

AI-Powered Digital Forensics and Evidence Analysis

^[1] T.N. Sriranjani, ^[2] B. Aakashsha Yadav, ^[3] B. Divya Sri, ^[4] C.H. Godha Tejaswini, ^[5] C. H. Naga Sudha, ^[6] K. Murudula

^[1] ^[2] ^[3] ^[4] ^[5] ^[6] Department of Computer Science and Engineering, Keshav Memorial Institute of Technology, Hyderabad, Telangana, India

Corresponding Author Email: ^[1] Sriranjani.kmit@gmail.com, ^[2] baakashsha18@gmail.com, ^[6] murudulakannan2@gmail.com

Abstract— Investigating large volumes of digital evidence from forensic reports such as Unified Forensic Data Reports (UFDR) is often time-consuming and complex when performed manually. This study proposes an AI-Powered Digital Forensics Assistant that leverages Artificial Intelligence (AI) and Natural Language Processing (NLP) to simplify and accelerate digital investigations. The system automatically processes forensic datasets to extract key information including messages, call logs, contacts, and media files. It supports multilingual detection and translation, enabling investigators to analyze diverse communication data efficiently. Additionally, users can perform natural language queries such as identifying conversations related to financial or suspicious activities. The system detects patterns, keywords, and behavioral trends to highlight potential evidence. An interactive visualization module presents communication networks, timelines, and activity graphs, providing clear and actionable insights. By integrating AI-driven automation with traditional forensic methods, the proposed system significantly enhances the efficiency and accuracy of digital investigations, reducing manual effort and enabling faster evidence discovery.

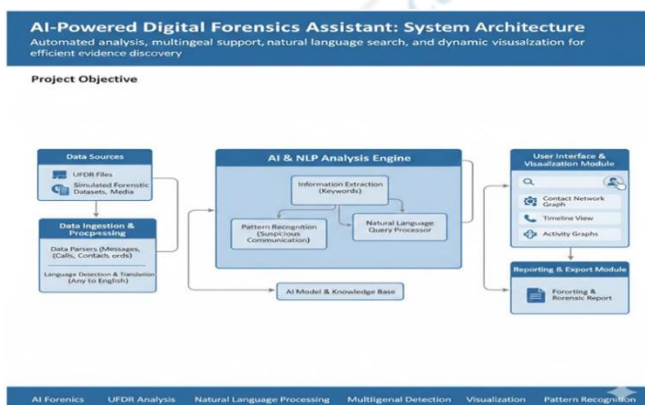
Index Terms— Artificial Intelligence, Digital Forensics, Natural Language Processing, UFDR Analysis, Data Visualization.

I. INTRODUCTION

Digital forensics is very important for modern investigations into cybercrime, fraud detection, and criminal analysis. As digital data grows at an exponential rate, it gets harder and harder to analyze forensic reports like Unified Forensic Data Reports (UFDR).

Old-fashioned forensic methods need people to look through large datasets by hand, which takes a lot of time, is prone to mistakes, and doesn't work very well. Investigators frequently encounter difficulties in discerning pertinent evidence concealed within extensive datasets of communication.

This research proposes an AI-Powered Digital Forensics Assistant to tackle these issues. The system uses both Artificial Intelligence (AI) and Natural Language Processing (NLP) to speed up investigations, automate evidence extraction, and make analysis more accurate.



II. METHODOLOGY

The suggested AI-Powered Digital Forensics Assistant has a systematic way of processing and analyzing forensic data quickly.

A. Gathering Data:

The system can read forensic datasets like Unified Forensic Data Reports (UFDR), which have both structured and unstructured digital evidence like messages, call logs, contacts, and media files. Standard forensic tools are used to collect these datasets from mobile devices and digital storage systems.

B. Preparing the Data:

There is often noise, redundancy, and inconsistencies in raw forensic data. So, preprocessing is done to:

- Get rid of entries that are the same or don't matter.
- Take care of missing or broken data.
- Make timestamps and formats the same.
- Turn unstructured text into structured text.
- This step makes sure that the dataset is clean and ready for analysis.

C. Getting Features:

Key features are taken from the dataset, such as:

- Information about the sender and receiver.
- Content of the message.
- How long and how often calls happen.
- Information about media files.

We use methods like TF-IDF or word embeddings to break up text data into tokens and turn it into formats that machines can read.

D. Analysis Based on AI and NLP:

The system uses AI and Natural Language Processing to look at the data that was pulled out:

- Extracting keywords to find suspicious words.
- Sentiment analysis to find out how someone feels.
- Named Entity Recognition (NER) to find names, places, and money-related things.

Using topic modeling to put conversations that are similar together. Machine learning models are used to find strange patterns and anomalies in how people communicate.

E. Processing in Multiple Languages:

To work with different datasets, the system has:

- Algorithms for language detection.
- Using NLP models for automatic translation.
- This lets investigators look at conversations in different languages without having to translate them by hand.

F. Module for Processing Queries:

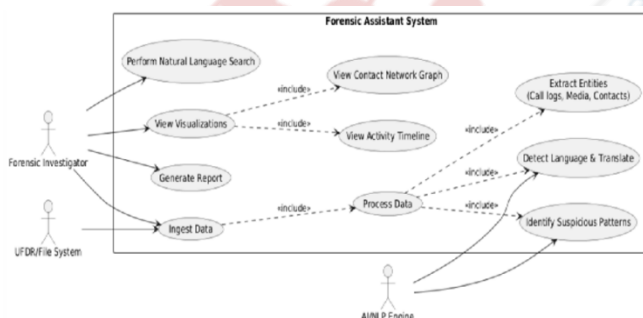
People can ask the system questions in plain English, like:
"Show financial transactions that look suspicious"
"Find contacts that have been made often in the last week"

Using NLP, the system understands the question and gets the right answers.

G. Reporting and Visualization:

The final results are shown by:

- Graphs that show how communication networks work.
- Charts that show activity patterns over time.
- These visual aids make it easier for investigators to quickly understand complicated data.



III. RESULTS AND DISCUSSIONS

The new system really speeds up digital forensic work.

A. Performance Evaluation:

We ran it on some sample forensic data—things like communication logs and media files. Here's what stood out: It processed data way faster than if you did it by hand. It pulled out the important information accurately. It spotted suspicious patterns quickly and efficiently.

B. Efficiency Improvement:

Old-school forensic analysis eats up a lot of time and effort. This system cuts down the hassle in a few key ways: It slashes investigation time because it automates data filtering. It lowers the chances of missing evidence or making mistakes. It makes it much simpler to tackle big, complex datasets.

C. Accuracy and Reliability:

AI models do a better job at picking out keywords, spotting patterns, and finding connections between data points. They deliver steady and reliable results across different datasets, so you can trust the output no matter where the data comes from.

D. Visualization Benefits:

Tools like graphs and timelines make things clearer. They break down complicated relationships, give you instant insight into how people communicate, and help investigators act quickly.

E. Limitations:

Even with all these strengths, the system isn't perfect. Everything depends on having good data. It still struggles with encrypted information, and sometimes AI models carry hidden biases. These are real challenges, but there's plenty of room to tackle them in the next round of updates.

IV. HELPFUL HINTS

The AI-Powered Digital Forensics Assistant brings a lot to the table in different fields:

A. Law Enforcement Agencies:

It lets police dig into digital evidence fast, which really speeds up investigations. They can use it to track down suspects by looking at how people communicate and connect the dots to solve cases.

B. Cybercrime Investigation:

This assistant stands out when it comes to spotting things like phishing, hacking, or online scams. It helps track down shady messages and builds a clearer picture by mapping out the digital trails left behind by cybercriminals.

C. Fraud Detection:

Spotting financial fraud and scams gets easier too. The tool can flag strange transactions and keeps an eye on business communications to catch anything that seems off.

D. Digital Evidence Management:

Keeps forensic data organized and easy to find. Lets you pull up important info fast when you need it. Makes sure legal cases have solid, reliable evidence.

V. SOME COMMON MISTAKES

Building and analyzing AI-driven digital forensic systems comes with some classic pitfalls, and if you want accurate,

reliable results, it's best to steer clear of them.

First, treat “data” as a plural—digital forensic datasets have plenty of records and entries, not just one. If you misinterpret evidence or ignore the broader context, you’ll end up drawing the wrong conclusions. That’s why you really need context-aware analysis. Entity mix-ups (like confusing people’s names, locations, or financial terms) often happen when NLP models aren’t trained well enough. Don’t just trust the machines—always back up AI findings with human review to catch errors machines might miss.

If your system can’t handle multilingual data the right way, bad translations will mess with your investigations. Visualization has its own risks; unclear graphs or badly mapped relationships can twist how you see communication patterns. Also, don’t forget that correlation isn’t causation. Mixing up the two when looking at behavioral trends leads to bad interpretations.

And, of course, if you feed your model noisy or incomplete data without cleaning it up first, the whole thing will underperform.

Notice these details and pay attention to them—you’ll end up with a far more reliable, effective AI-powered forensic system.

VI. EDITORIAL POLICY

The presenting author, **T.N. Sriranjani**, is responsible for the preparation and submission of this paper.

All co-authors have contributed to the work and approved the final manuscript. The research is original and properly cites relevant prior studies. Ethical guidelines have been followed in handling forensic data and system development. The authors are willing to revise and improve the paper based on reviewer feedback.

VII. PUBLICATION PRINCIPLES

This research paper, led by the presenting author **T.N. Sriranjani**, contributes to the advancement of digital forensic analysis by integrating Artificial Intelligence and Natural Language Processing techniques.

1. Contribution to Knowledge:

Under the guidance of the presenting author, this work introduces an AI-based approach for automating forensic evidence analysis. The system enhances traditional methods by enabling faster and more efficient investigation of digital data.

2. Scope and Depth of Work:

The study covers all essential components, including data preprocessing, feature extraction, AI-driven analysis, and visualization. The depth of the research reflects the complexity of handling large-scale forensic datasets.

3. Scientific and Technical Merit:

The proposed system demonstrates strong technical significance through the application of AI and NLP techniques. It improves accuracy, reduces manual effort, and provides meaningful insights for investigators.

4. Reproducibility and Transparency:

The methodology is clearly structured and documented under the supervision of the presenting author, allowing other researchers to replicate and extend the work. The system design and workflow are described in a transparent manner to support future research.

This study adheres to standard publication principles and reflects collaborative research conducted under the leadership of the presenting author, contributing to the field of AI-driven digital forensics.

VIII. CONCLUSION

The explosion of digital data has left old-school forensic methods struggling to keep up. This research introduces an AI-Powered Digital Forensics Assistant built with AI and Natural Language Processing to automate and sharpen the analysis of evidence. This system handles massive forensic datasets with ease. It pulls out important details, spots hidden patterns, and delivers clear, user-friendly visualizations. Investigators spend less time sifting through data, make fewer errors, and move through cases much faster. With multilingual support and simple natural language queries, the system becomes much more accessible for investigators everywhere. Blending AI with modern forensic techniques isn’t just a technical upgrade—it’s a big leap forward for how we analyze digital evidence.

REFERENCES

- [1] Di Domenico, G.; Sit, J.; Ishizaka, A.; Nunan, D. Fake news, social media and marketing: A systematic review. *J. Bus. Res.* 2021, 124, 329–341. [CrossRef]
- [2] Grubl, T.; Lallie, H.S. Applying Artificial Intelligence for Age Estimation in Digital Forensic Investigations. *arXiv* 2022, arXiv:2201.03045.
- [3] Suryanto, H.; Degeng, I.N.S.; Djatmika, E.T.; Kuswandi, D. The effect of creative problem solving with the intervention social skills on the performance.