Artificial Intelligence in Digital Forensics: Augmented Analysis and Emerging Evidence

Albi Marini, Danny Trainor

Motivation and Goals

-Assess this new frontier of digital forensics as a result of the technological wave of AI

-The project goal was to have a paper that would assess this motivation in some way

-A paper that was both **written** and **revised** over the weeks in which the project was to be worked on. (COMPLETED)

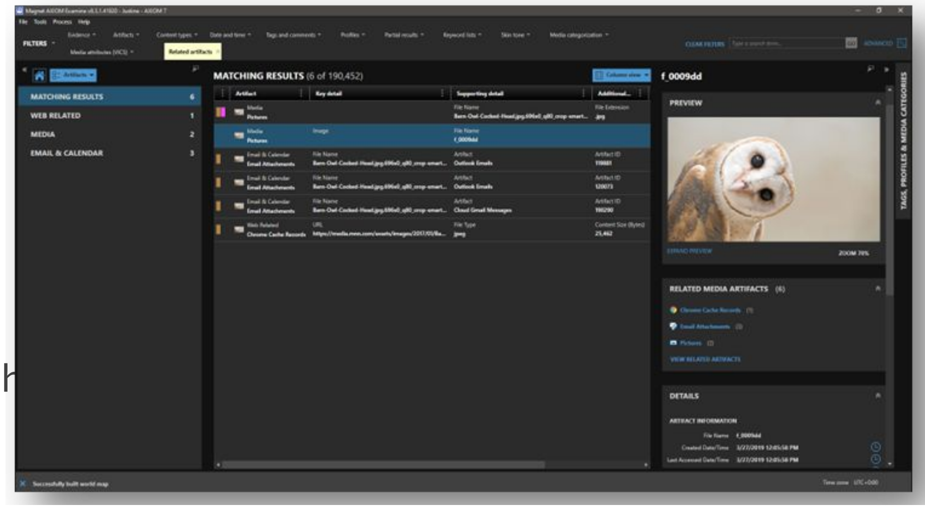
AI in Digital Forensics Overview

-Enhances forensic (or anti-forensic!) tools and techniques across all types of digital evidence

-Creates new types of evidence we can look at

Forensic Enhancement: Magnet AXIOM

- Magnet AXIOM is an example of a tool enhanced by AI Developments
  - Extremely popular forensics discovery tool in the mobile device space especially in federal entities
- Since 2018, AXIOM has integrated ***Magnet.AI*** into their software, with latest versions using ***Magnet Copilot***
  - **Offline data recognition and media authenticity capabilities that automatically gather evidence from device disk images**



MATCHING RESULTS (6 of 190,452)

Artifact	Key detail	Supporting detail	Additional...
Media Pictures		File Name Barn-Owl-Cocked-Head.jpg.696x0_q80_crop-smart...	File Extension .jpg
Media Pictures	Image	File Name f_0009dd	
Email & Calendar Email Attachments	File Name Barn-Owl-Cocked-Head.jpg.696x0_q80_crop-smart...	Artifact Outlook Emails	Artifact ID 119881
Email & Calendar Email Attachments	File Name Barn-Owl-Cocked-Head.jpg.696x0_q80_crop-smart...	Artifact Outlook Emails	Artifact ID 120073
Email & Calendar Email Attachments	File Name Barn-Owl-Cocked-Head.jpg.696x0_q80_crop-smart...	Artifact Cloud Gmail Messages	Artifact ID 190290
Web Related Chrome Cache Records	URL https://media.mnn.com/assets/images/2017/05/Ba...	File Type jpeg	Content Size (Bytes) 25,462

f_0009dd

DETAILS

ARTIFACT INFORMATION

File Name	f_0009dd
Created Date/Time	3/27/2019 12:05:58 PM
Last Accessed Date/Time	3/27/2019 12:05:58 PM

Successfully built world map

Figure 1: Example of *Magnet Copilot* automatically identifying and categorizing evidence from a device image in the *Artifact Explorer* of Magnet AXIOM

Cellebrite Pathfinder

- Cellebrite tools have been used since 1999 in law enforcement agencies for data collection/extraction, collaboration & data points
- Since 2021, their newest tool, ***Pathfinder***, has been touted as platform to “**automate the ingestion and analysis of digital evidence**”
  - Provides a **graph-like visualization** of key investigative points tracing activity from the victim and suspect
  - Automatically links data points to GPS data from photos, cell towers, Wi-Fi hotspots to construct a full incident report

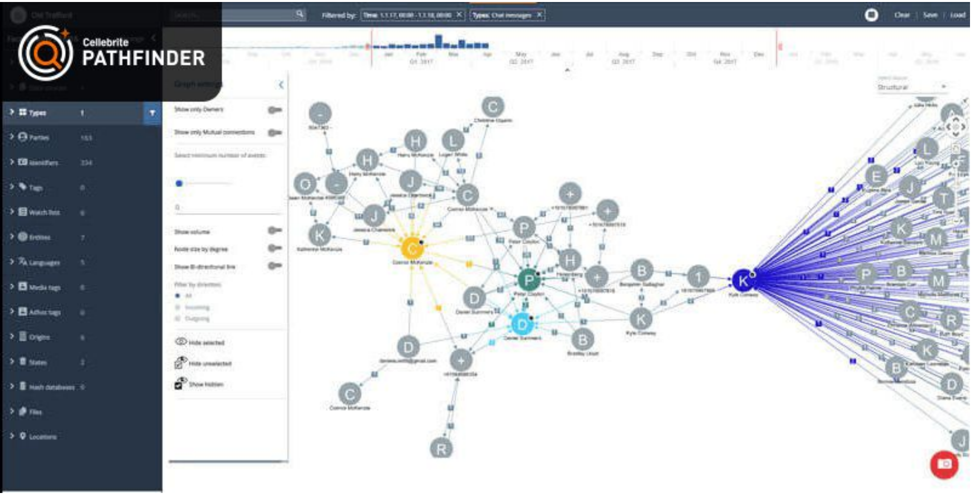


Figure 2: Example of Cellebrite Pathfinder's graph interface that automatically collects and draws connections between data points from multiple data sources to create a case-web of involved parties

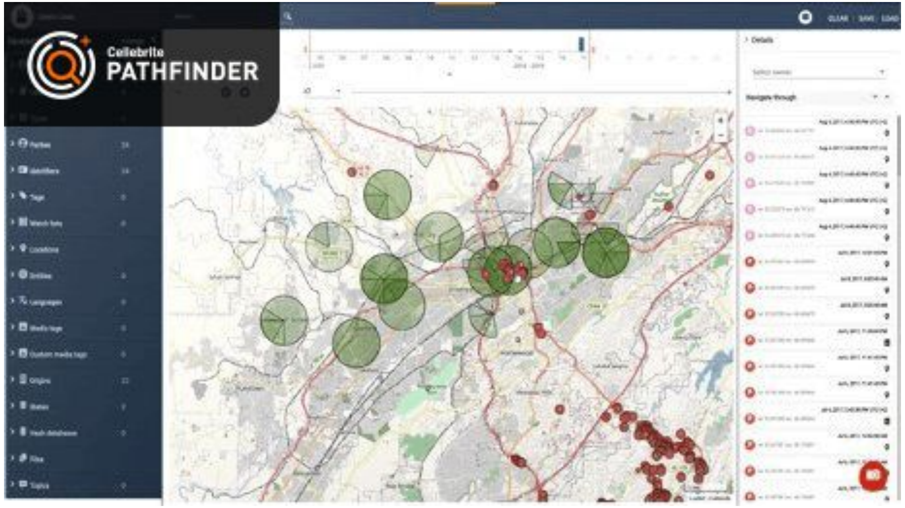
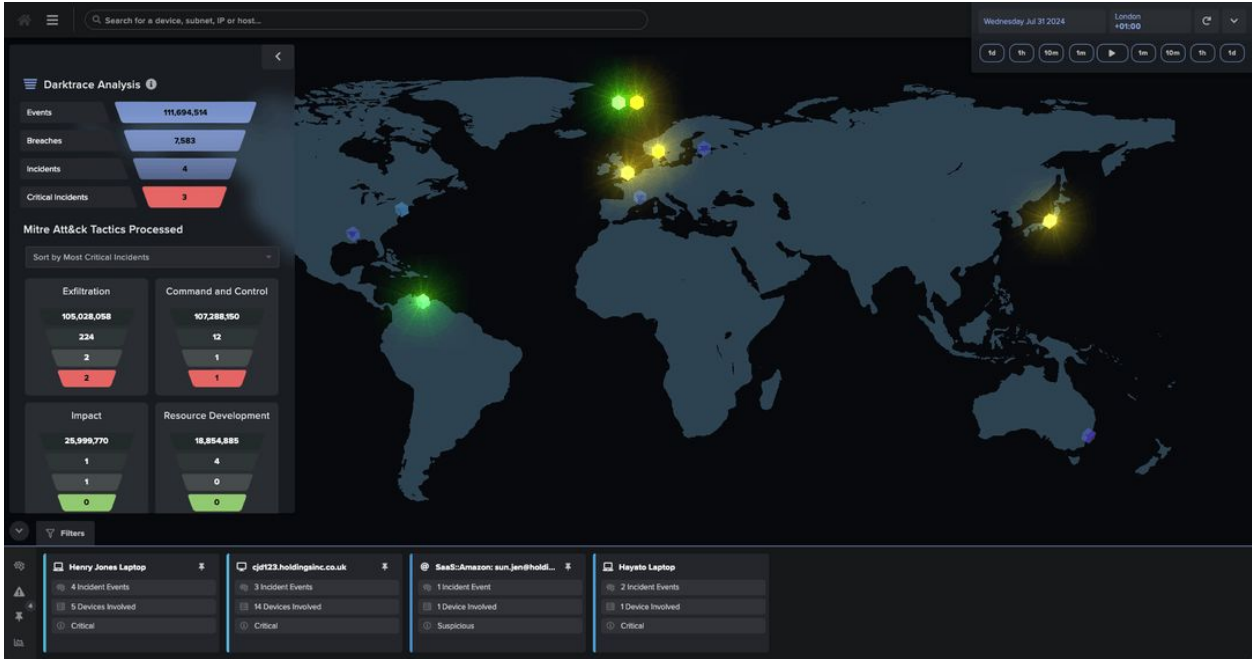


Figure 3: Automatic geo-location identification of where investigative data points occurred in Cellebrite Pathfinder

Darktrace

- **AI cybersecurity platform** that offers network forensic capabilities
  - Newer company, example of AI-enabled startup
- **Reinforcement learning powered “self-learning AI transformer-based language model”** are leveraged in incident detection and response
  - They provide a globe visualization for network incidents, picturing each device, its connections, and the path an attack took through the nodes of a network
- ***Cyber AI* automatically triages, interprets, and provides a report of the incident**
  - Eg: ransomware outbreak: trace the spread across the network node to provide an investigator with a list of affected machines and a complete attack timeline



Darktrace Analysis

Events	111,694,514
Breaches	7,583
Incidents	4
Critical Incidents	3

Mitre Att&ck Tactics Processed

Sort by Most Critical Incidents

Exfiltration	Command and Control
105,028,058	107,288,150
224	12
2	1
2	1

Impact	Resource Development
25,999,770	18,854,885
1	4
1	0
0	0

Henry Jones Laptop	cjd123.holdingsinc.co.uk	SaaS::Amazon: sun.jen@holdi...	Hayato Laptop
4 Incident Events	3 Incident Events	1 Incident Event	2 Incident Events
5 Devices Involved	14 Devices Involved	1 Device Involved	1 Device Involved
Critical	Critical	Suspicious	Critical

Figure 4: Example of the Darktrace network tool visualization interface that provides a map of network nodes, as well as their state and log of previous incidents

Large-Language Model Assisted Analysis

- Cornell University in 2025 (Yin et al., 2025): LLMs have become widely used across the field across all steps of the investigative process
  - **ChatGPT-3.5, 4.5** and **Google Gemini** “process enormous volumes of text-based evidence” (chat logs, documents, emails) much faster than humans
  - Categorize text, recognize patterns, and connect names with timestamps or address to “identify complex interrelationships”
- Example of **LLM-driven Mobile Evidence Contextual Analysis:** feeding mobile chat logs into LLM to perform contextual analysis
  - Find criminal intent through pattern identification or coded language in messages automatically, previously very challenging without AI

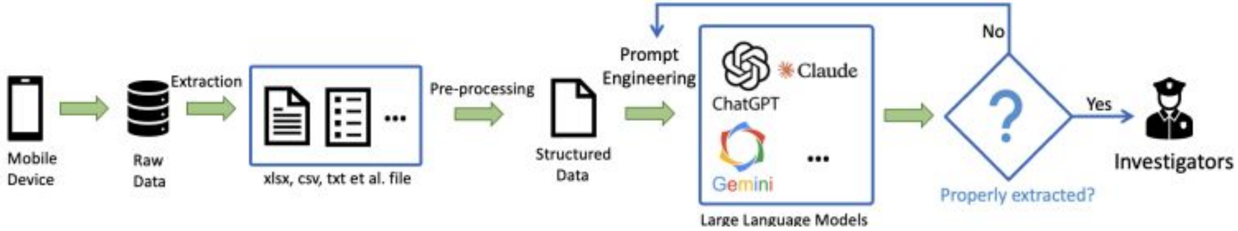


Figure 5: Overview of LLM-driven Mobile Evidence Contextual Analysis Framework

volGPT

- Research example of AI-assisted volatile memory forensics (Oh et al., 2024)
    - “first prompt-based large language model for memory forensics, providing analysts with triage automation and explanations for triage reasons” using GPT-like reasoning models as a base
    - Proved very effective in detecting ransomware on infected memory dumps of Windows systems (accuracy 87-99%, avoids false positives)
        - Can detect techniques like process masquerading
- Takes a memory image, filter down the list of suspicious processes, and provide text explanations for each software as to why it was flagged as suspicious

Pros and Cons of AI as a tool in Digital Forensics

Pros:

- Frees up time in an investigation, especially in data discovery and analysis allowing analysts to focus on interpretation of evidence

Cons:

- Cornell University 2024 (Sanna et al., 2024), these AI tools are not robust - do not guarantee perfect accuracy in their findings, overlook evidence and can give false insight, ie misclassifying deep fakes as the real person
  - Inherit drawback of AI tools as the technology is still improving
  - Risks are dangerous in forensics: all evidence presented in court must be without faults
- Errors due to bias in the data used to train the models used in these tools
  - Eg AI tends to classify conversation from some groups as malicious when its not
- Smart adversaries can alter images and texts to make them undetectable by AI (“anti-forensics”)
- Black-box AI models do not disclose how they work make them difficult to trust/verify

LLM Invocation Logs as Digital Forensic Evidence

- LLMs are susceptible to manipulation attacks, and can be vessels for malicious intent
- Logs are real-time events that are collected instantaneously
- Invocation logs can tell us what the client/web service the client was using was prompting the LLM. (Gemini, GPT,etc.)

Generative Adversarial Networks, Deepfakes, and Digital Media

- GAN: Machine learning frameworks that aim to create realistic, synthetic data
  - Deepfake Technology
- Can be used to manipulate images, potentially exploiting forensic classifiers