From Mass Extraction To Ethical Triage: Multi-agent AI For Privacy-preserving Computer Forensics

Md Tamjid Hossain, CISSP, PhD[1] and Shafkat Islam, PhD[2]

Texas A&M University-San Antonio[1], Purdue University Northwest[2]
Email: mhossain@tamusa.edu[1], islam59@pnw.edu[2]

Introduction

Computer forensics investigations often relies on mass evidence extraction from disks, mobile devices, logs, and cloud sources. Although this helps avoid missing key evidence, it exposes large amounts of private, non-case data to investigators [1]. To address this challenge and facilitate data minimization during investigations, we explore an ethics-by-design forensic triage approach using Qwen2.5-7B-Instruct, which achieved 70% accuracy in our initial experiments.

Research Question

- **RQ1:** How can AI-assisted triage support data minimization in computer forensic investigations by reducing unnecessary exposure to non-case-related data?
- **RQ2:** Can such a framework preserve relevant evidence while ensuring controlled, authorized access to sealed data?

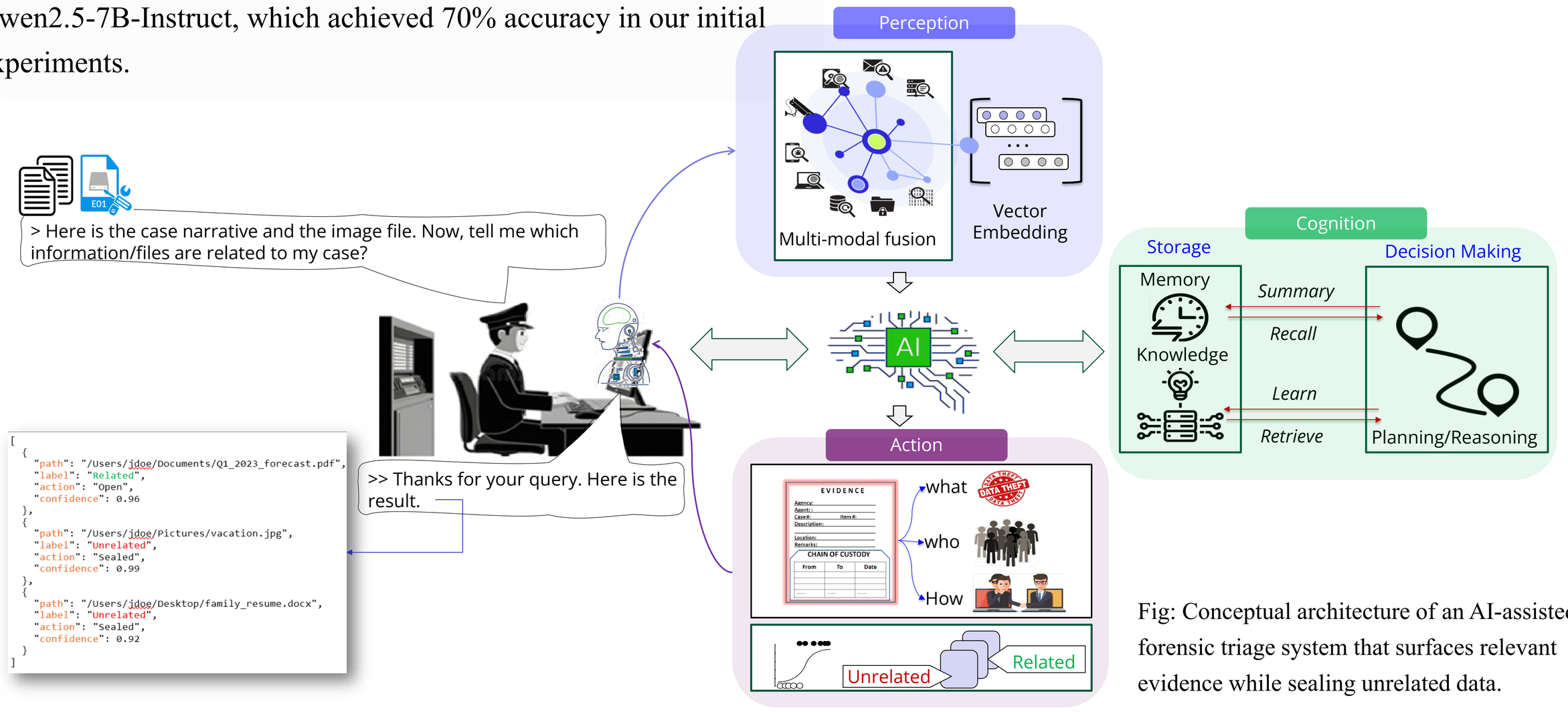

Fig: Conceptual architecture of an AI-assisted forensic triage system that surfaces relevant evidence while sealing unrelated data.

Proposed Method/Workflow

[Case Narrative]

Timeline • Entities • Indicators

↓

[Context-Aware LLM Agent]

Assesses artifact relevance

↓

[Artifact Classification]

Relevant • Non-Relevant • Uncertain

↓ ↓ ↓

Visible Sealed Human Review

Preliminary Results

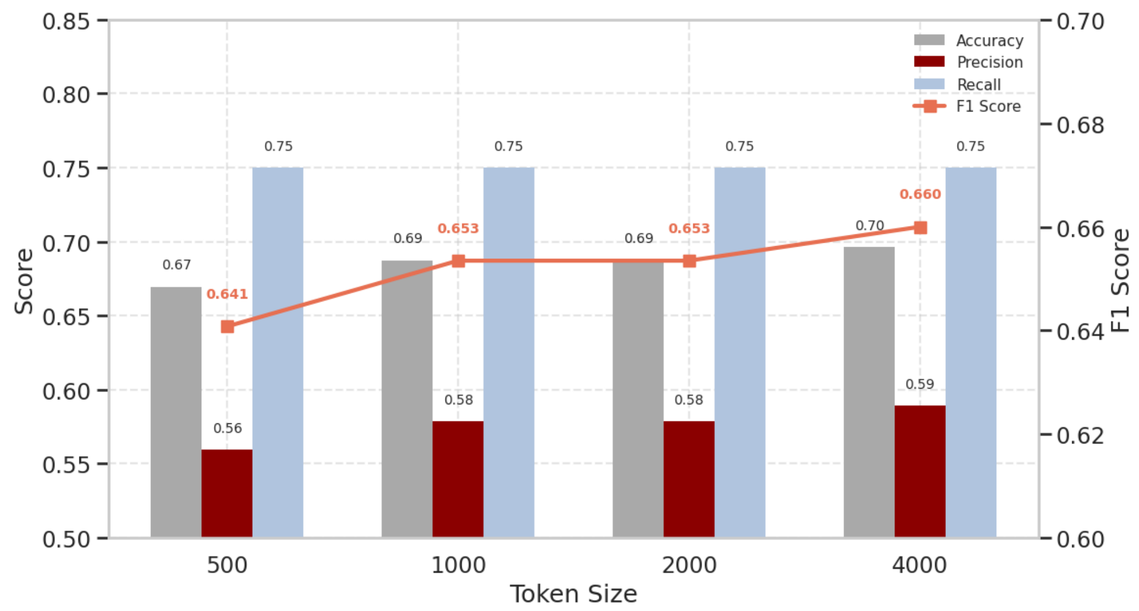

Fig: Performance across token size (Qwen2.5-7B-Instruct)

Sample Output


```
{
  "File Path": "scripts_n_stuff/passwords_backup.txt",
  "MIME": "text/plain",
  "Specialist": "txt",
  "LLM-generated Flag": "Relevant",
  "Ground truth Flag": "Relevant",
  "Reason": "```json\n{\n  \"relevant\": true,\n
  \"reason\": \"The file contains sensitive information
  such as usernames and passwords for various systems,
  which could be crucial for understanding potential
  security breaches or unauthorized access.\"\n}\n,
  "File Size": 417,
  "Analysis Time": 80.1878,
},
```


Discussion: Potential Challenges

- Relevant or exculpatory evidence must not be incorrectly hidden; low-confidence cases should require human review
- The framework must handle diverse forensic artifacts

Conclusion

We aim to move computer forensics from reactive mass review toward privacy-aware, context-guided investigation. By combining AI-based triage with auditable access control, the approach seeks to reduce unnecessary exposure to private data while preserving evidentiary integrity.

Reference

[1] Scientific Working Group on Digital Evidence. (2023). *Considerations for required minimization of digital evidence seizure*.