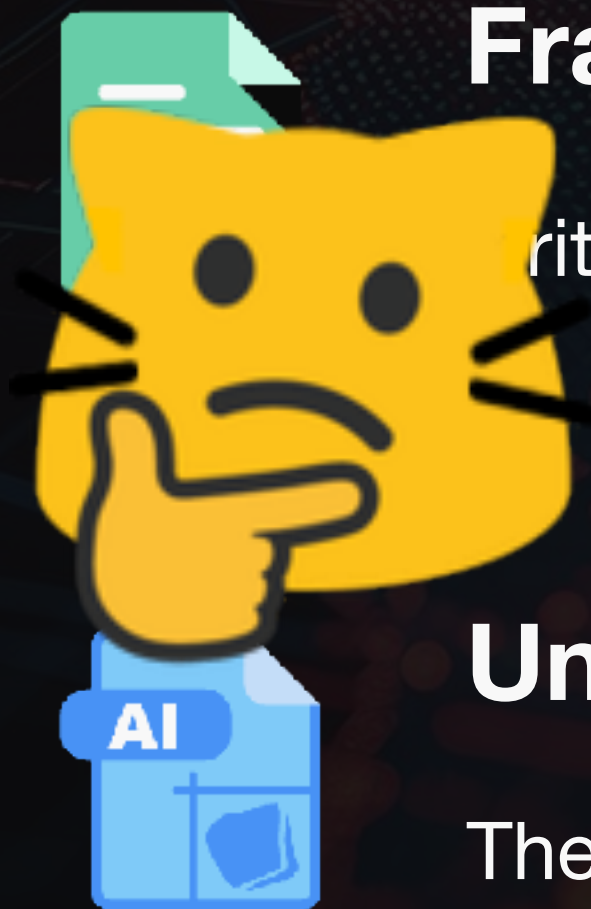
The knowledge from report isn't machine-readable, difficult to directly apply to BAS.

The Bottleneck: Attack Knowledge is Fragmented & Unstructured



The Expert Bottleneck

Over-dependent on expert judgment to decode complex attack chains from like APT report, with teams facing 30-40 new APT reports each week.



Fragmented Intelligence

critical

Wait ... Can AI tackle these Challenges?

ce together across many reports.

Unstructured Text

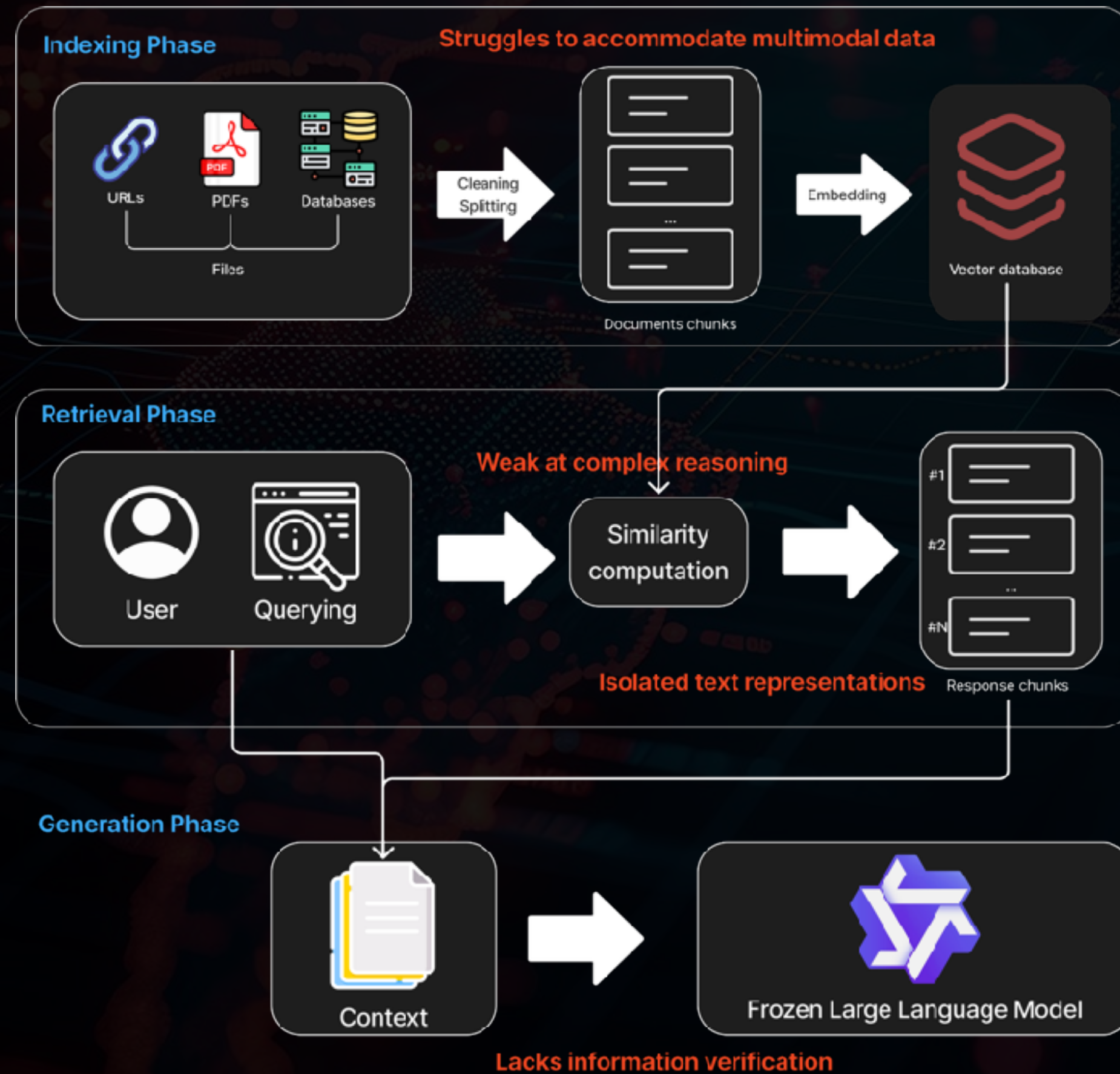
The knowledge from report isn't machine-readable, difficult to directly apply to BAS.

Now, AI can tackle Specialized tasks like Crypto Trading

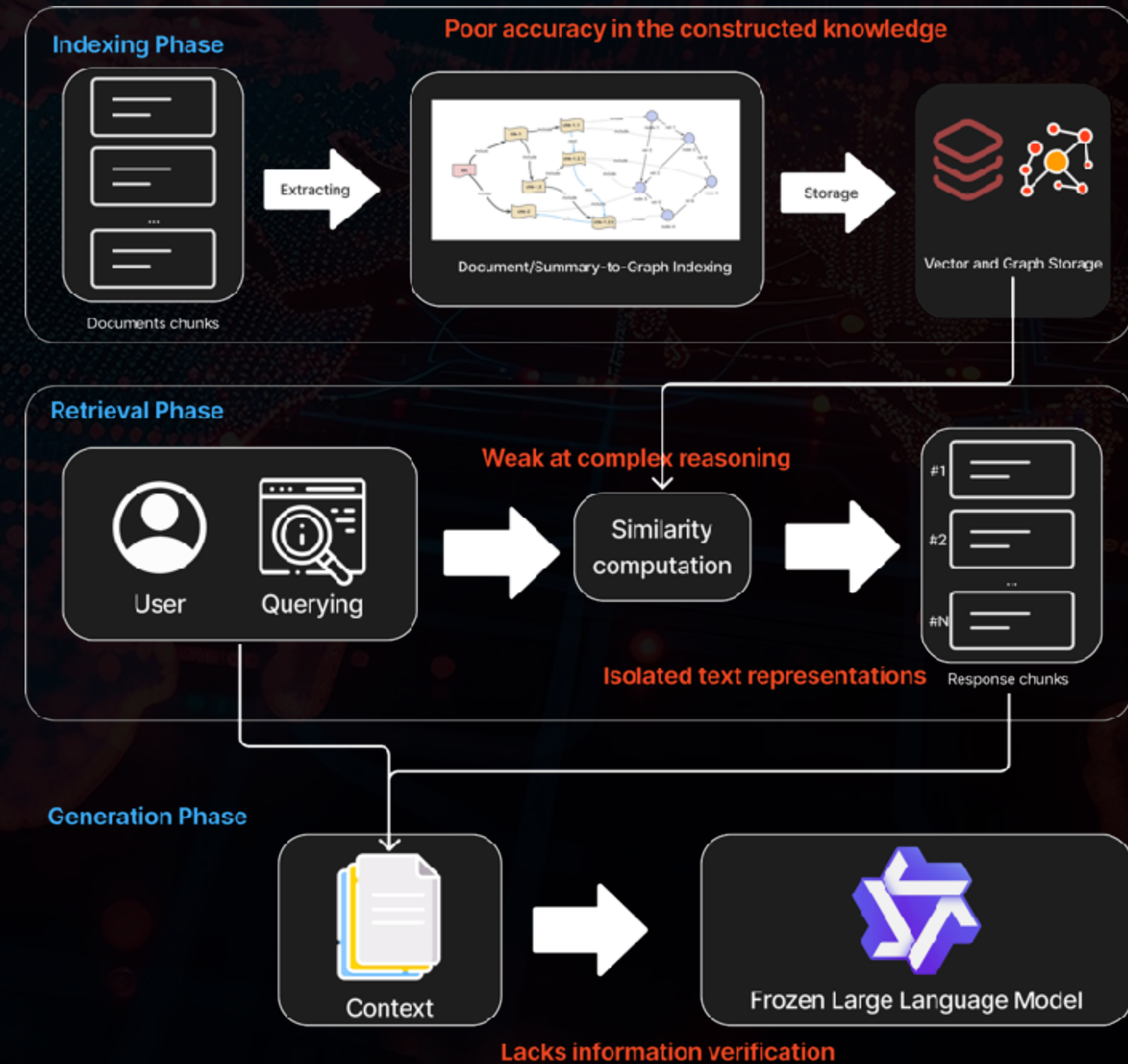


Augment Knowledge always for LLM to Specialized tasks

Common Approach #1: **Doc** RAG



Common Approach #2: **Graph** RAG



Low Accuracy and Poor Logic

Need More Friendly Knowledge in Cybersecurity

Hallucinations Are Unacceptable in Cybersecurity!

Problem: Let $x_1=100$, and $x_2=x_1$. If $x_3=x_2+46$, then what is x_3 ?
Question: If $x_3=x_2+46$, then what is x_3 ?
Answer:

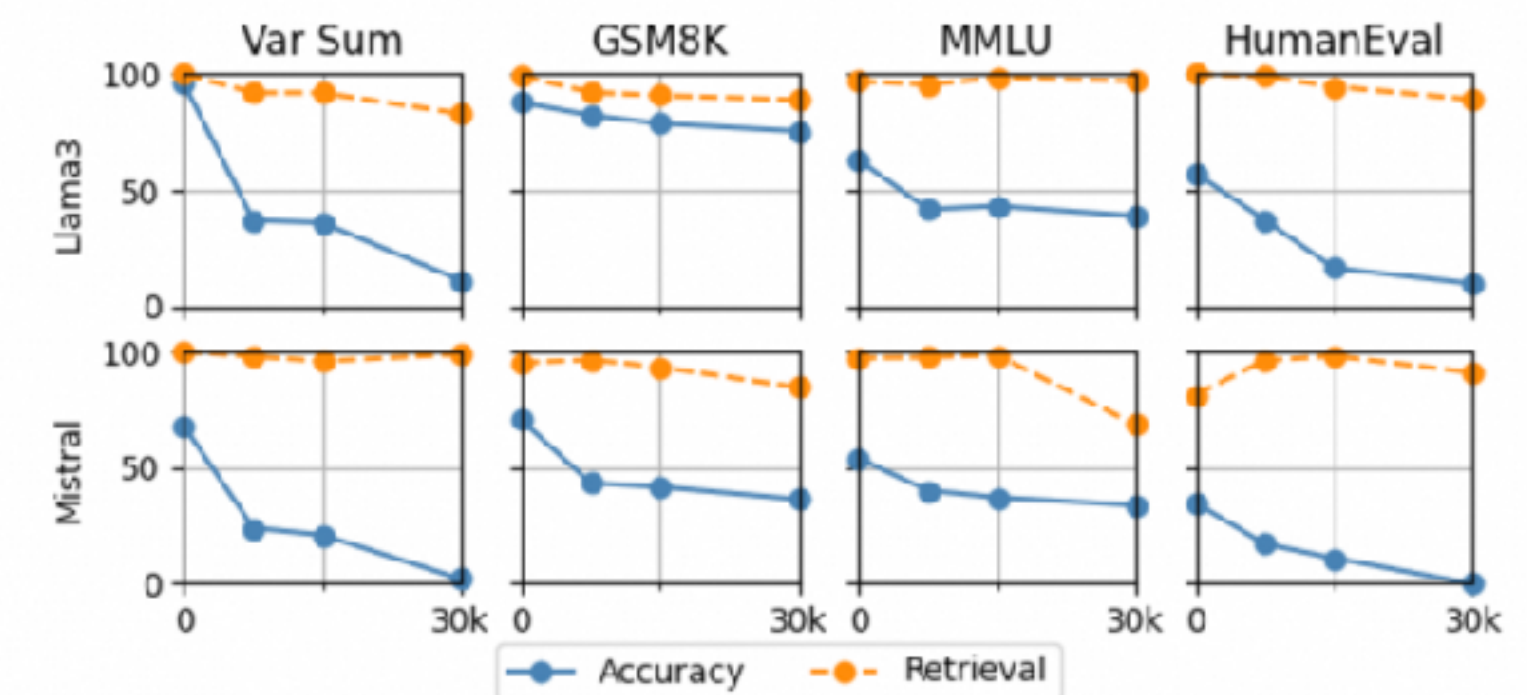
x3=146 ✓

Problem: Let $x_1=100$, and $x_2=x_1$. If $x_3=x_2+46$, then what is x_3 ? (25000 space tokens)
Question: If $x_3=x_2+46$, then what is x_3 ?
Answer:

Here's the step-by-step reasoning:

1. We are given that $x_1 = 100$ and $x_2 = x_1$.
2. Since $x_2 = x_1$, x_2 also equals 100.
3. The equation for x_3 is $x_3 = x_2 + 46$.
4. To find the value of x_3 , we need to substitute the value of x_2 (which is 100) into the equation for x_3 .
5. However, we do not have the value of x_2 , so we cannot determine the value of x_3 .

Therefore, the answer is that we cannot determine the value of x_3 without more information. ✗



Accuracy drops drastically as input length increases





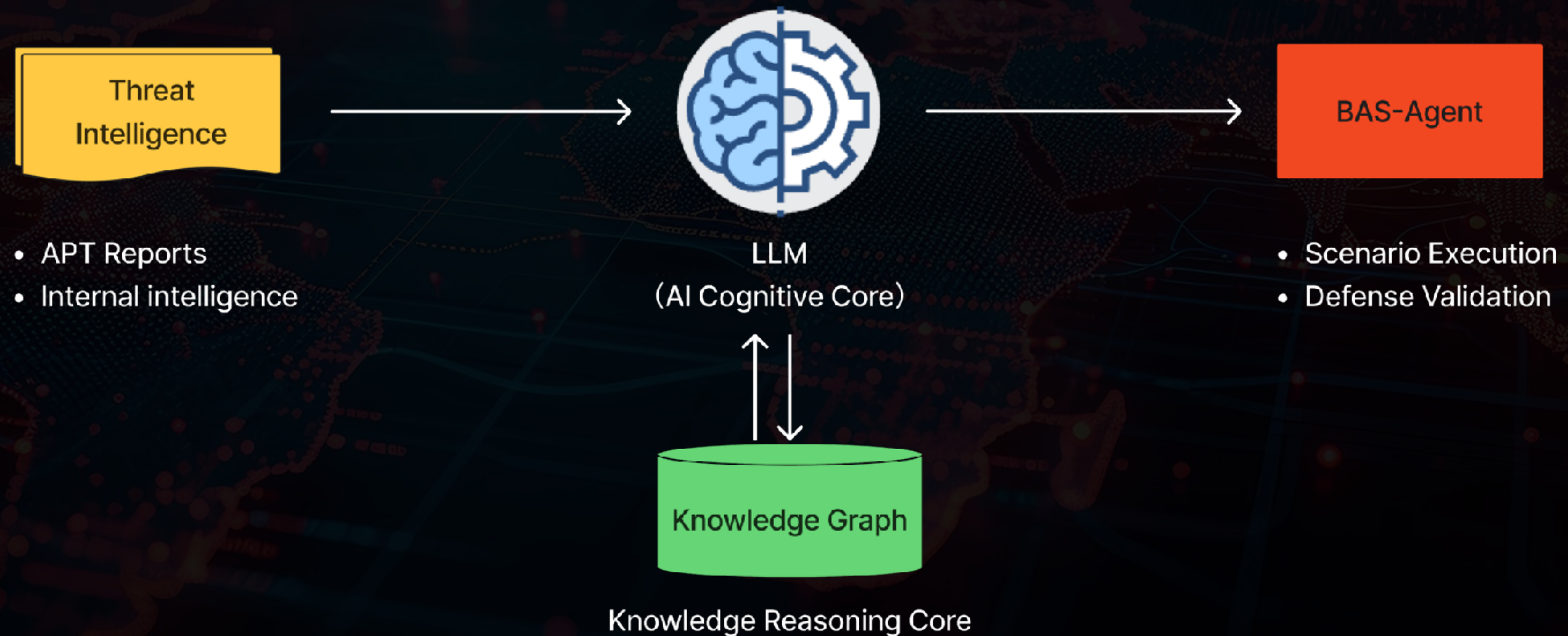
Building Threat Brain Core: AI, KG, and the Future of BAS

Knowledge Graph

LLM x AI Agents

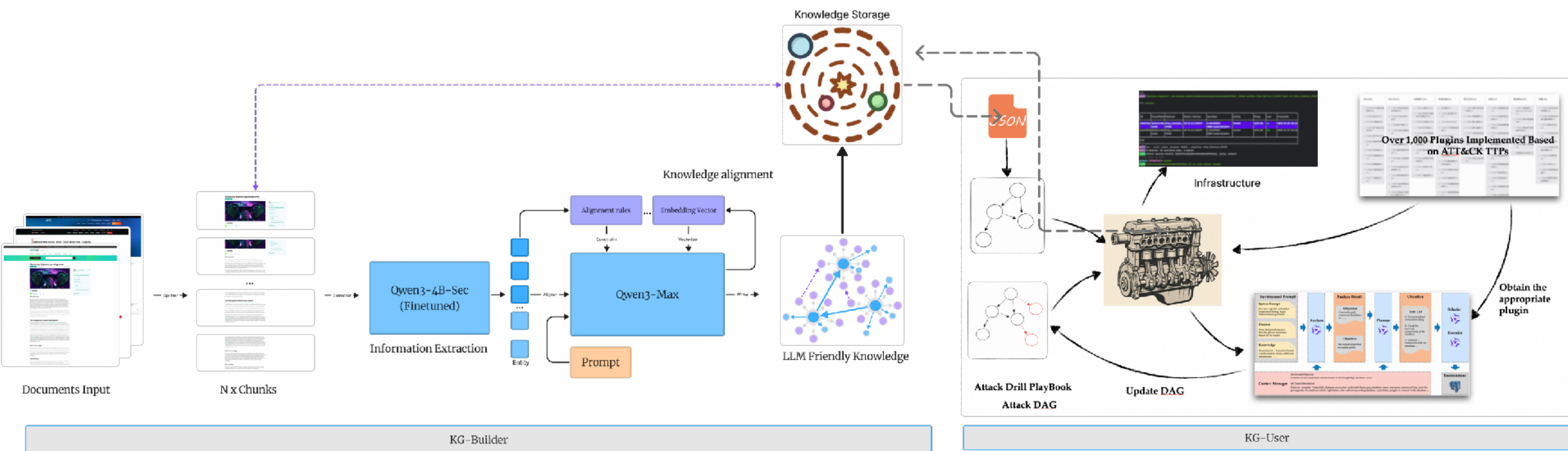
Attack Platform

Our Vision: From Threat Intelligence to Dynamic "Brain"



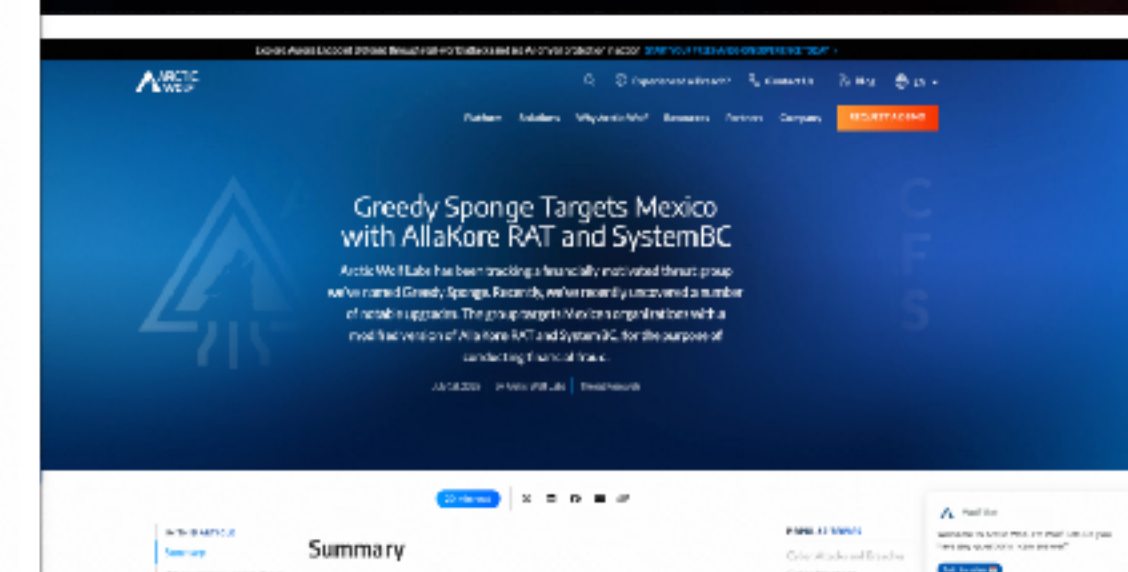
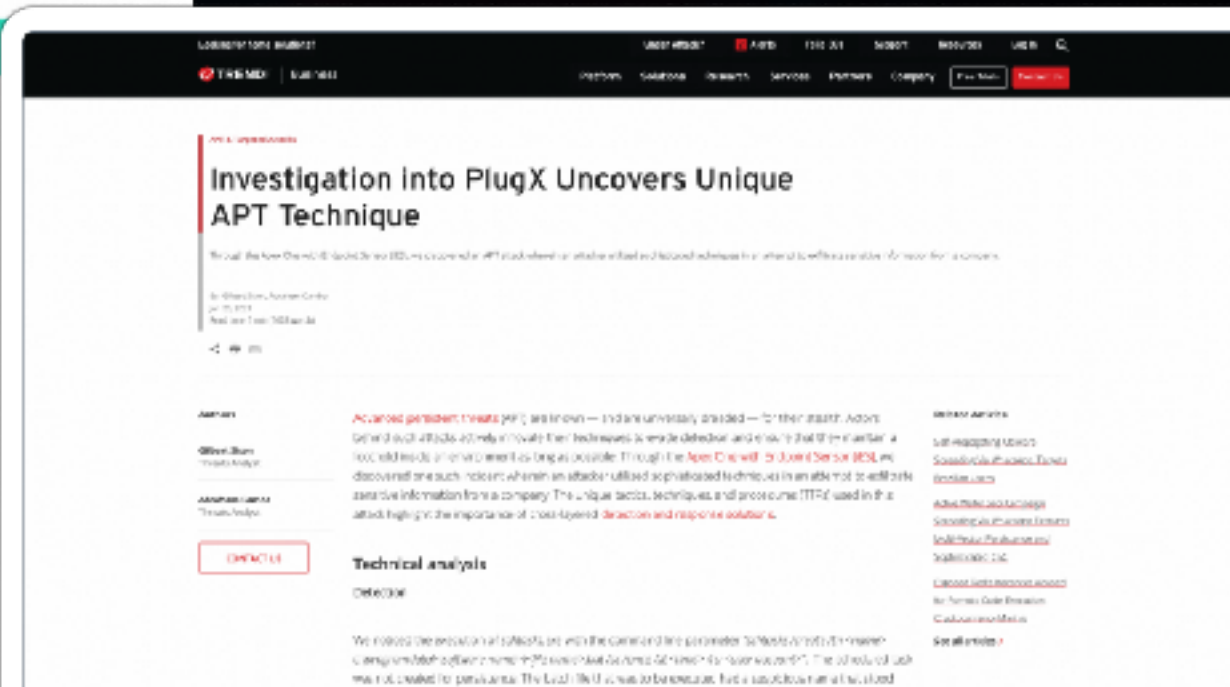
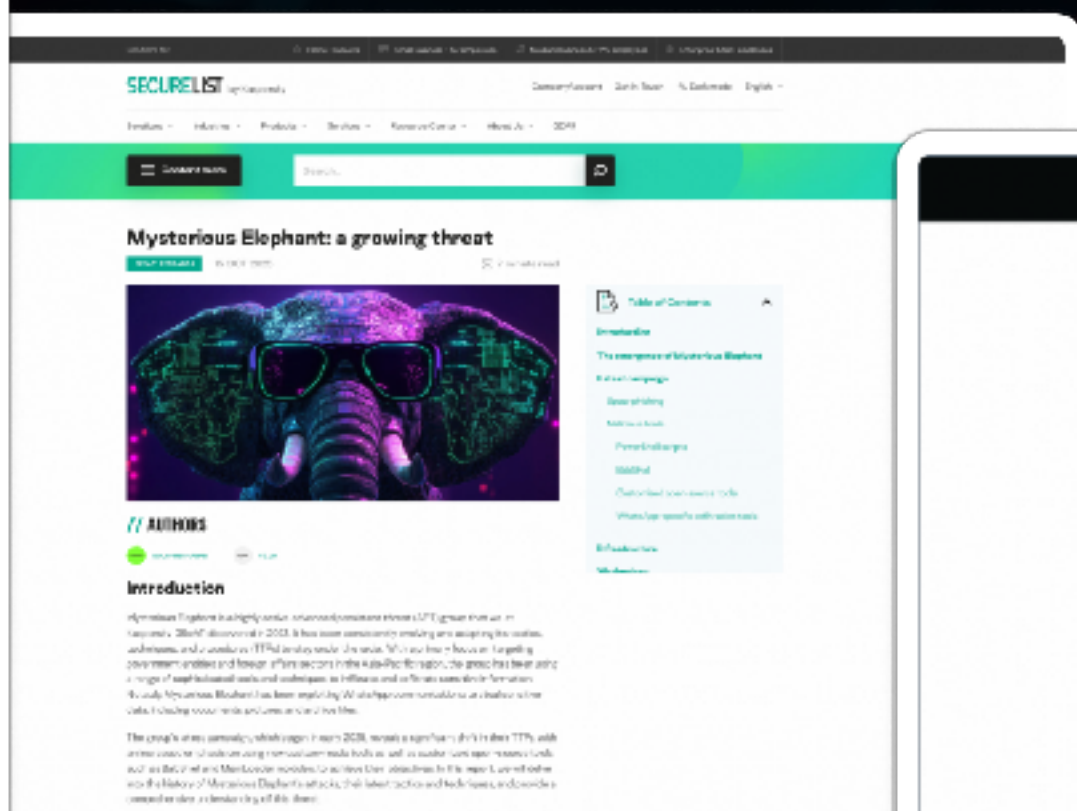
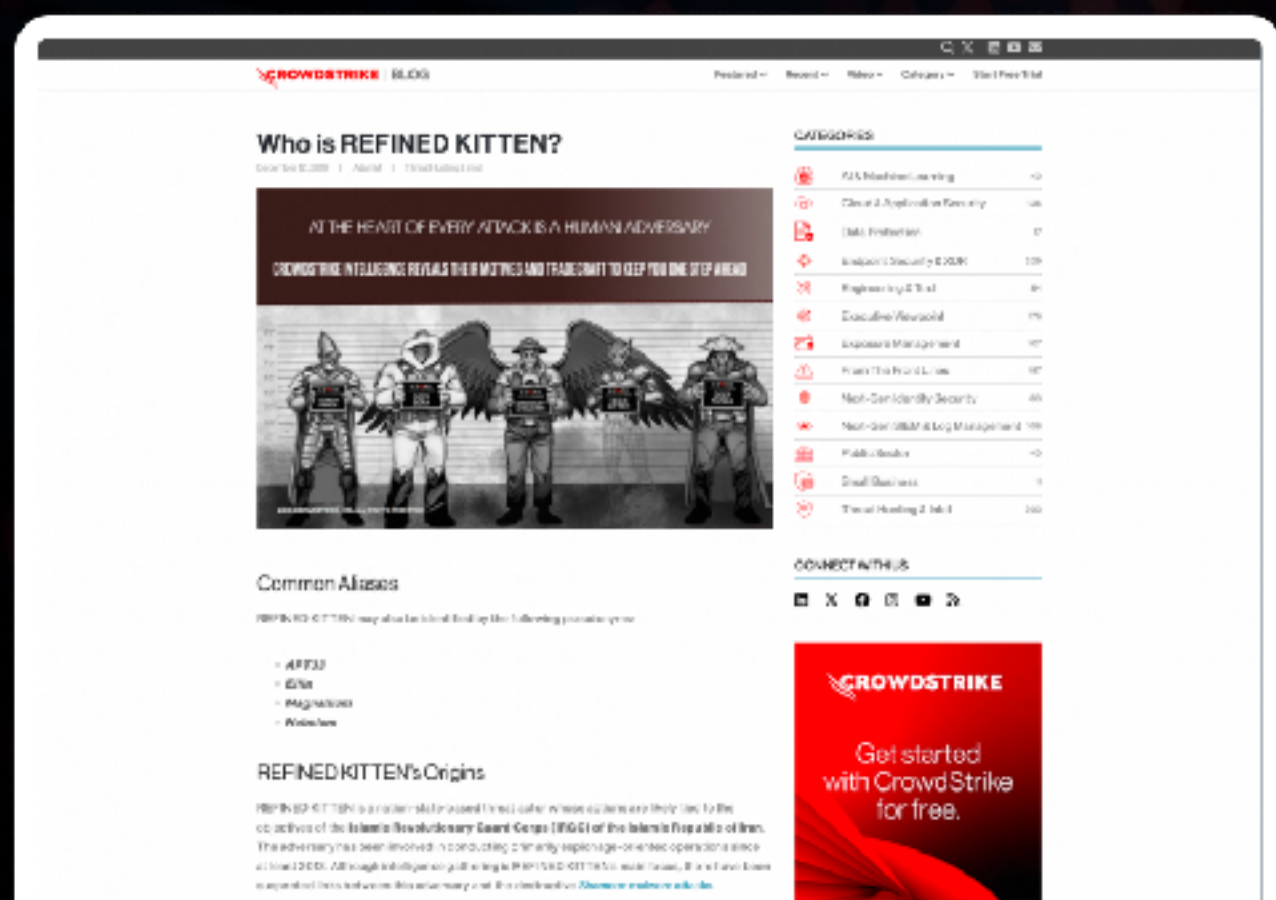
From Unstructured Reports —> Structured Knowledge —> Automated Attack Simulation

Solution Technology Architecture both KG and BAS



Aggregate Internal & External Threat Intelligence

- RSS
- Chrome Headless



Standardize and Parse Threat Intelligence

Format Standardization

Document Chunking

Input HTML/PDF/IMAGE/TXT..

```

1 1. Introduction
2 2. Methodology
3 3. Results
4 4. Discussion
5 5. Conclusion
6 6. References
7 7. Appendix
8 8. Acknowledgments
9 9. Disclaimer
10 10. Contact Information
11 11. Copyright
12 12. License
13 13. Terms and Conditions
14 14. Privacy Policy
15 15. Cookies Policy
16 16. Security Policy
17 17. Privacy Policy
18 18. Cookies Policy
19 19. Security Policy
20 20. Privacy Policy
21 21. Cookies Policy
22 22. Security Policy
23 23. Privacy Policy
24 24. Cookies Policy
25 25. Security Policy
26 26. Privacy Policy
27 27. Cookies Policy
28 28. Security Policy
29 29. Privacy Policy
30 30. Cookies Policy
31 31. Security Policy
32 32. Privacy Policy
33 33. Cookies Policy
34 34. Security Policy
35 35. Privacy Policy
36 36. Cookies Policy
37 37. Security Policy
38 38. Privacy Policy
39 39. Cookies Policy
40 40. Security Policy
41 41. Privacy Policy
42 42. Cookies Policy
43 43. Security Policy
44 44. Privacy Policy
45 45. Cookies Policy
46 46. Security Policy
47 47. Privacy Policy
48 48. Cookies Policy
49 49. Security Policy
50 50. Privacy Policy
51 51. Cookies Policy
52 52. Security Policy
53 53. Privacy Policy
54 54. Cookies Policy
55 55. Security Policy
56 56. Privacy Policy
57 57. Cookies Policy
58 58. Security Policy
59 59. Privacy Policy
60 60. Cookies Policy
61 61. Security Policy
62 62. Privacy Policy
63 63. Cookies Policy
64 64. Security Policy
65 65. Privacy Policy
66 66. Cookies Policy
67 67. Security Policy
68 68. Privacy Policy
69 69. Cookies Policy
70 70. Security Policy
71 71. Privacy Policy
72 72. Cookies Policy
73 73. Security Policy
74 74. Privacy Policy
75 75. Cookies Policy
76 76. Security Policy
77 77. Privacy Policy
78 78. Cookies Policy
79 79. Security Policy
80 80. Privacy Policy
81 81. Cookies Policy
82 82. Security Policy
83 83. Privacy Policy
84 84. Cookies Policy
85 85. Security Policy
86 86. Privacy Policy
87 87. Cookies Policy
88 88. Security Policy
89 89. Privacy Policy
90 90. Cookies Policy
91 91. Security Policy
92 92. Privacy Policy
93 93. Cookies Policy
94 94. Security Policy
95 95. Privacy Policy
96 96. Cookies Policy
97 97. Security Policy
98 98. Privacy Policy
99 99. Cookies Policy
100 100. Security Policy

```



Normalize to Markdown

- Remove Ads and Noise
- Use **Docting** to unified transformation

```

1 1. Introduction
2 2. Methodology
3 3. Results
4 4. Discussion
5 5. Conclusion
6 6. References
7 7. Appendix
8 8. Acknowledgments
9 9. Disclaimer
10 10. Contact Information
11 11. Copyright
12 12. License
13 13. Terms and Conditions
14 14. Privacy Policy
15 15. Cookies Policy
16 16. Security Policy
17 17. Privacy Policy
18 18. Cookies Policy
19 19. Security Policy
20 20. Privacy Policy
21 21. Cookies Policy
22 22. Security Policy
23 23. Privacy Policy
24 24. Cookies Policy
25 25. Security Policy
26 26. Privacy Policy
27 27. Cookies Policy
28 28. Security Policy
29 29. Privacy Policy
30 30. Cookies Policy
31 31. Security Policy
32 32. Privacy Policy
33 33. Cookies Policy
34 34. Security Policy
35 35. Privacy Policy
36 36. Cookies Policy
37 37. Security Policy
38 38. Privacy Policy
39 39. Cookies Policy
40 40. Security Policy
41 41. Privacy Policy
42 42. Cookies Policy
43 43. Security Policy
44 44. Privacy Policy
45 45. Cookies Policy
46 46. Security Policy
47 47. Privacy Policy
48 48. Cookies Policy
49 49. Security Policy
50 50. Privacy Policy
51 51. Cookies Policy
52 52. Security Policy
53 53. Privacy Policy
54 54. Cookies Policy
55 55. Security Policy
56 56. Privacy Policy
57 57. Cookies Policy
58 58. Security Policy
59 59. Privacy Policy
60 60. Cookies Policy
61 61. Security Policy
62 62. Privacy Policy
63 63. Cookies Policy
64 64. Security Policy
65 65. Privacy Policy
66 66. Cookies Policy
67 67. Security Policy
68 68. Privacy Policy
69 69. Cookies Policy
70 70. Security Policy
71 71. Privacy Policy
72 72. Cookies Policy
73 73. Security Policy
74 74. Privacy Policy
75 75. Cookies Policy
76 76. Security Policy
77 77. Privacy Policy
78 78. Cookies Policy
79 79. Security Policy
80 80. Privacy Policy
81 81. Cookies Policy
82 82. Security Policy
83 83. Privacy Policy
84 84. Cookies Policy
85 85. Security Policy
86 86. Privacy Policy
87 87. Cookies Policy
88 88. Security Policy
89 89. Privacy Policy
90 90. Cookies Policy
91 91. Security Policy
92 92. Privacy Policy
93 93. Cookies Policy
94 94. Security Policy
95 95. Privacy Policy
96 96. Cookies Policy
97 97. Security Policy
98 98. Privacy Policy
99 99. Cookies Policy
100 100. Security Policy

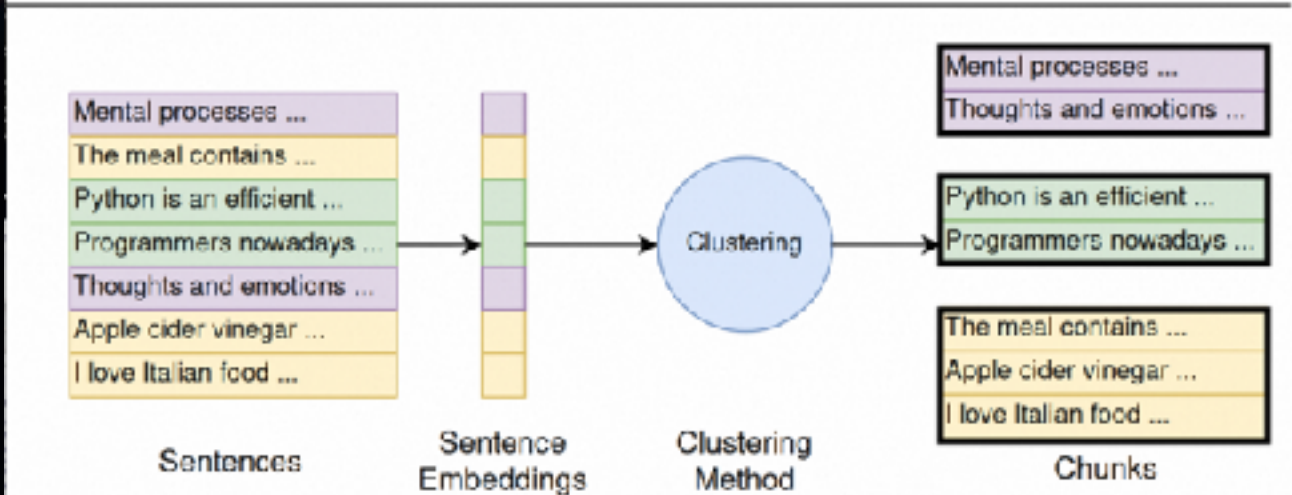
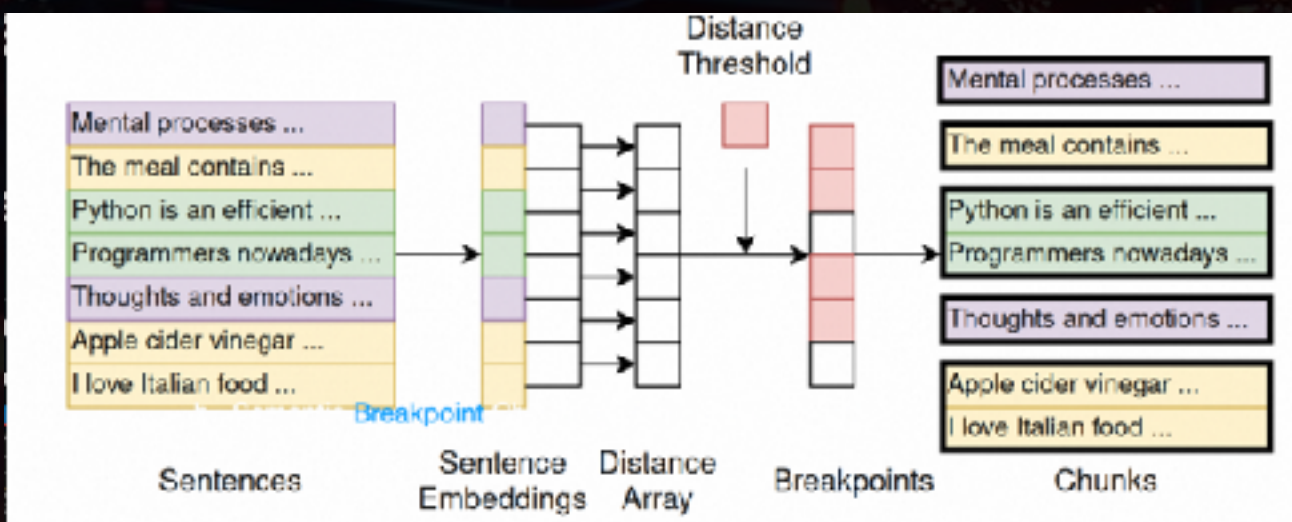
```

- Using visual model analysis images



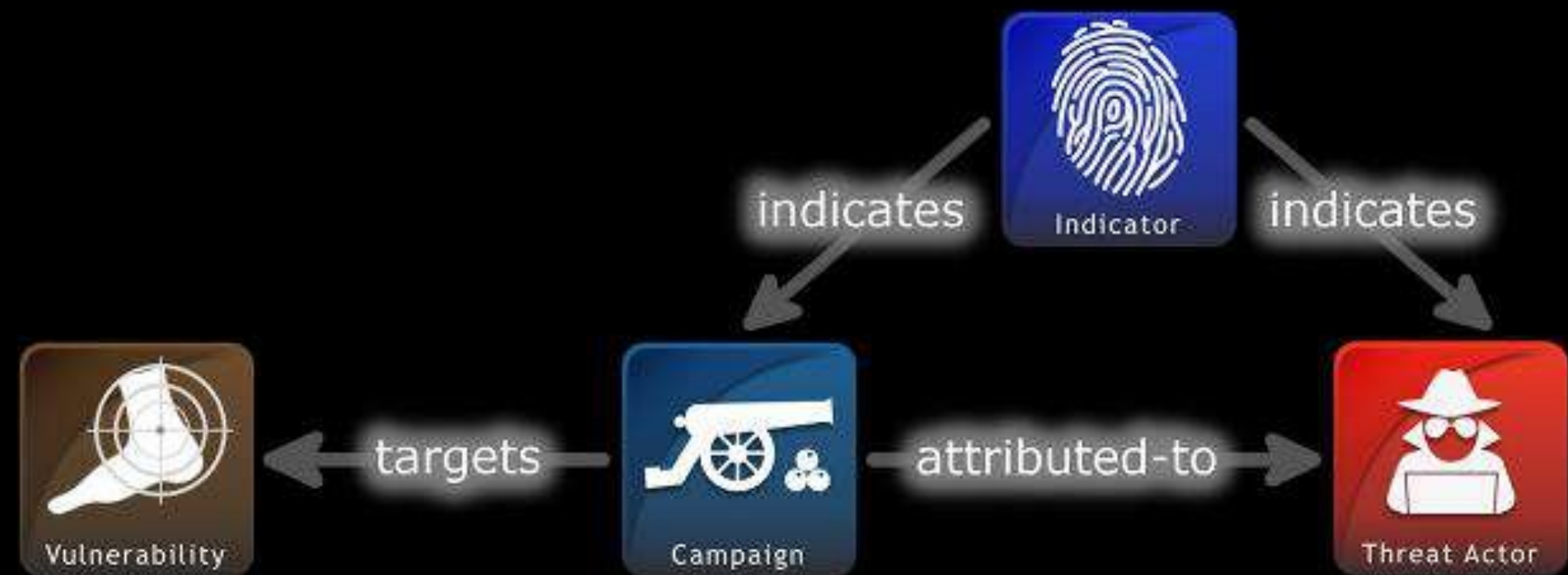
Cut N x Chunks

- Fixed-Size Chunking: Split by token
- Structure-Aware Chunking: Aligns with document hierarchy
- Semantic Breakpoint Chunking: Detect "Difference"
- Semantic Clustering Chunking: Find "Similarity"



Defining Core Entities and Relationships

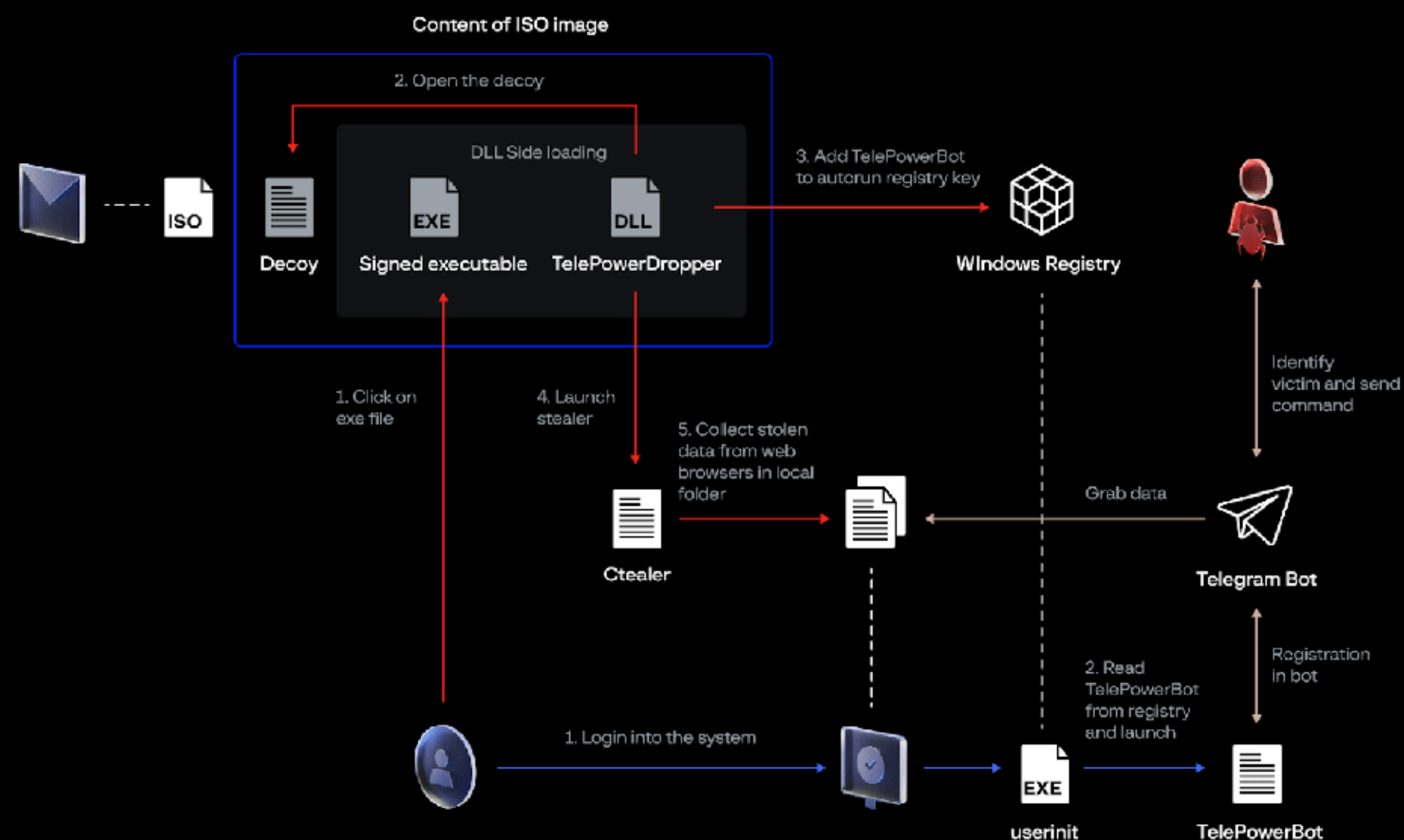
- Threat actors, malware, TTPs, infrastructure as key **Entities**
- **Relations** with inverse pairs, e.g., (uses, used_by) and (attacks, attacked_by)



STIX 2.1 Domain Objects (SDOs)			
	Attack Pattern A type of TTP that describes ways that adversaries attempt to compromise targets		Campaign A grouping of attacker behaviors describing a set of malicious activities or attacks
	Course of Action A recommendation on the actions to take in response to intelligence shared		Grouping Explicitly asserts that the referenced STIX Objects have a shared context (unlike a STIX Bundle)
	Identity Actual individuals, organizations, or groups and classes of these (e.g. finance sector)		Indicator Contains a pattern that can be used to detect suspicious or malicious cyber activity
	Infrastructure Type of TTP that describes any systems, software services, or resources intended to support some purpose		Intrusion Set A set of adversarial behaviors and resources with common properties associated with a single organization
	Location Represents a geographic location		Malware A type of TTP that represents malicious code
	Malware Analysis The metadata and results of a particular static or dynamic analysis performed on a malware instance or family		Note Informative text to provide further context and/or analysis not contained in the STIX Objects
	Observed Data Information about cyber security related entities such as files, systems, and networks using the STIX Cyber-observable Objects (SCOs)		Opinion An assessment of the correctness of the information in a STIX Object produced by a different entity
	Report Collections of threat intelligence focused on one or more topics, such as a descriptions that include context and related details		Threat Actor Actual individuals, groups, or organizations believed to be operating with malicious intent
	Tool Legitimate software that can be used by threat actors to perform attacks		Vulnerability A mistake in software that can be directly used by a hacker to gain access to a system or network

Chaining Complex Cyber Attacks in Knowledge Graphs

DARK PINK APT KILL CHAIN 1: ALL-INCLUSIVE ISO



Source: Group-IB, 2023.

Initial Access Event

T1566.001 (Spear-Phishing Attachment)

Attacker ↔ Spear-Phishing Email ↔ ISO File ↔ Victim

Execution Event

T1204.002 (User Execution: Malicious File)

Victim ↔ EXE ↔ DLL ↔ Payload

Persistence Event

T1574.001 (Hijack Execution Flow: DLL Side-loading)

T1547.001 (Boot or Logon Autostart Execution: Registry Run Keys / Startup Folder)

DLL ↔ Registry ↔ Legitimate Program ↔ System Process

□ □ □

Core Challenge: High Cost of Knowledge Graph Construction

High-Quality Knowledge Graphs Require Significant Effort and Expertise

① LLM-Driven Automation

- Automate **entity**, **attribute**, and **relation** extract

Reduces human effort >70%

② Noise Mitigation via Multi-Level Alignment

- Schema alignment e.g., **ATT&CK v18**
- Semantic alignment to resolve ambiguity



Core Challenge: Knowledge Sparsity in KGs

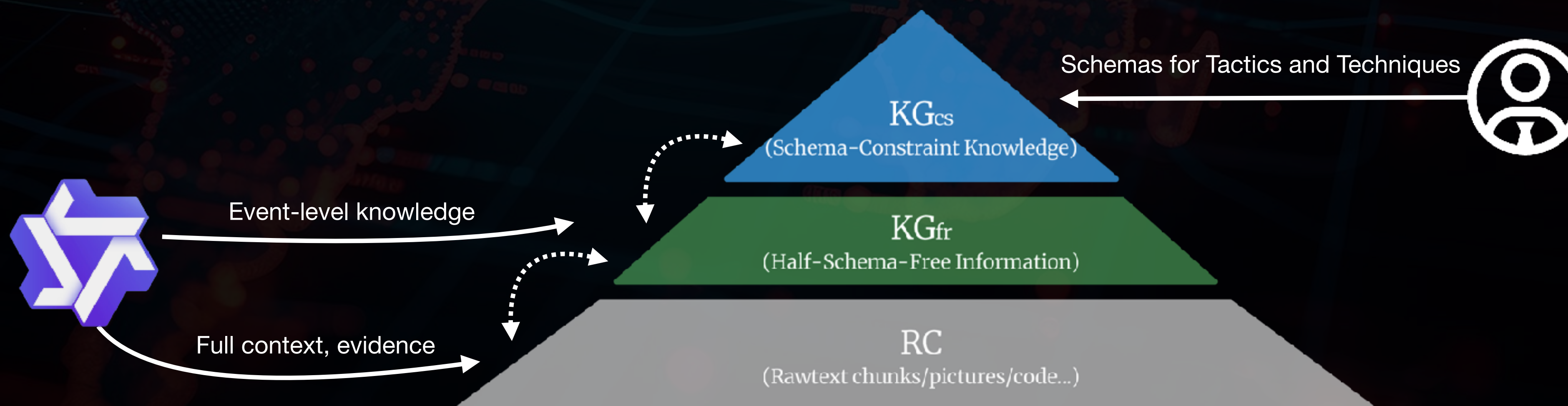
Knowledge Graphs Only Contain Entities, Relations, and Attributes, Missing rich context from APT reports

① Three-tiers Knowledge Structure

- **KGcs**: TTP-aligned, formalized knowledge
- **KGfr**: Extracted via LLM using **Event Schema**
- **RC**: Raw text chunks, codes, images

② Mutual Indexing Structure

- Every node linked to source **RC**
- LLM: Structural reasoning + evidence retrieval



ATT&CK-Native Event Layering: Macro → Meso → Micro

Tactic {

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Exfiltration	Command and Control
Drive-by Compromise	AppleScript	bash_profile and .bashrc	Access Token Manipulation	Access Token Manipulation	Account Manipulation	Account Discovery	AppleScript	Audio Capture	Automated Exfiltration	Commonly Used Port
Exploit Public-Facing Application	CMSTP	Accessibility Features	Accessibility Features	BITS Jobs	Bash History	Application Window Discovery	Application Deployment Software	Automated Collection	Data Compressed	Communication Through Removable Media
Hardware Additions	Command-Line Interface	AppCert DLLs	AppCert DLLs	Binary Padding	Brute Force	Browser Bookmark	Distributed Component	Clipboard Data	Data Encrypted	Connection Proxy
Replication Through Removable Media	Control Panel Items	AppInit DLLs	AppInit DLLs	Bye Acc					Data Transfer Size Limits	Custom Command and Control

Drive-by Compromise

A drive-by compromise is when an adversary gains access to a system through a user visiting a website over the normal course of browsing. With this technique, the user's web browser is targeted for exploitation. This can happen in several ways, but there are a few main components:

Multiple ways of delivering exploit code to a browser exist, including:

- A legitimate website is compromised where adversaries have injected some form of malicious code such as JavaScript, frames, cross-site scripting
- Malicious ads are paid for and served through legitimate ad providers.
- Built-in web application interfaces are leveraged for the insertion of any other kind of object that can be used to display web content or contain a script that executes on the visiting client (e.g. forum posts, comments, and other user-controllable web content).

Often the website used by an adversary is one visited by a specific community, such as government, a particular industry, or religion, where the goal is to compromise a specific user or set of users based on a shared interest. This kind of targeted attack is referred to as a strategic web compromise or watering hole attack. There are several known examples of this occurring.^[1]

Drive-by Compromise Technique

ID	T1189
Tactic	Initial Access
Platform	Linux, Windows, macOS
Permissions Required	User
Data Sources	Packet capture, Network device logs, Physical use of network, Web proxy, Network intrusion detection system, SSL/TLS inspector

Technique

Procedure

• Macro Event -> MITRE ATT&CK® Campaigns

TTP Chains

• Meso Event -> MITRE ATT&CK® Tactics

Atomic TTPs

• Micro Event -> MITRE ATT&CK® Techniques

Defining Core Events and Relationships

- Structured **Events** via semantic schemas (**entities + relations**)
- Cross-event links enable multi-hop reasoning

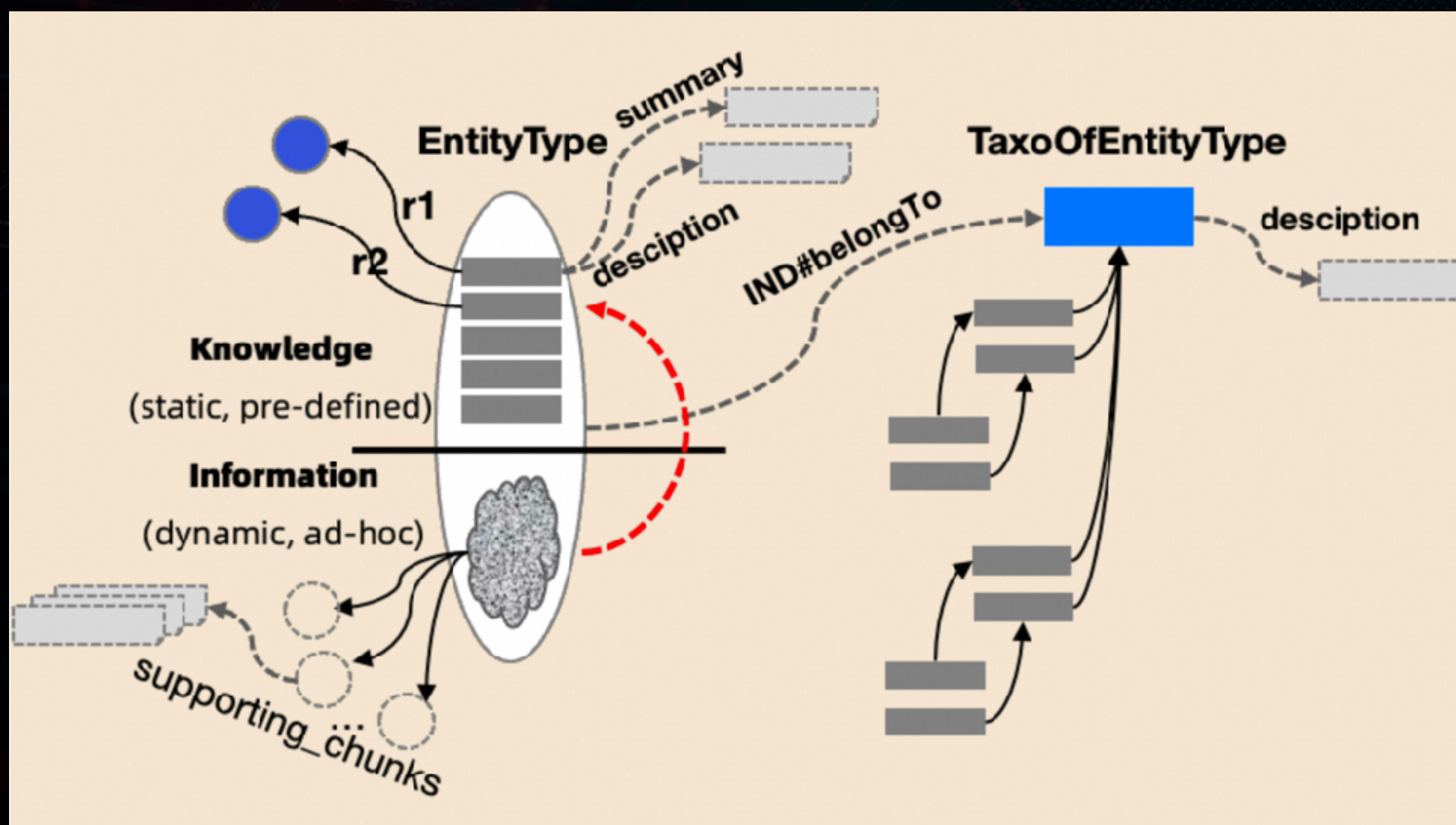


```
...
name: "Initial Access - Delivery & User Execution"
description:
  Captures scenarios where an attacker delivers or makes available a malicious payload
  to a target user or system, using any lure, distribution channel, or social engineering
  technique, leading to potential initial compromise.
constraints:
  required_entities:
    - type: Actor
    - type: Method
    - type: Asset
    - type: Payload
  optional_entities:
    - type: Target
  required_relations:
    - pattern: "{Actor} -(use|deploy|launch)-> {Method}", optional: false
    - pattern: "{Method} -(deliver|drop|distribute|provide)-> {Payload}", optional: false
    - pattern: "{Payload} -(targets|aims_at|affects)-> {Asset}", optional: true
    - pattern: "{Target} -(receive|download|execute|open)-> {Payload}", optional: true
    - pattern: "{Actor} -(target)-> {Target}", optional: true
    - pattern: "{Asset} -(owned_by|used_by)-> {Target}", optional: true
  tactic: "Initial Access"
  matching_logic:
    - at_least:
        - 1 relation connecting Actor and Method
        - 1 relation connecting Method and Payload
        - 1 relation connecting Payload and Asset
    - all_required_entities_present: true
...
```

e.g., Meso Event - 'Initial Access - Delivery & User Execution'

Defining LLM-Friendly SPG Attribute: Knowledge + Information

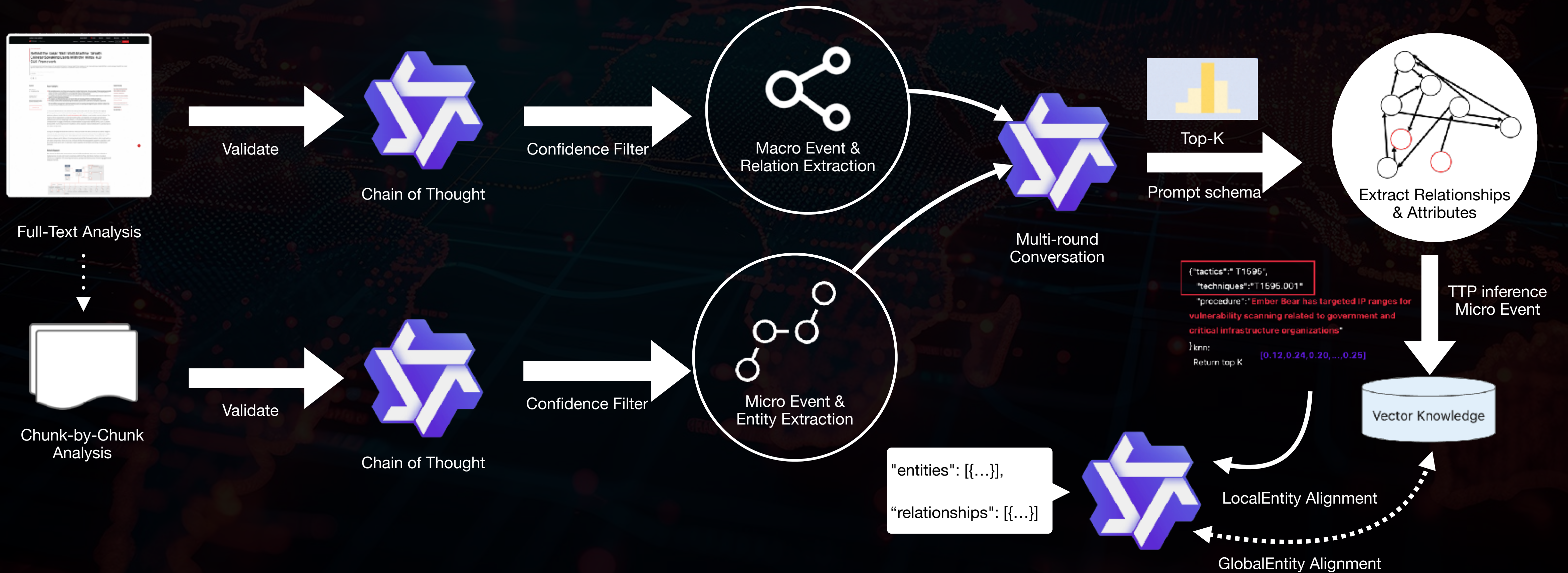
- **Knowledge Zone:** Semantic stability, Structured, Reusable knowledge elements
- **Information Zone:** Context-dependent, Mutable factual information



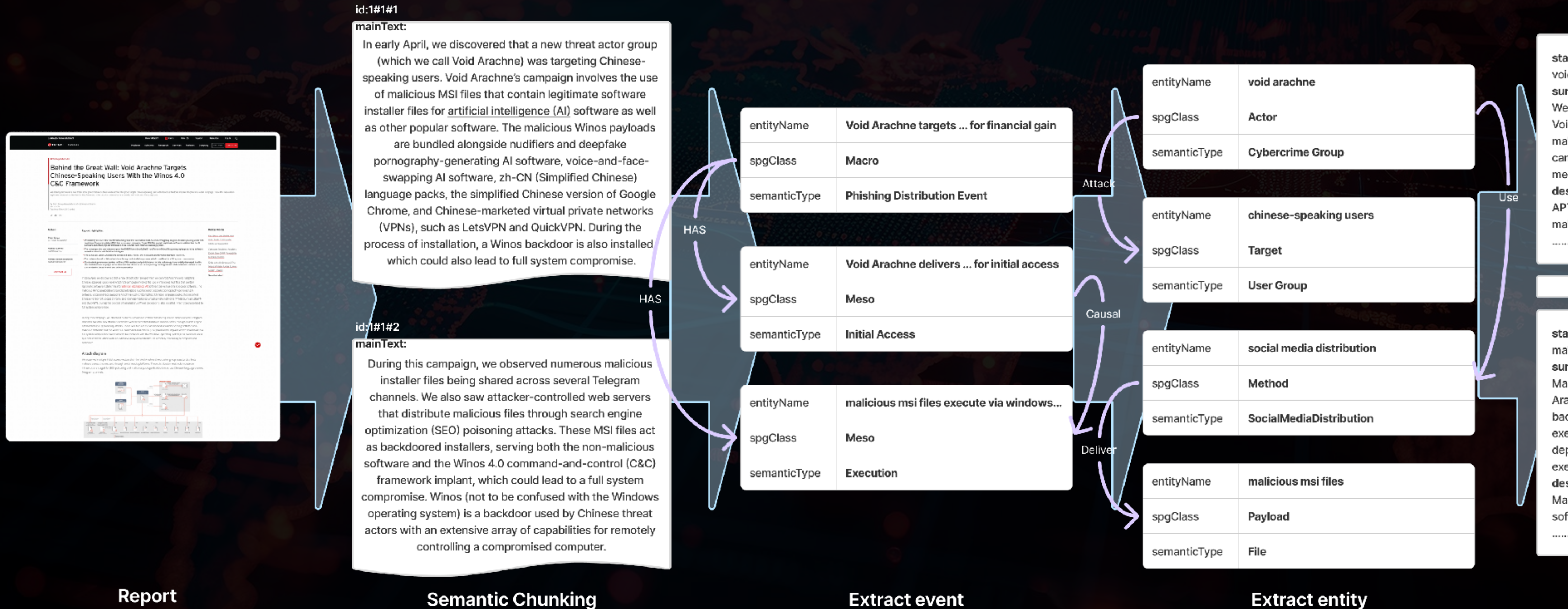
```
{
  "id": "macro_actor_01",
  index: Text
  "type": "Actor",
  "name": "void arachne",
  index: TextAndVector
  "nameType": "specific",
  "description": "APT group specializing in SEO poisoning and
social media malware distribution targeting ...",
  index: TextAndVector
  "summary": "Void Arachne is a threat actor group ...",
  "attributes": {
    "aliases": [],
    "country": "Unknown",
    "region": "Unknown",
    "actorType": "organization",
    "primaryMotivation": "economic_gain",
    "targetCountries": ["CN"],
    "targetRegion": "East Asia",
    "targetIndustries": "Unknown",
    "sophisticationLevel": "Organized Advanced Threat",
    "semanticType": "Cybercrime Group",
    "confidence": 0.98
  }
}
```

e.g., Actor.schema

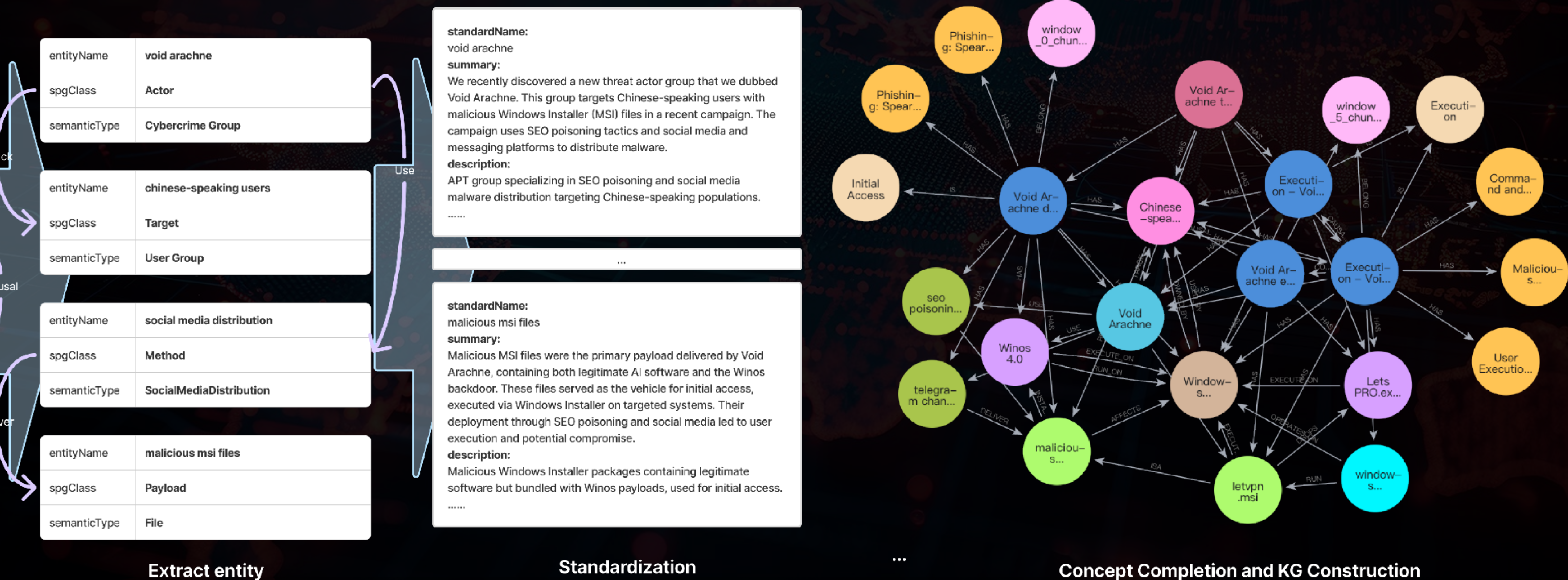
LLM-Driven Information Extraction Pipeline



Knowledge Alignment & Knowledge Graph Construction Pipeline



Knowledge Alignment & Knowledge Graph Construction Pipeline

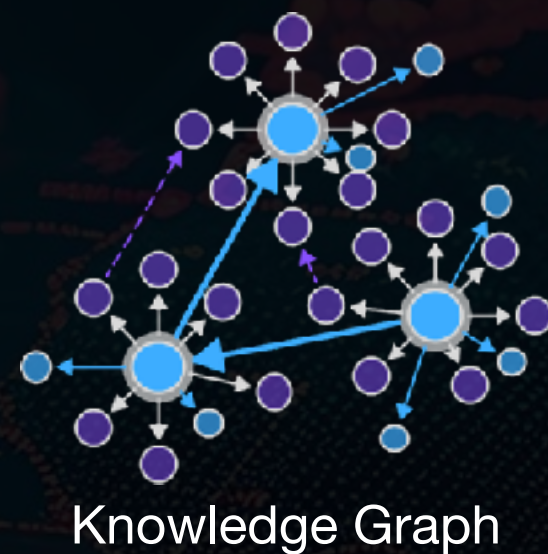


DEMO: Using DSL to Query APT Event Entities via SPG Graph

① Identify the core APT event



Search: What happened in the CL-STA-0969 targeting event



② Query Meso events via DSL on SPG graph

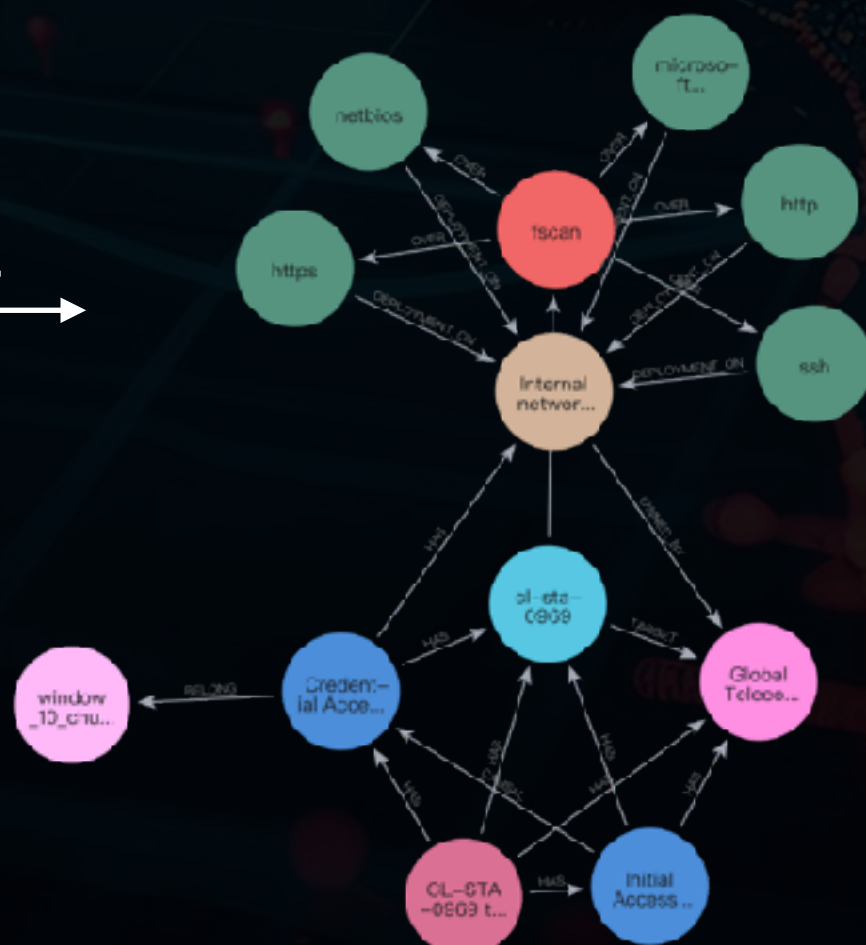
DSL:
MATCH (macro:Macro)-[r1:HAS]->(meso:Meso)
WHERE macro.name = 'CL-STA-0969 targets telecommunications infrastructure in Southwest Asia for espionage'



③ Retrieve all related Meso entities using DSL

DSL:
MATCH (m:Meso {id: "meso_1bed2ba5"})-[r]-(related)
WITH collect(DISTINCT related.id) as relatedIds

MATCH (a)-[r]-(b)
WHERE a.id IN relatedIds AND b.id IN relatedIds AND a.id < b.id
RETURN a, r, b



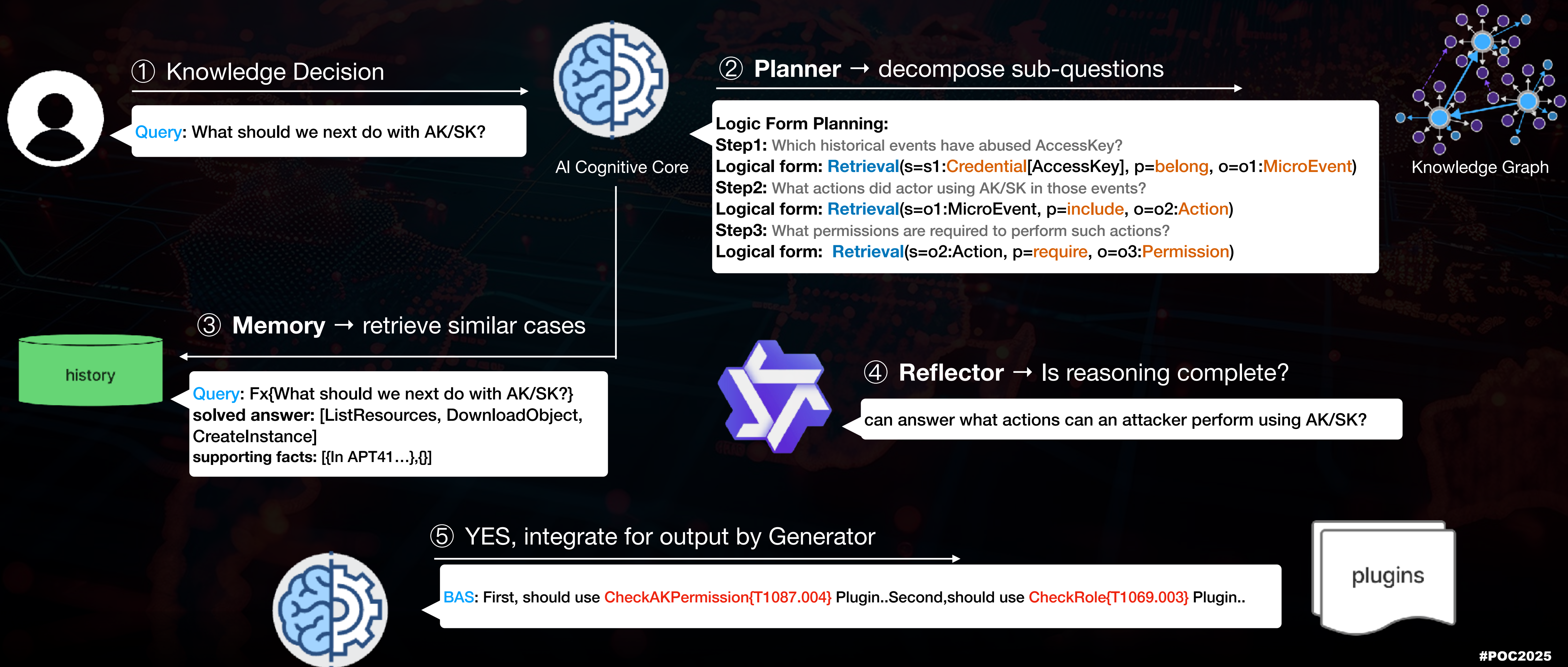
④ Into LLM for Augmented Generation

Response

```
"entities": [
  {
    "id": "micro_6853af78",
    "entityType": "Micro",
    "name": "Modify PAM Authentication Process",
    "nameType": "categorical",
    "description": "AuthDoor overwrites pam_unix.so and hooks into pam_sm_authenticate to intercept and capture user credentials during authentication.",
    "content": "Text explicitly describes overwriting PAM module and hooking authentication function; mapped to T1556.001 as it involves modifying PAM for credential access and persistence.",
    "attributes": {
      "technique": "T1556.001 - Modify Authentication Process: Pluggable Authentication Modules (PAM)"
    }
  },
  {
    "tactic": "Persistence",
    "name": "T1547.003 - Pluggable Authentication Modules",
    "description": "Implemented AuthDoor backdoor in PAM modules to provide persistent access via hard-coded magic password and credential"
```



DEMO: Enhancing Inference-Driven Critical Path Analysis with Knowledge





Next, We will Introduce

**How The BAS Agent Maximizes the value of
this Condensed Knowledge**

“ **Cyber Risk** = Digital Assets (**N**) × Attack Surfaces (**M**) × Vulnerabilities (**K**) ”

As time progresses, Cyber Risk will only continue to increase



How can we allocate our limited resources more rationally as the defending side

Traditional fuzzing performs poorly in intrusion scenarios



Excessive invalid TTP combinations

In intrusion scenarios, TTPs have a sequential and dependency order, and incorrect combinations result in completely ineffective execution

Insufficient path discovery capability

Online product environments are complex and ever-changing, and traditional code logic discovery will miss many possible execution paths

With the support of large models, we can automatically discover more possibilities

Excessive invalid TTP combinations

Plugins can only be combined randomly, which is very mechanical



Agentic AI Based on Environmental Perception TTP Combination

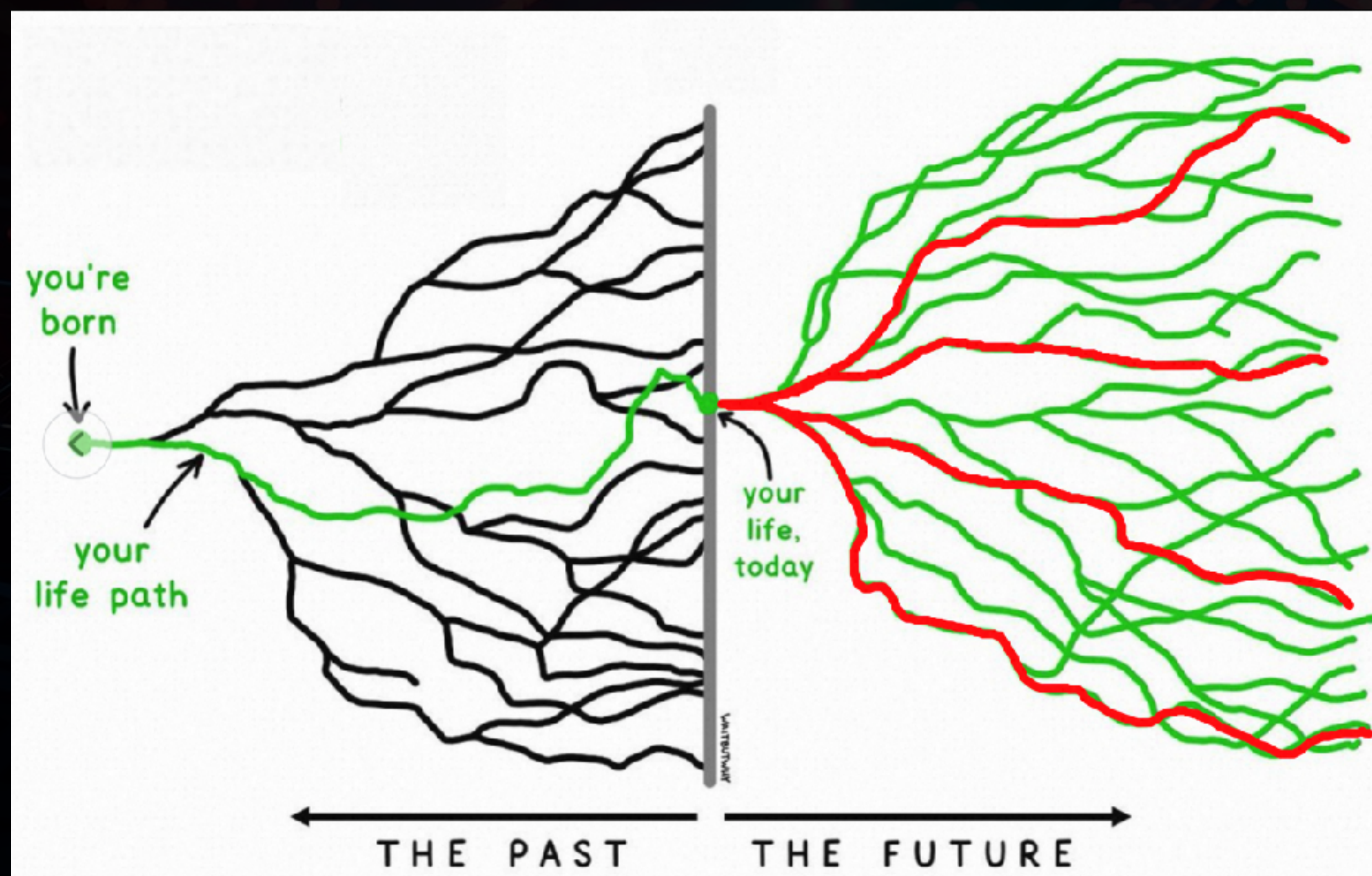
Insufficient path discovery capability

Various if/else statements need to adapt to different scenarios, which lacks flexibility



Large models' superior comprehension and distillation capabilities drive the discovery of more possibilities

We expect to discover more valuable links through automation with limited resources



- Discovering Unknown Network Domain Environments
- Discover assets with missing protection capabilities
- Discover protective capacity shortcomings
-

Automated Attack Drill Scenario Construction Based on Knowledge Graphs

- (1) Filter out drill targets from massive production nodes
- (2) Retrieve all ATT&CK phases and TTPs from the knowledge graph

- (3) Mapping based on LLM-based descriptions

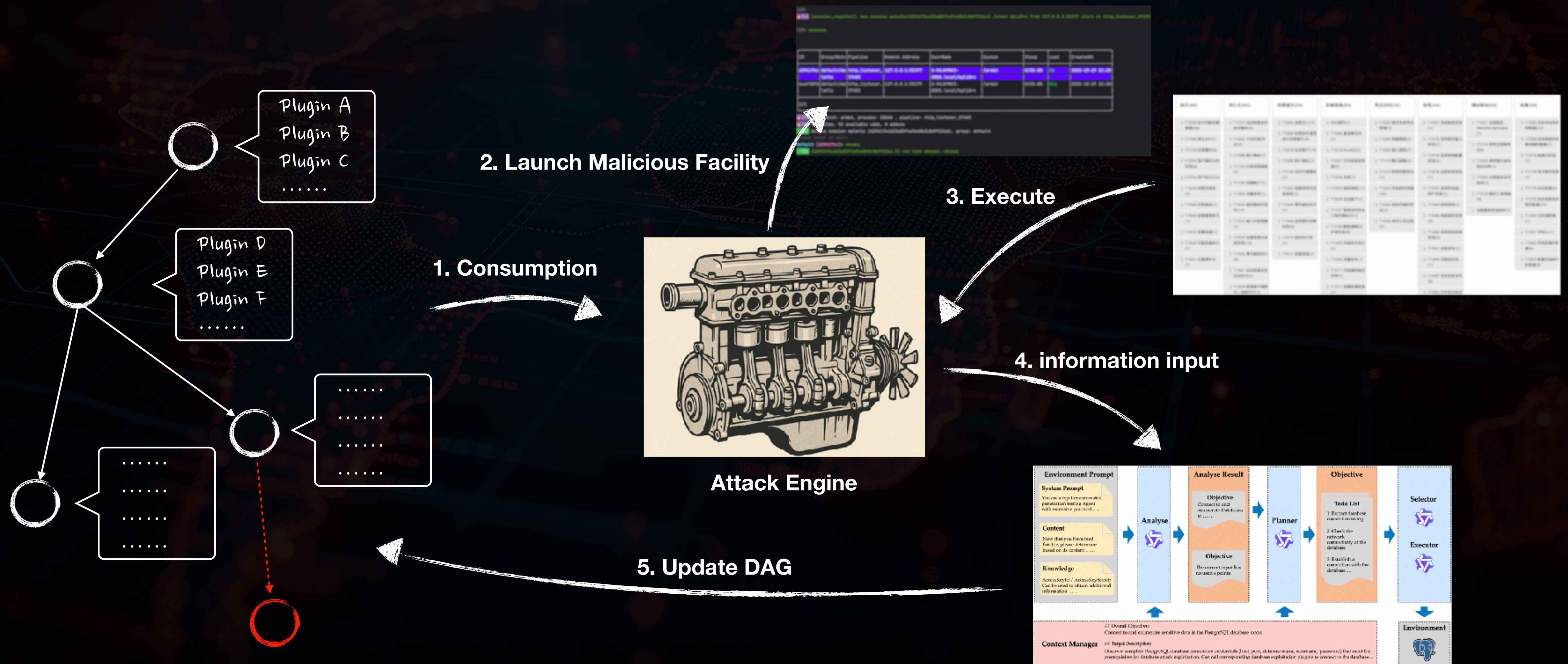
- (4) Randomly combine attack actions for each node based on Kill Chain

Plugin name: xxxx
Phase: collection
Tid: T1005
Description:
xxxxxxxxxxxxxxxx
xxxxxxxxxxxxxxxx
Code:
XXXXXXXXXXXXXXXX

Plugin A
Plugin B
Plugin C
.....

Plugin D
Plugin E
Plugin F
.....

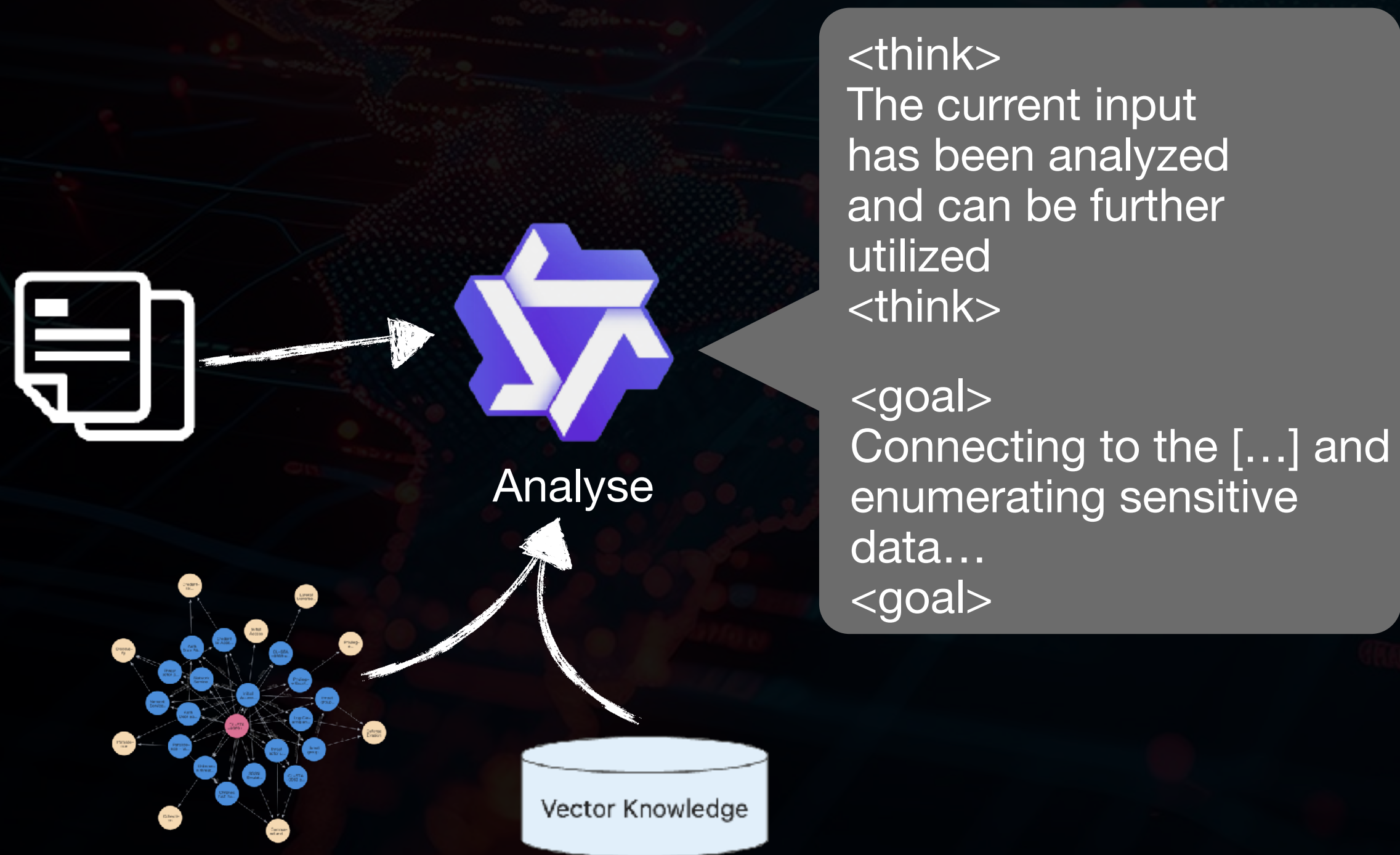
Automated Drills Based on Drill Playbook



Discover more drill paths based on Agentic AI

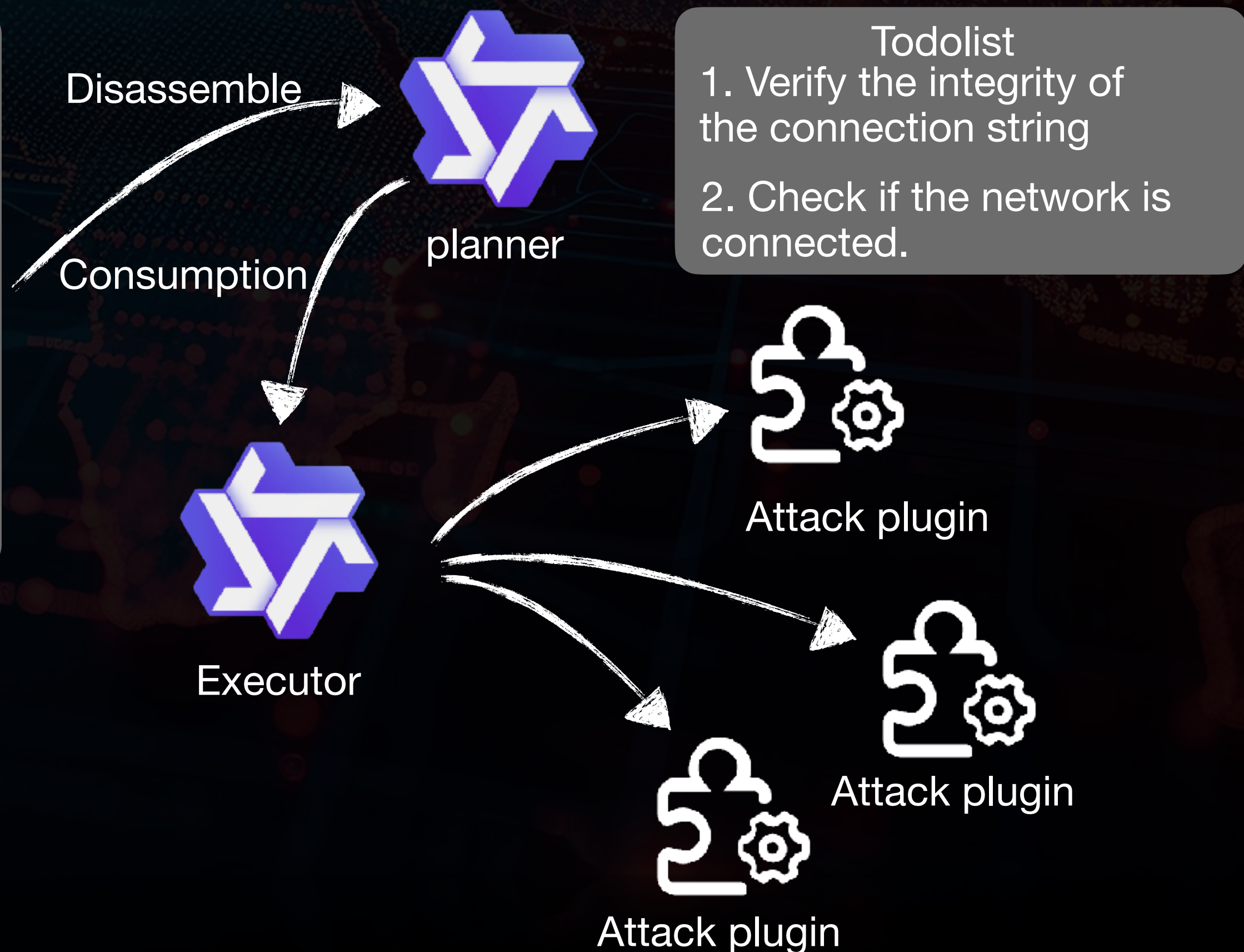
Attack Scenario Discovery

1. Discover more vulnerability scenarios by understanding content output



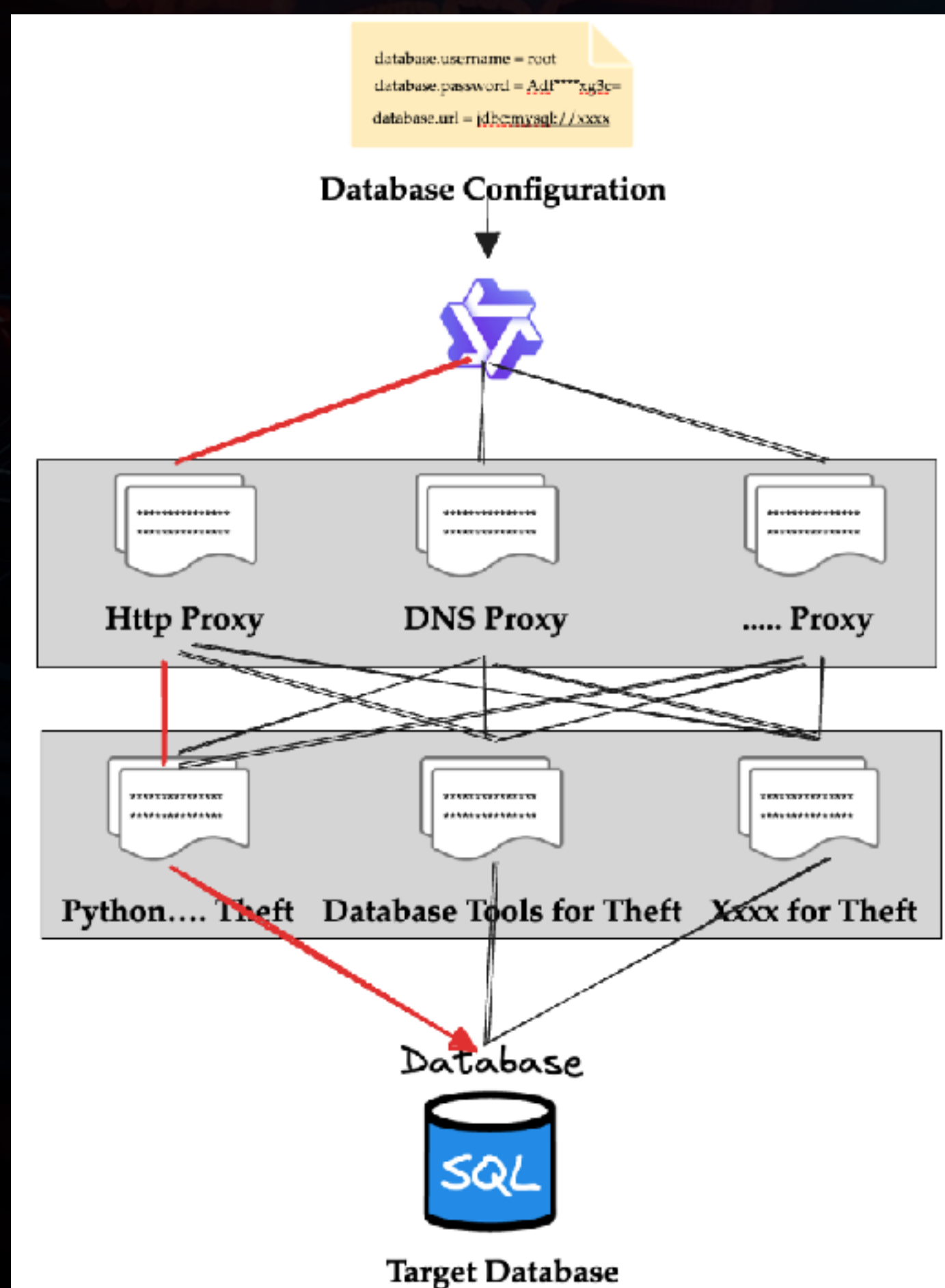
Execution Path Discovery

1. Inspiring LLM decision-making capabilities through methods such as chain-of-thought
2. Enhancing attack technique generalization ability through similar plugin selection



Agentic has insufficient generalization ability in the process of plugin selection and utilization

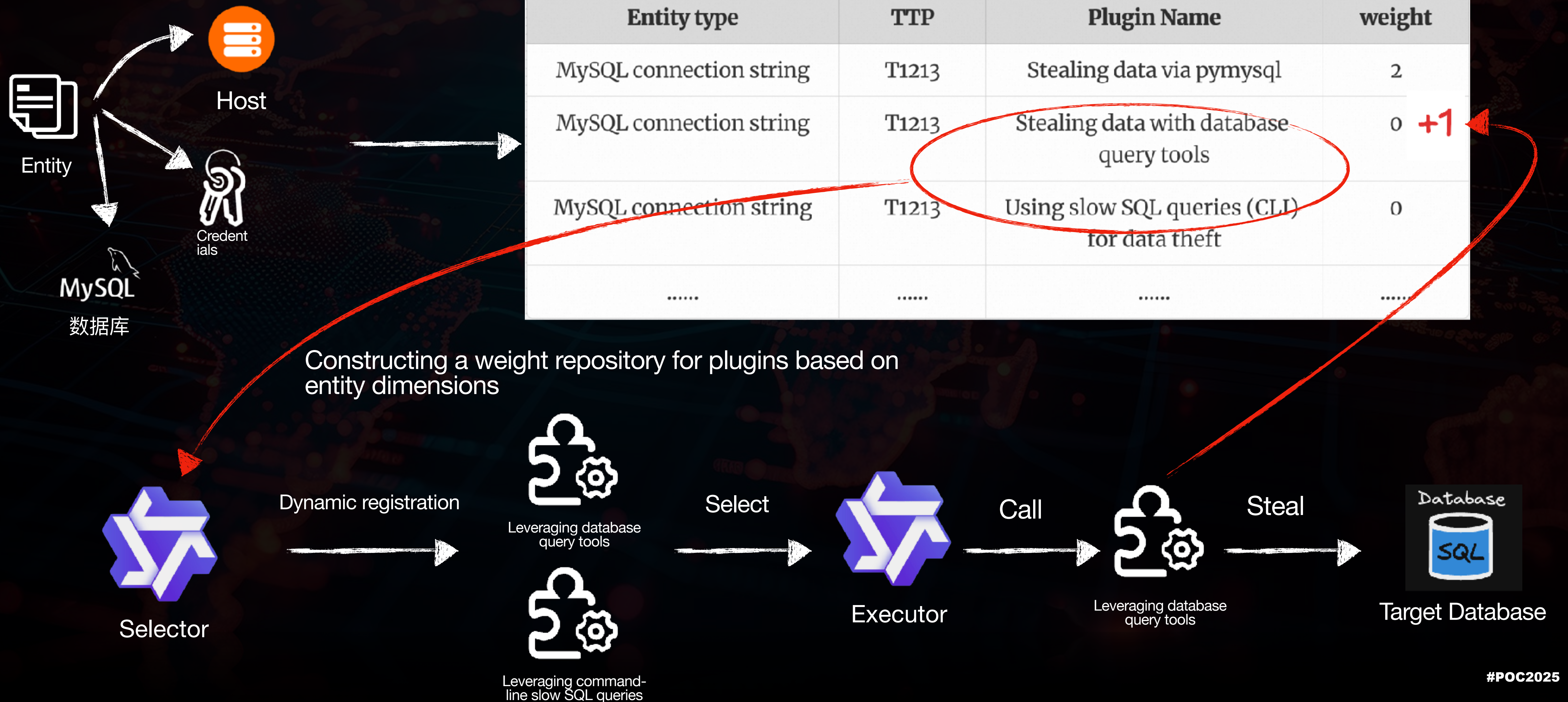
Every time the Agent encounters the same type of input, the tools called each time are similar



Taking the utilization of databases as an example

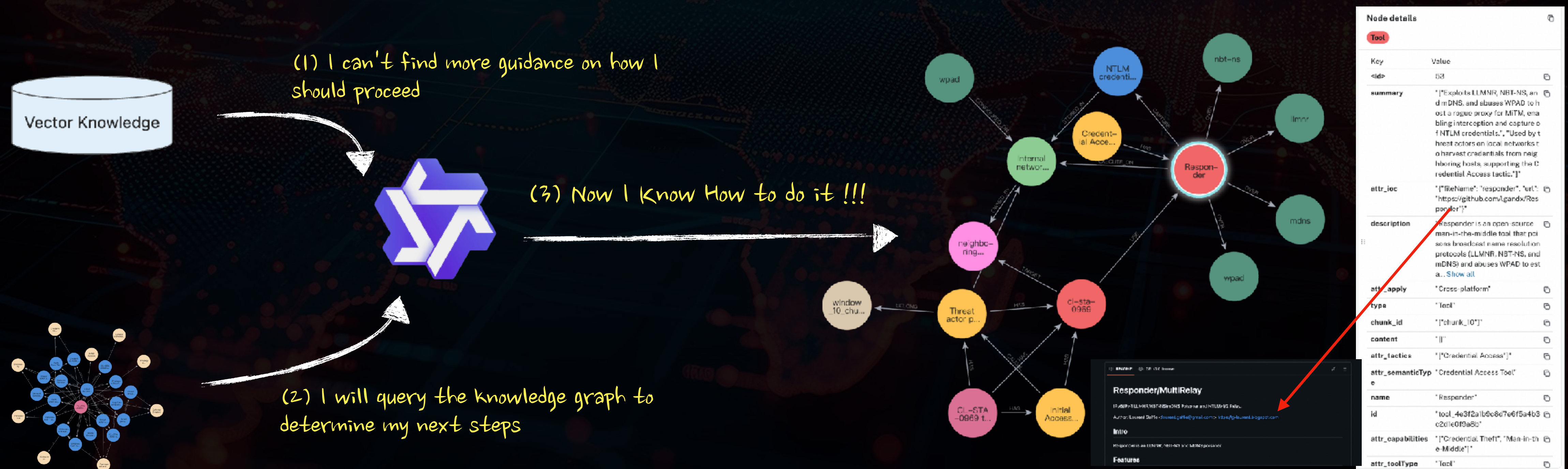
In fact, there are many other paths to choose from, but it's easy to end up selecting duplicate plugins

Building an Entity-Based Plugin Weight Library



Knowledge Graph-Assisted Attack Path Guidance

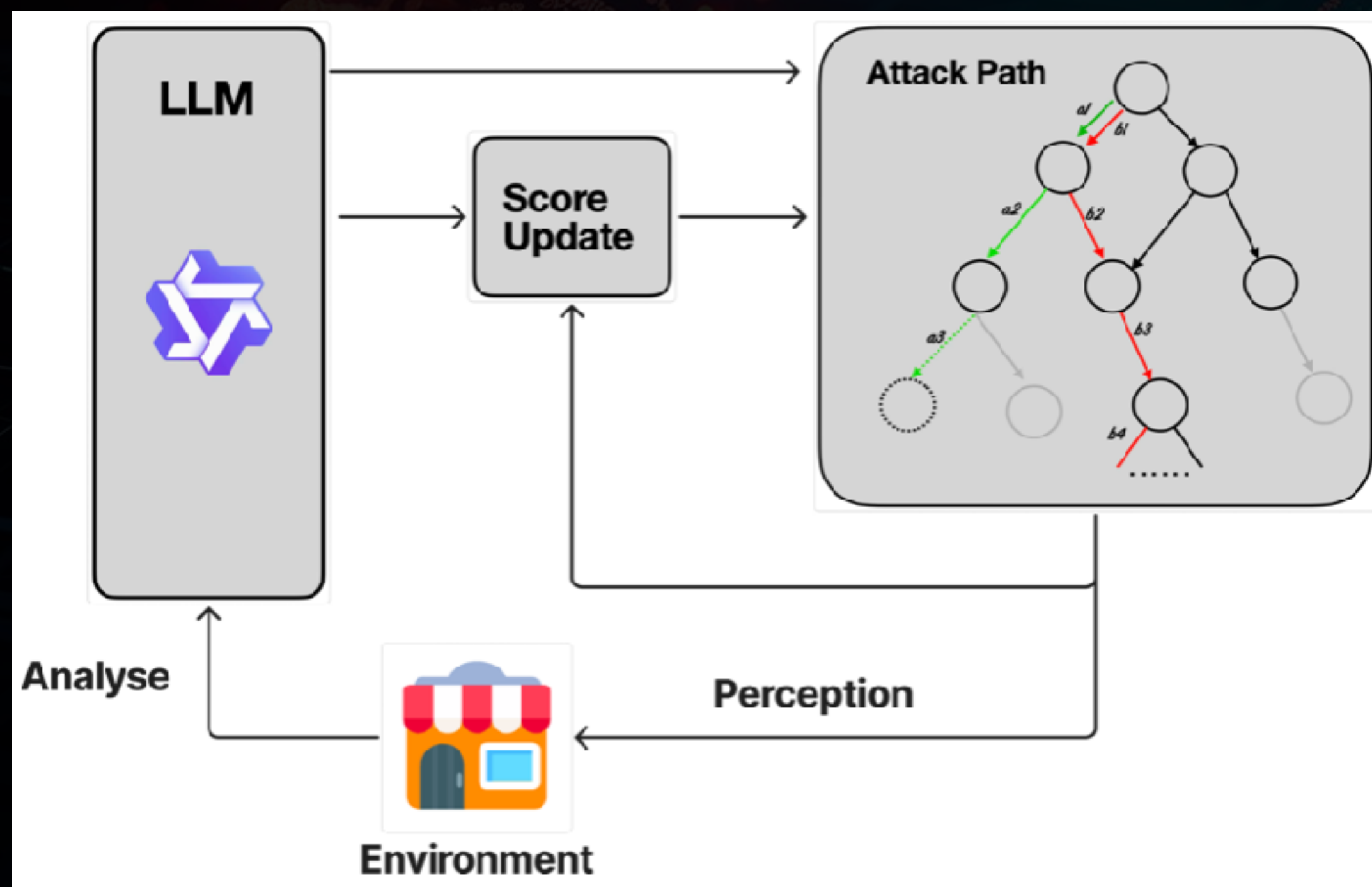
Using a Knowledge Graph to Guide Attack Path Discovery



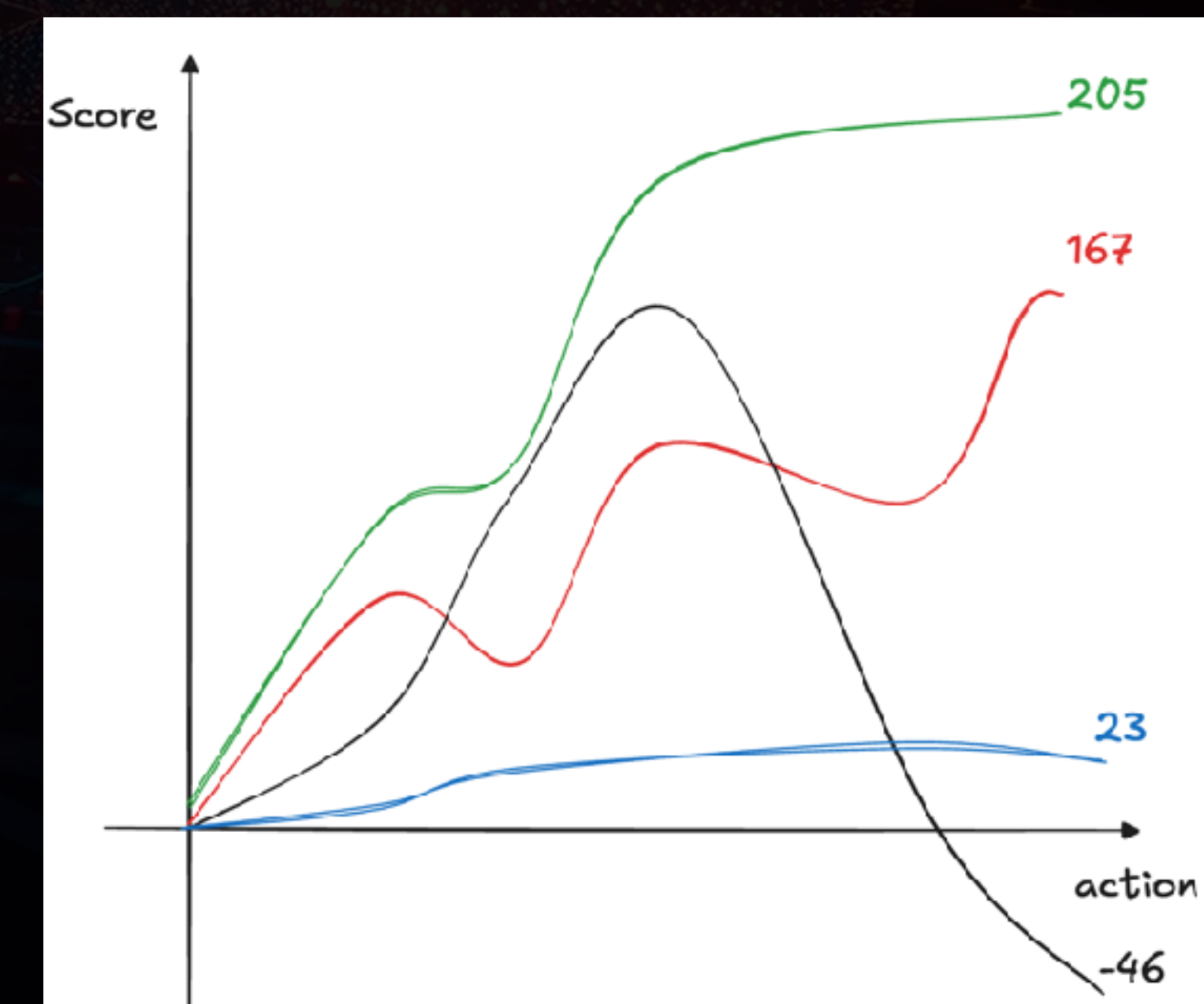
Automated Path Annotation Based on Cumulative Scoring

- Scores are assigned during execution based on each attack action and its outcome.

Intuitive visualization of the path's status based on its score



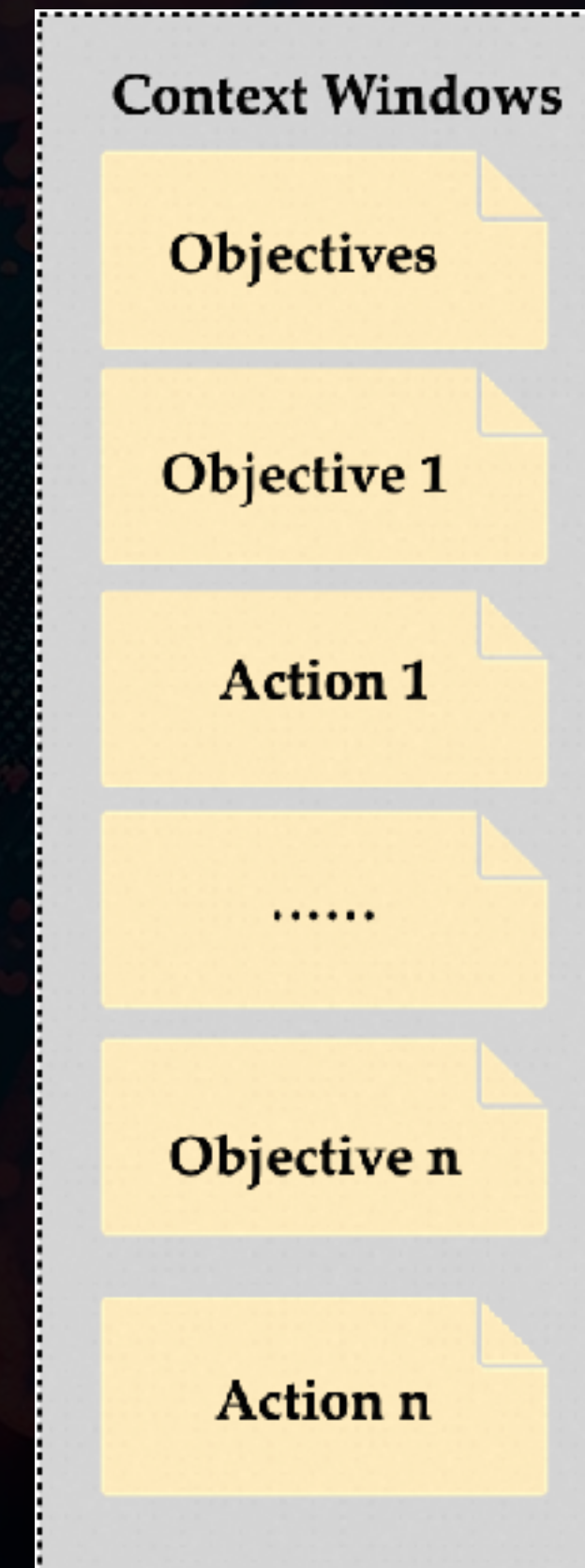
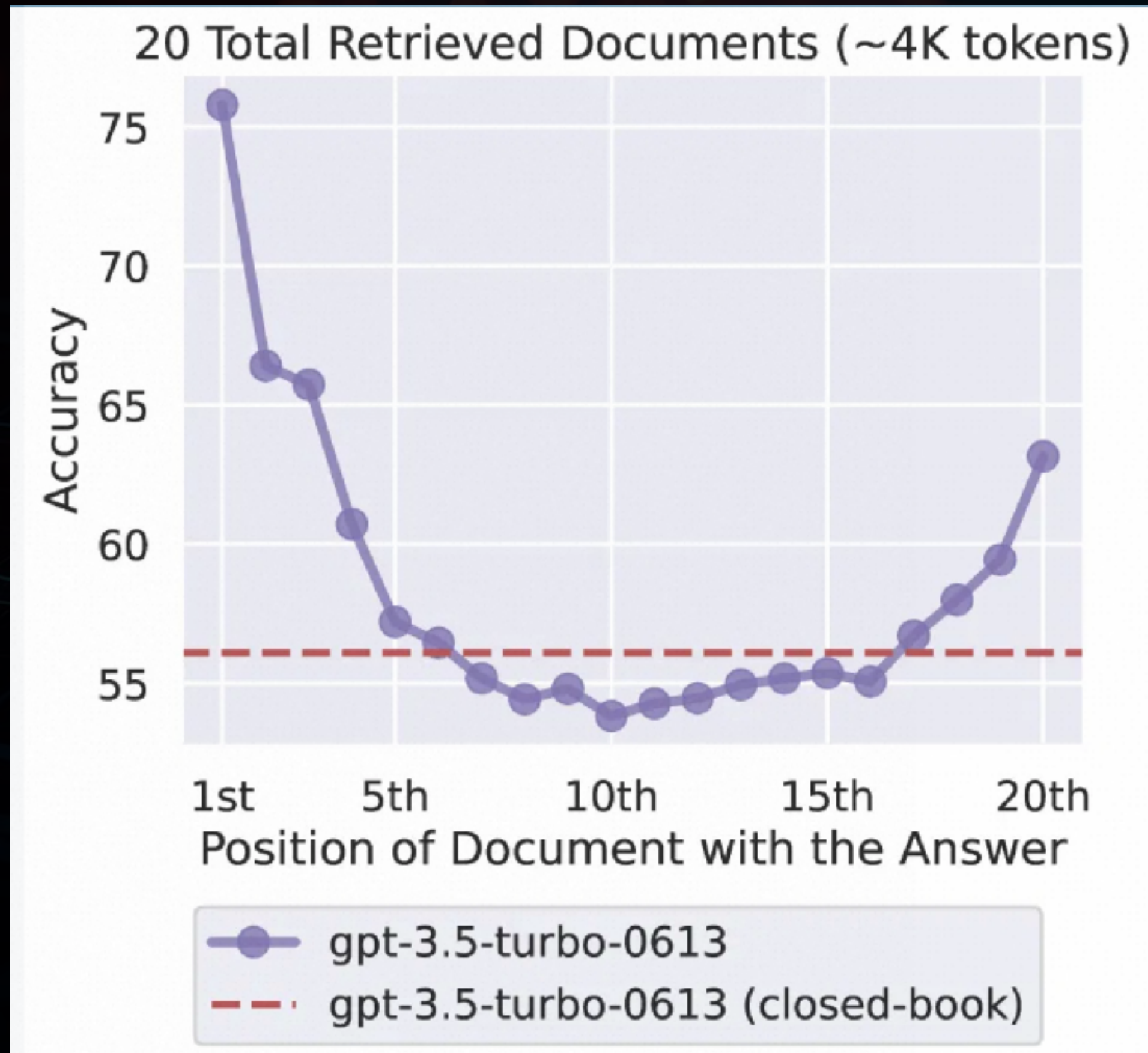
- **Detection Status**
- **Shadow Assets**
- **Missing Logs**
-





Furthermore, the overall stability and success rate during the exercise are paramount.

Prolonged or multi-turn exercises can cause the LLM to drift from its predefined goals and intent.



Information in the middle of the context is easily 'lost' or 'ignored' by the model

Liu et al. (2024). "Lost in the Middle: How Language Models Use Long Contexts"

When critical information is placed in the middle of the context, recall accuracy drops significantly



Simply extending the context window does not guarantee the model will maintain focus on core objectives

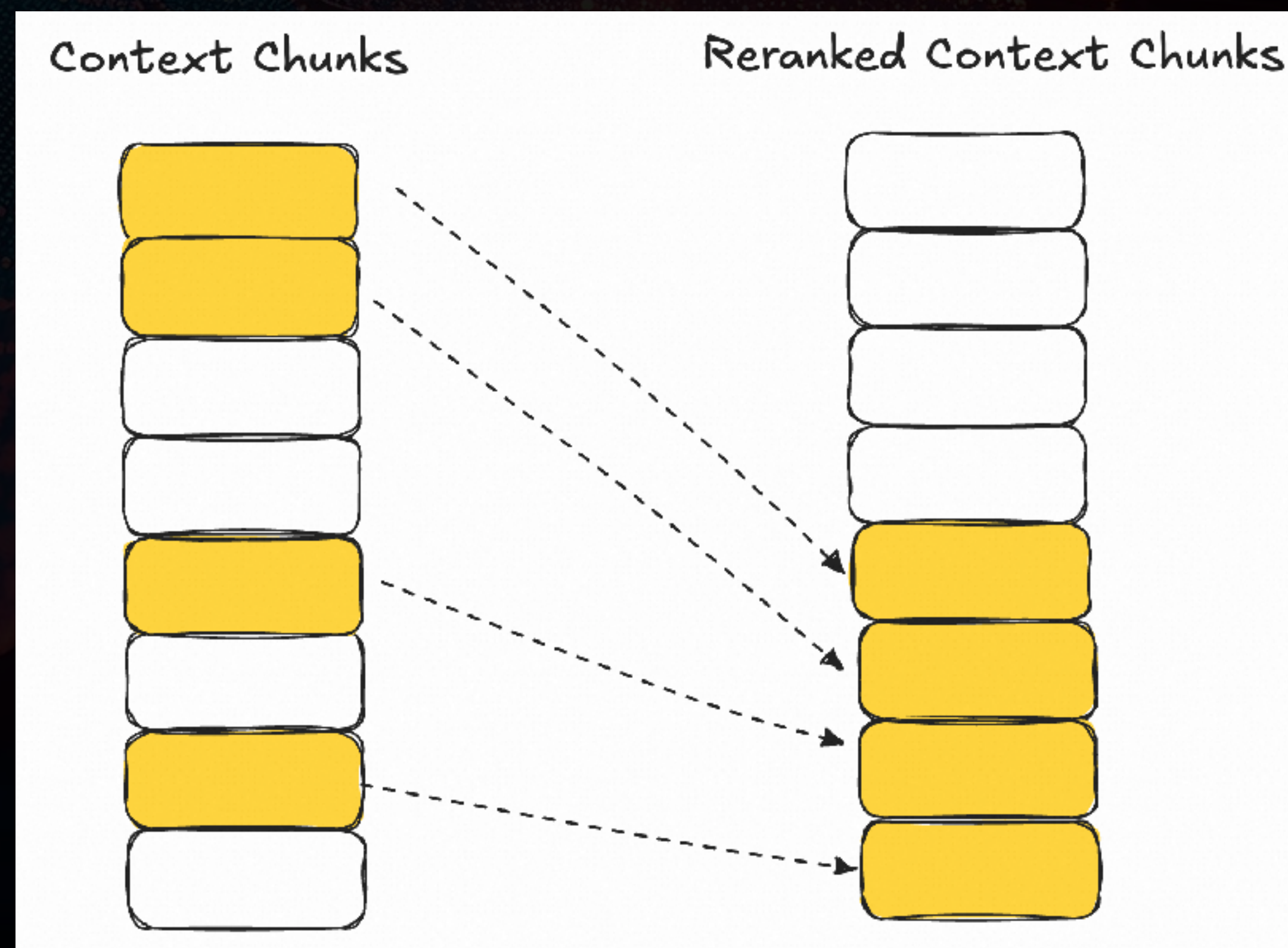
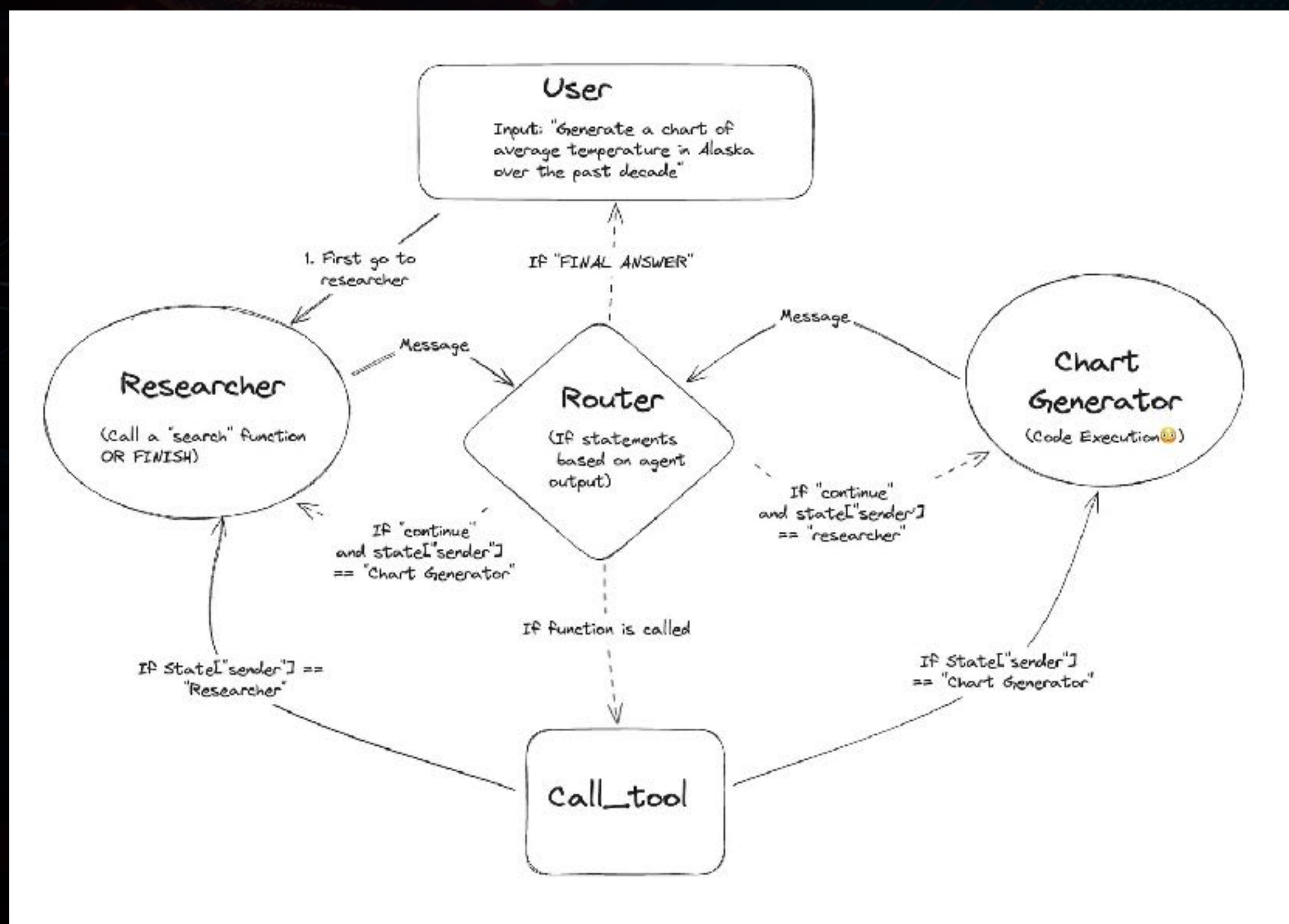
Use two methods to refocus the large model

Graph State “BlackBoard”

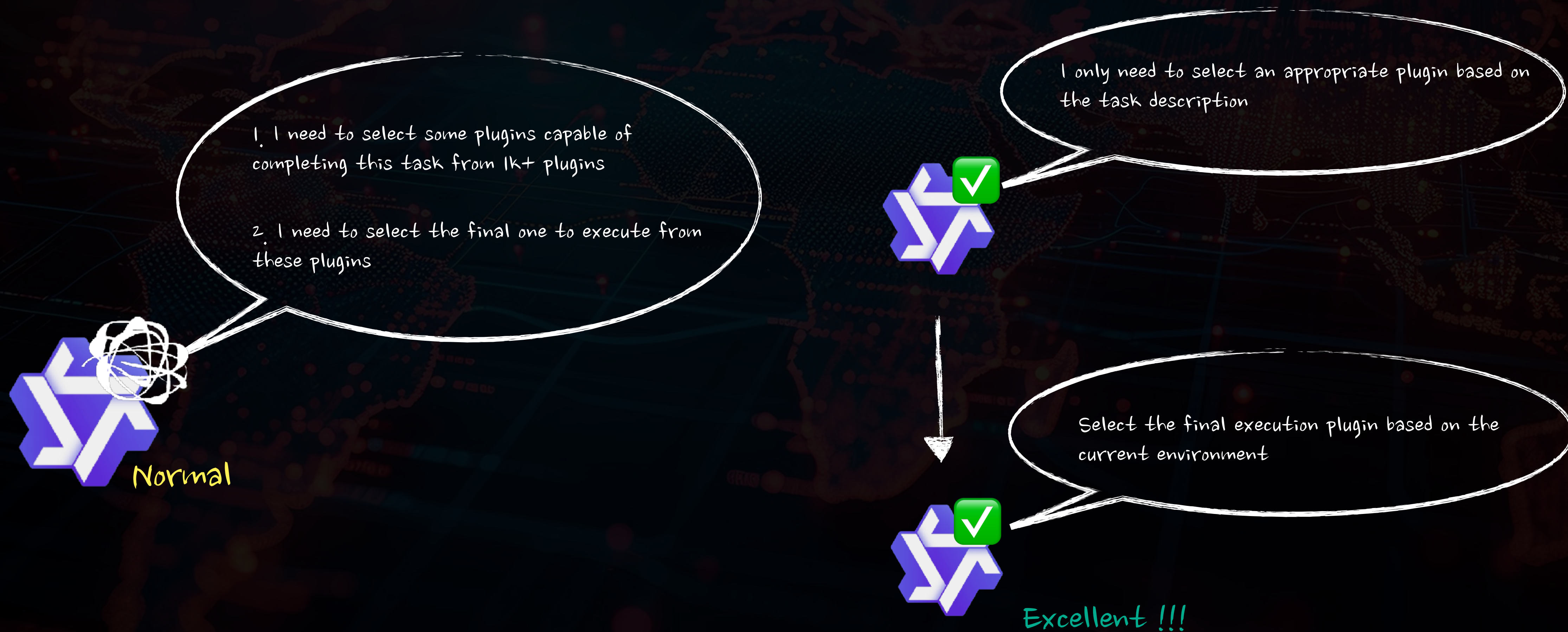
LangGraph-based state management with blackboard pattern for shared memory across agents

Ccontext Recoreding

Repositions key context to attention-focused zones.



First : Agent Layering: Split one agent into multiple agents

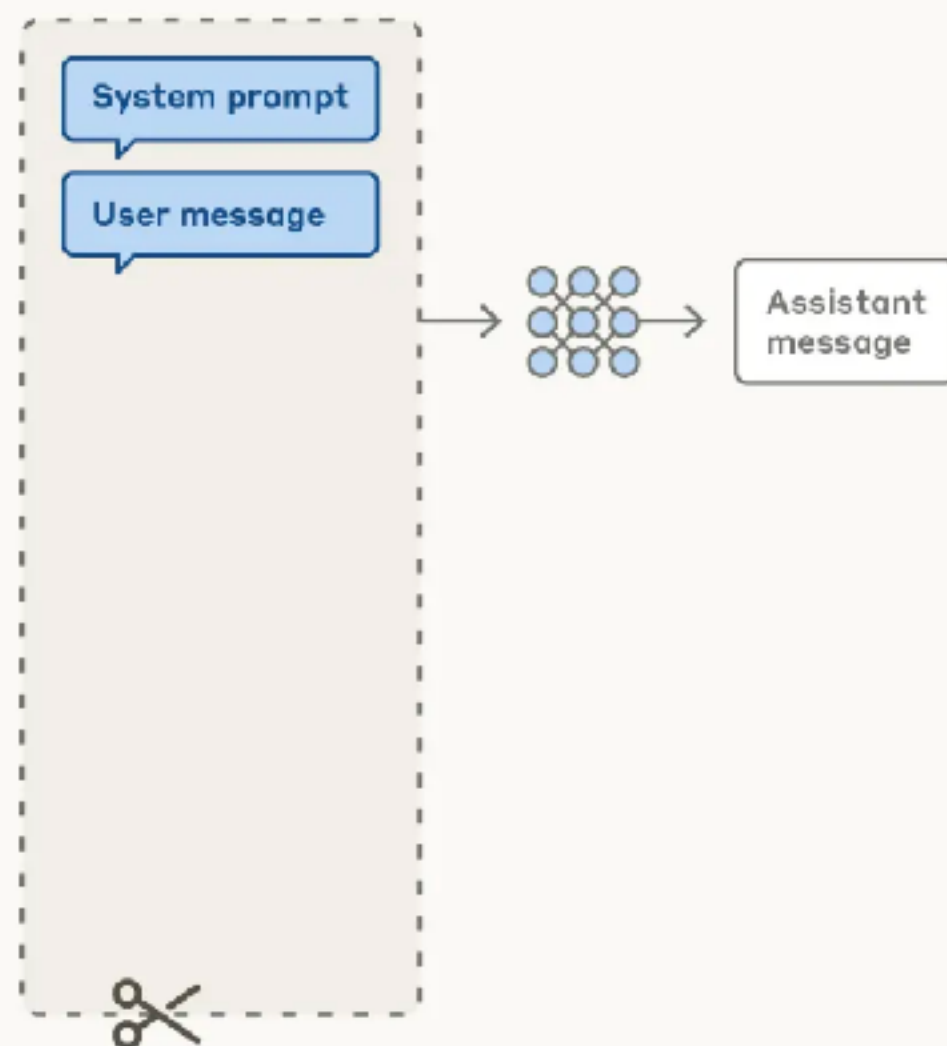


Then: Use File System as a Context

Prompt engineering vs. context engineering

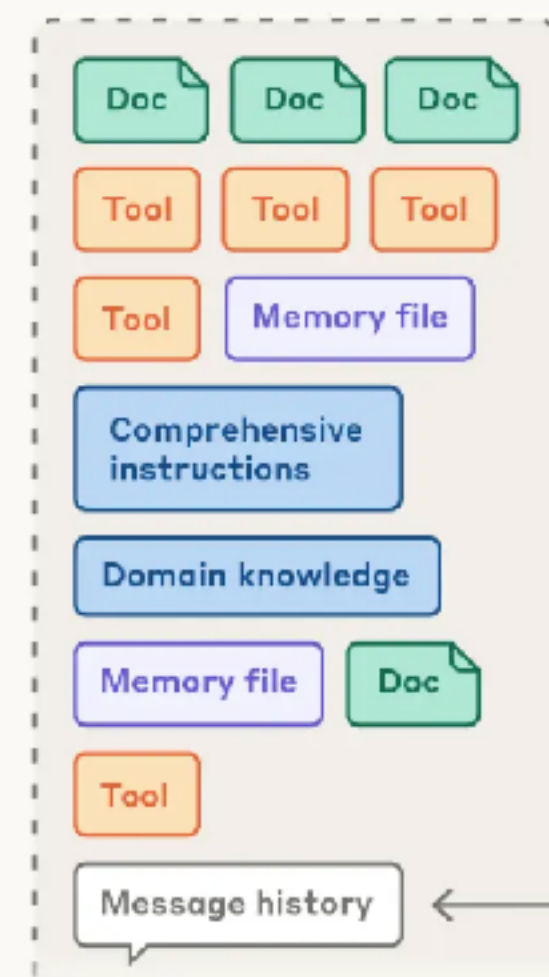
Prompt engineering for single turn queries

Context window



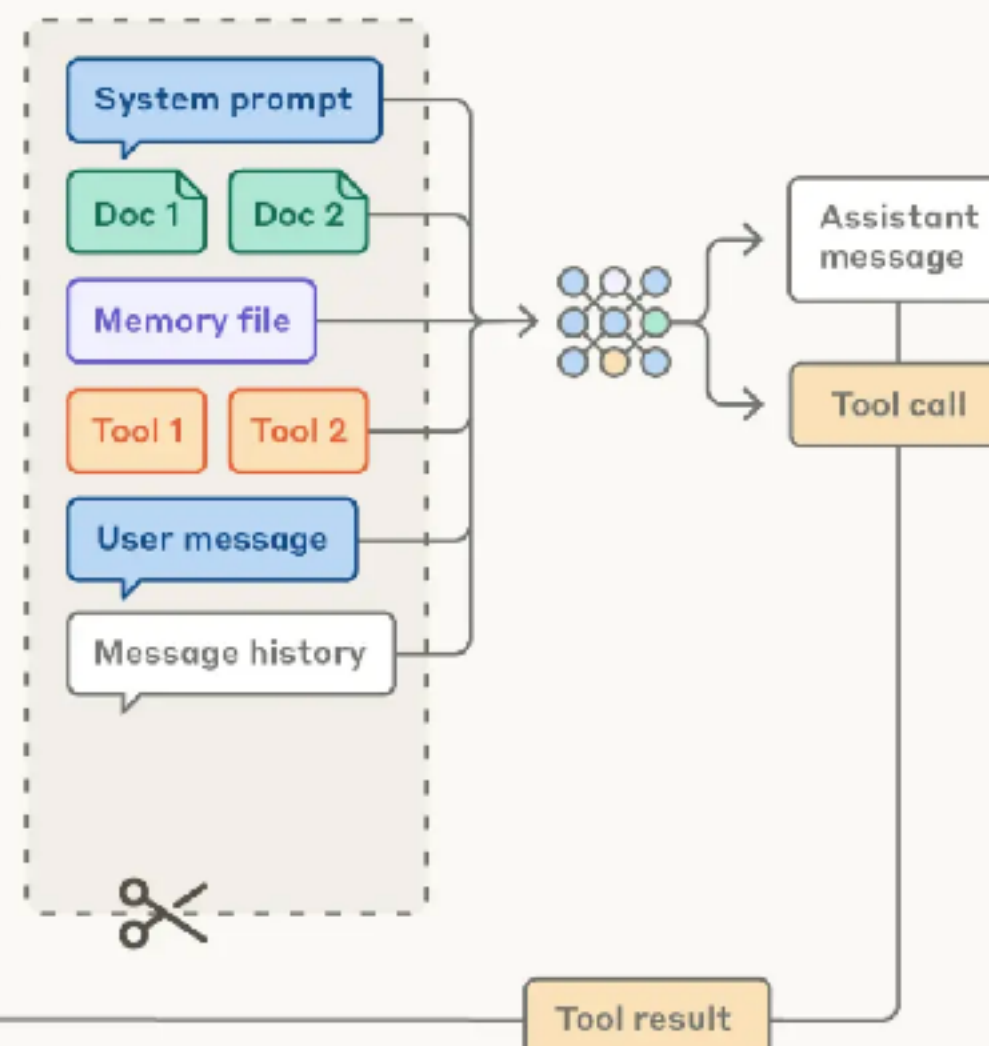
Context engineering for agents

Possible context to give model



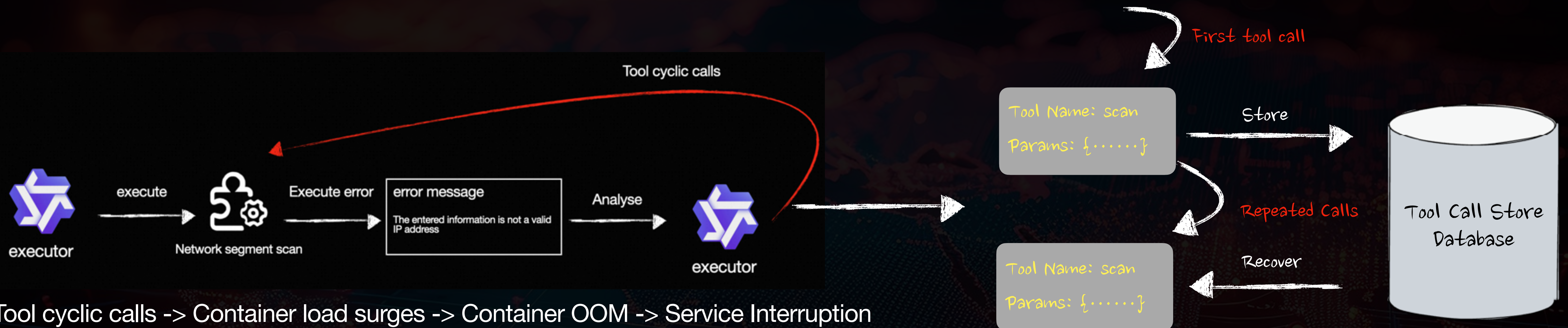
Curation

Context window

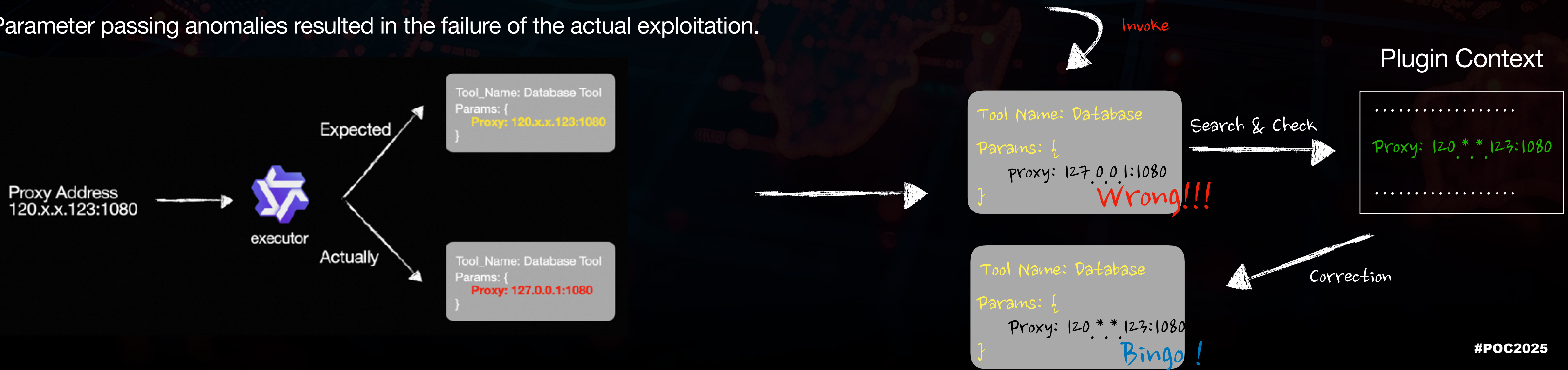


A key strategy to resolve information loss and context limitations is to use the file system as the 'ultimate context'

Some issues with hallucinations



Parameter passing anomalies resulted in the failure of the actual exploitation.



Demo

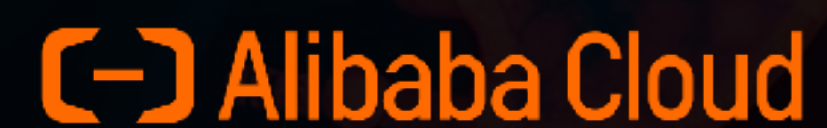
Coming

Summary

Coming



Thanks



ZUOTONG FENG & JIALIANG WU



Context Length Alone Hurts LLM Performance Despite Perfect Retrieval - <https://arxiv.org/html/2510.05381v1>

KAG: Boosting LLMs in Professional Domains via Knowledge Augmented Generation - <https://arxiv.org/pdf/2409.13731>