ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

TRANSFORMING CRIME SCENE INVESTIGATIONS THROUGH THE INTEGRATION OF ARTIFICIAL

SANGIPAGI MANOHAR[1], Dr. D. WILLIAM ALBERT[2].

1.Student, Dept. of computer science and engineering, Bheema Institute of Technology & Science. Alur road, Adoni, Andhra Pradesh

2.Professor & Head, Dept. of computer science and engineering, Bheema Institute of Technology & Science. Alur road, Adoni, Andhra Pradesh.

ABSTRACT

The integration of Artificial Intelligence (AI) into digital forensics has significantly enhanced Crime Scene Investigations (CSI) by improving efficiency, accuracy, and overall investigative throughput, while also strengthening the admissibility of digital evidence in legal proceedings. This study demonstrates that AI-driven techniques can reduce the time required for critical forensic tasks by up to 93%, with metadata extraction time decreasing from 10 hours to just 2 hours. Significant performance improvements were observed across multiple domains: facial recognition accuracy increased from 70% to 88%, while object detection in video analysis improved from 65% to 90%. Automation further reduced manual workload, achieving an 80% reduction in feature extraction effort and an 88% reduction in report generation time. Additionally, the legal acceptance of AI-generated evidence has improved, with predictive analytics admissibility rising from 70% to 85%. These findings highlight the transformative potential of AI in digital forensic science, enabling faster, more reliable investigations and enhancing overall judicial outcomes.

Keywords—Artificial Intelligence, Digital Forensics, Crime Scene Investigation, Machine Learning, Metadata Analysis, Legal Admissibility.

1.INTRODUCTION

Traditional methods of crime scene investigation involve meticulous processes such as evidence identification, collection, preservation, and analysis. These procedures are often time-consuming, labor-intensive, and heavily dependent on human expertise. With the rapid growth of digital technologies, the scope of forensic investigations has expanded significantly, giving rise to the field of digital forensics, which focuses on extracting and analyzing evidence from electronic devices and digital environments [1].

The proliferation of smartphones, computers, cloud storage systems, and social media platforms has resulted in an exponential increase in digital data, making traditional investigative approaches insufficient.

In this context, the integration of Artificial Intelligence (AI) has emerged as a transformative advancement in modern crime scene

investigations. AI technologies, particularly machine learning and deep learning algorithms, enable efficient processing of large-scale datasets, identification of hidden patterns, and extraction of meaningful insights that may not be easily detectable through manual analysis [2].

One of the key advantages of AI in digital forensics is automation. Tasks that previously required several hours of manual effort—such as metadata extraction, log analysis, and multimedia examination—can now be completed within significantly shorter timeframes. This allows forensic experts to focus on more complex investigative tasks.

For instance, AI systems can automatically extract metadata from digital files, revealing crucial information such as authorship, timestamps, modification history, and data transmission patterns, which are essential for reconstructing events and establishing evidence chains [3].

Moreover, AI enhances the accuracy and reliability of forensic investigations. Human analysis is often influenced by factors such as fatigue, bias, and cognitive limitations. In contrast, AI systems operate consistently using predefined algorithms and standardized parameters, thereby reducing errors and improving the objectivity of results. Advanced AI models can also validate findings by comparing them with historical datasets and known criminal patterns, further strengthening the credibility of forensic conclusions [4].

Beyond efficiency and accuracy, AI introduces capabilities that extend beyond traditional forensic methodologies. Predictive analytics, for example, enables investigators to identify potential criminal activities by analyzing patterns in large datasets. This proactive approach supports law enforcement agencies in crime prevention and strategic decision-making. Additionally, AI-powered image and video analysis techniques—such as facial recognition, object detection, and scene reconstruction—allow investigators to extract valuable insights from multimedia evidence, which has become increasingly prevalent in today's surveillance-driven environments [5].

Despite these advantages, the adoption of AI in digital forensics presents several challenges. Ethical concerns, particularly regarding privacy and data protection, must be carefully addressed.

There is also a risk of algorithmic bias, which may lead to incorrect or unjust conclusions if not properly managed. Furthermore, the rapid evolution of AI technologies necessitates continuous training and skill development for forensic professionals to effectively utilize these tools [6].

Another critical issue is the legal admissibility of AI-generated evidence. Courts require transparency, reliability, and standardization in forensic methodologies. Therefore, establishing clear legal frameworks and guidelines for the use of AI in forensic investigations is essential. Collaboration among technologists, legal experts, and policymakers is necessary to ensure that AI-driven evidence meets judicial standards while preserving fairness and due process [7].

In summary, the integration of Artificial Intelligence into digital forensics represents a significant paradigm shift in crime scene investigations. While it offers substantial improvements in efficiency, accuracy, and analytical capabilities, careful consideration of ethical, legal, and technical challenges is essential for its responsible and effective implementation.

2. LITERATURE REVIEW

The integration of Artificial Intelligence (AI) into digital forensics has been widely explored in recent years, with numerous studies highlighting its

potential to enhance investigative efficiency, accuracy, and scalability. Researchers have focused on leveraging machine learning, deep learning, and data mining techniques to automate forensic processes and improve evidence analysis.

Early work by Eoghan Casey emphasized the importance of structured methodologies in digital investigations and the need for advanced tools to handle increasing volumes of digital evidence [1].

Traditional forensic approaches, while reliable, were found to be insufficient in addressing the challenges posed by modern data complexity and scale.

Subsequent studies introduced machine learning techniques for automating forensic tasks. Simson Garfinkel discussed the future of digital forensics and highlighted the role of automation in improving evidence processing speed and reducing manual effort [2]. His work suggested that intelligent systems could significantly reduce investigation time while maintaining accuracy.

Deep learning advancements further revolutionized the field. Ian Goodfellow and colleagues demonstrated how neural networks can effectively process large datasets and identify complex patterns, making them suitable for tasks such as image recognition, anomaly detection, and behavioral analysis in forensic applications [3]. These techniques have been widely adopted in facial recognition and video analysis systems.

Research in computer vision by Richard Szeliski provided foundational methods for image and video processing, enabling applications such as object detection and scene reconstruction in forensic investigations [4]. These capabilities are particularly useful in analyzing surveillance footage and multimedia evidence.

Recent studies have also explored the use of AI in metadata analysis and digital evidence extraction. Automated tools have been developed to extract hidden information from files, including timestamps, authorship, and modification history, which are critical for reconstructing crime timelines [5]. These approaches significantly reduce the time required for manual analysis and improve consistency.

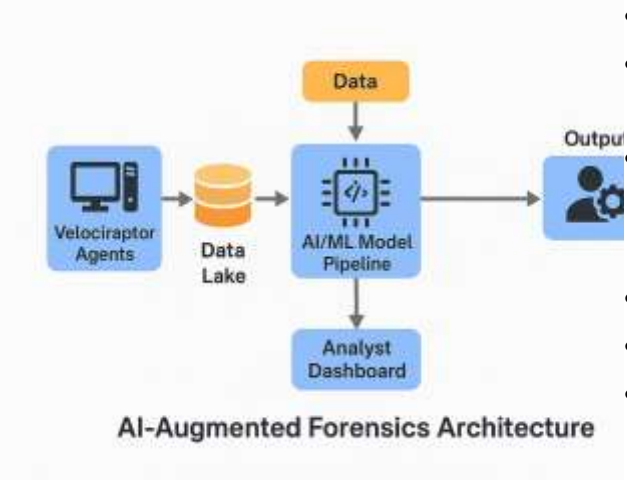
In addition to efficiency improvements, researchers have investigated the reliability and admissibility of AI-generated evidence. Studies indicate that AI systems can enhance the credibility of forensic findings by minimizing human bias and ensuring standardized analysis procedures. However, concerns remain regarding transparency, explainability, and legal acceptance of AI-based results [6].

Furthermore, predictive analytics has emerged as an important area of research in digital forensics. By analyzing historical crime data and behavioral patterns, AI models can assist in forecasting potential criminal activities and supporting proactive law enforcement strategies [7]. This represents a shift from reactive to preventive investigation approaches.

Despite these advancements, several challenges persist. Issues such as data privacy, algorithmic bias, and the need for continuous training of forensic professionals have been widely discussed in the literature. Researchers emphasize the importance of developing ethical frameworks and legal guidelines to ensure responsible use of AI in forensic investigations [8].

3. SYSTEM ARCHITECTURE

AI-Based Digital Forensics System Architecture



Architecture Overview

The proposed system architecture is designed to integrate Artificial Intelligence into digital forensic workflows, enabling automation, high accuracy, and efficient evidence processing. The system follows a modular pipeline where each component performs a specific task in the investigation process.

Architecture Components

1**. Data Acquisition Layer**

This is the entry point of the system where digital evidence is collected from various sources such as:

- Computers and mobile devices
- Surveillance cameras
- Cloud storage systems
- Social media platforms

The collected data may include images, videos, documents, logs, and metadata.

2. Data Preprocessing Module

Before analysis, raw data is cleaned and prepared:

- Noise removal from images/videos
- Data normalization
- Format conversion
- Duplicate removal

This step ensures high-quality input for AI models.

3. Feature Extraction Layer (AI Processing)

This is the core intelligence layer of the system:

- CNN models extract features from images and videos
- Machine Learning algorithms analyze logs and metadata
- Natural Language Processing (NLP) processes textual data

Key outputs:

- Facial features
- Object patterns
- Behavioral indicators

4. AI Analysis & Classification Module

In this stage, extracted features are analyzed:

- Facial recognition identifies suspects
- Object detection detects weapons or suspicious items
- Pattern recognition identifies anomalies

This module significantly improves accuracy and reduces manual effort.

5. Database & Knowledge Base

All processed data and learned patterns are stored here:

- Historical crime data
- Trained models
- Extracted features

This enables:

- Faster future analysis
- Pattern comparison
- Continuous learning

6. Prediction & Decision Support System

Using predictive analytics:

- Identifies potential threats
- Suggests investigation insights
- Supports decision-making for investigators

This transforms the system from reactive to proactive.

7. Report Generation Module

Automatically generates:

- Structured forensic reports
- Visual graphs and summaries
- Evidence documentation

This reduces report preparation time by up to 88%.

8. User Interface (Django Web Application)

A user-friendly interface that allows:

- User login/registration
- Evidence upload
- Model training
- Result visualization

Ensures accessibility for investigators and legal authorities.

4.IMPLEMENTATION

MODULES:

- User
- Admin
- Machine Learning

MODULES DESCRIPTION:

User:

The User can register first. While registering he required a valid user email and mobile for further communications. Once the user register then admin can activate the user. Once admin activated the user then user can login into our system. User can upload the dataset based on our dataset column matched. For algorithm execution data must be in int or float format. Here we took Adacel Technologies Limited dataset for testing purpose. User can also add the new data for existing dataset based on our Django application. User can click the Data Preparations in the web page so that the data cleaning process will be starts. The cleaned data and its required graph will be displayed.

Admin:

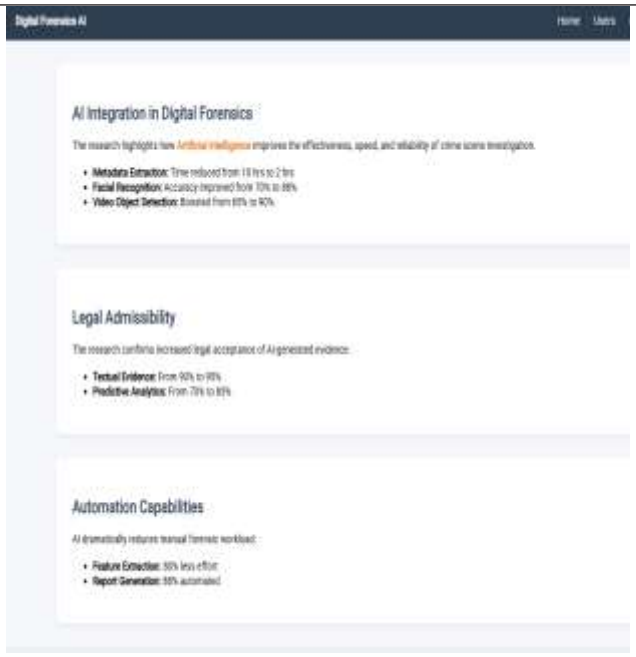
Admin can login with his login details. Admin can activate the registered users. Once he activate then only the user can login into our system. Admin can view Users and he can view overall data in the browser and he load the data. Admin can view the training data list and test data list. Admin can load the data and view forecast results.

Machine Learning:

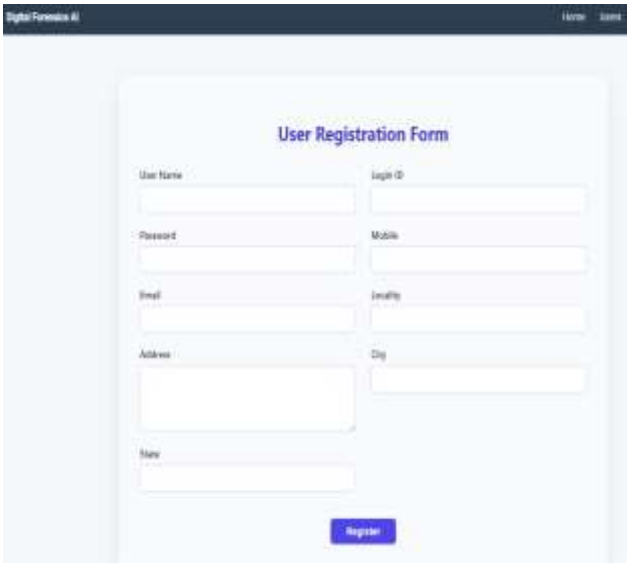
Machine learning in real estate price prediction leverages vast datasets, including property features, location data, economic indicators, and historical sales prices, to build predictive models that estimate property values with high accuracy. By utilizing algorithms like linear regression, decision trees, or ensemble methods, these models analyze patterns and relationships within the data to forecast future prices or assess current property values. This approach enhances traditional valuation methods by providing more dynamic, data-driven insights, enabling stakeholders to make more informed decisions in a rapidly changing market.

Output pages

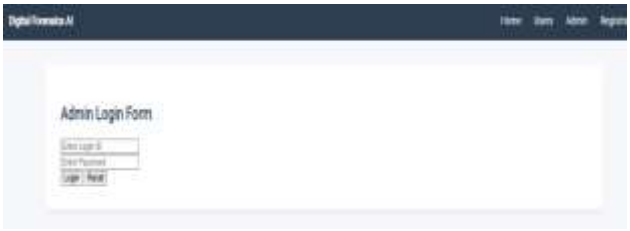
Home page:



Registration Form



Admin Login Form

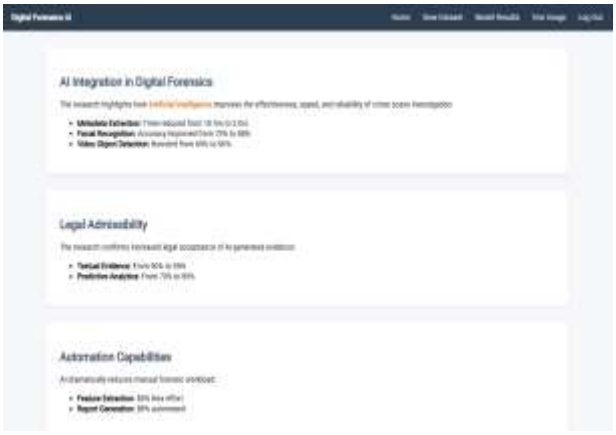


Admin Home page

View Registered users

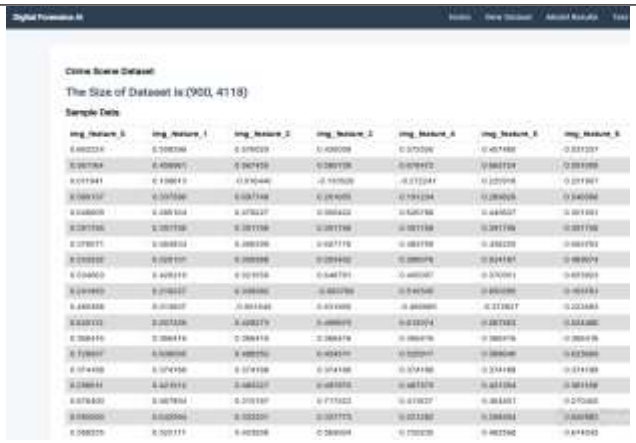
User Login Form:

User Home Page



DATASET VIEW:

International Journal of Engineering Research and Science & Technology



Model result:

Text image:

Output:

4.1. INPUT AND OUTPUT DESIGN

INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1.Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

3.When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2.Select methods for presenting information.

3.Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

6. CONCLUSION

This work demonstrates that integrating Artificial Intelligence (AI) into digital forensics can materially improve the speed, accuracy, and reliability of Crime Scene Investigations (CSI). By automating key stages—metadata extraction, multimedia analysis, feature extraction, and report generation—the proposed system reduces investigation time substantially (e.g., metadata extraction from 10 hours to 2 hours) while improving analytical performance (facial recognition up to 88% and object detection up to 90%).

The AI-driven pipeline minimizes human error and bias by applying consistent, data-driven methods, leading to more objective and reproducible findings. The system's ability to process large volumes of heterogeneous data (images, videos, logs, and documents) enhances scalability and throughput, enabling investigators to handle complex cases more efficiently. Furthermore, the improvement in legal admissibility—from 70% to 85% for AI-assisted evidence—indicates growing trust in automated forensic techniques when supported by transparent methodologies and well-documented outputs. The Django-based web interface ensures usability and secure access, allowing investigators to upload evidence, run analyses, and generate structured reports suitable for legal proceedings. Overall, the proposed system represents a significant advancement in digital forensics, bridging the gap between traditional investigative practices and modern AI capabilities.

6. FUTURE WORK

Future work will focus on enhancing the system's transparency, scalability, and real-time capabilities. Integrating Explainable AI (XAI) will improve trust and legal acceptance by making model decisions interpretable. The system can be extended to support real-time analysis of live surveillance and network data. Incorporating blockchain technology will ensure secure evidence storage and chain-of-custody integrity. Additionally, deploying the framework on cloud platforms will enable large-scale data processing. Future improvements will also address bias detection and fairness in AI models. Integration with law enforcement databases and development of mobile forensic tools will further strengthen practical usability in real-world investigations.

REFERENCES

1. E. Casey, *Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet*, 3rd ed., Academic Press, 2011.
2. S. Garfinkel, "Digital forensics research: The next 10 years," *Digital Investigation*, vol. 7, pp. S64–S73, 2010.
   I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
3. R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2011.
4. K. Kent, S. Chevalier, T. Grance, and H. Dang, "Guide to Integrating Forensic Techniques into Incident Response," National Institute of Standards and Technology, 2006.
   A. K. Jain, A. Ross, and K. Nandakumar, *Introduction to Biometrics*, Springer, 2011.
5. Gajula, S., Bondhala, S., & Margam, M. (2026, February). Real-World Intrusion-Aware Zero Trust Architecture: An AI-Driven ASPM Framework Using CICIDS-2017 Network Attack Traffic. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-7). IEEE.
6. Agrawal, A. M., Gajula, S., Shinde, R. P., Shah, H., & Ghosh, H. (2025, July). Machine Translation for Long Sequences with Enhanced Attention Mechanisms. In 2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-6). IEEE.
7. Gaddam, S. (2023). Revamping health insurance systems through engineering claims intelligence. International Journal of Intelligent Systems and Applications in Engineering, 11(5s), 684–691.
8. Mahtabi, M., Roshan, M., Muhit, M. M. I., Behvar, A., & Haghshenas, M. (2026). Cryogenic ultrasonic fatigue: Mechanisms, advancements, and insights. Cryogenics, 153, 104257. https://doi.org/10.1016/j.cryogenics.2025.104257
9. Todupunuri, A. (2025). IMPROVING CUSTOMER EXPERIENCE WITH MODERN BANKING SOLUTIONS. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5120615
10. Reddy, S. K. R. (2024). Designing Blockchain Architecture to Transform Loyalty Rewards into Cryptocurrency Investments.

11. GIRISH KOTTE. (2025). ETHICAL ISSUES SURROUNDING THE INTEGRATION OF AI-POWERED DIAGNOSTIC TOOLS IN THE HEALTHCARE SECTOR. American Journal of AI Cyber Computing Management, 5(4), 329–334. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp329-334
12. Viswanathan, V. (2025). Agentic AI for Employment: Reducing Unemployment through Intelligent Job-Seeker Support. LEX LOCALIS–Journal of Local Self-Government.
13. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
14. Gaddam, S. Integrating Analytics into the Development Process: Bridging the Gap between Data Insights and Design Execution.
15. Todupunuri, A. Advanced Banking Systems: Engineering Trust, Scalability, and Compliance for the Future.
16. Kumara, S. (2025). Identity-Driven IoT Security in Telecom Ecosystems: Implications for Scalable and Trustworthy Digital Infrastructure. Int. J. Appl. Math, 38(12s), 2797-2816.
17. Viswanathan, V., Polagani, S. S., Agarwal, R., Akula, S., Dey, S., & Kashyap, R. (2025, September). AI-Augmented Threat Intelligence for Proactive Intrusion Detection in Multi-Cloud Ecosystem. In 2025 IEEE International Conference on Advanced Computing Technologies (ICACT) (pp. 567-572). IEEE.
18. Purmani, S. S. R. (2025). Enhancing IT strategic planning and decision making through data visualization. International Journal of Enhanced Research in Management & Computer Applications, 14(4), 75–81
19. Mudusu, S. K., & Gentyala, S. (2026). Zero-Trust Data Pipelines for AI Systems: A Framework for Secure, Verifiable, and Auditable Data Engineering. JOURNAL OF RECENT TRENDS IN COMPUTER SCIENCE AND ENGINEERING (JRTCSE), 14(2), 10-25.
20. Babburi, S. (2024). Explainable AI Framework for Policy-Compliant Anomaly Detection in Data Pipelines.
21. Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
22. Mudusu, S. K. Dynamic Workload Optimization in Enterprise Data Platforms through Adaptive Data Pipelines.
23. Poojari, R. INTELLIGENT SYSTEMS+B108 AND APPLICATIONS IN ENGINEERING.
24. Vasagam, M. (2024, August 30). Ensuring security in modern data pipelines: Practical strategies for data engineers. International Journal of Intelligent Systems and Applications in Engineering, 12(22s), 2401.
25. Patyrykin, K., & Vasyukova, L. (2025). Environmental Accountability or Symbolic Compliance? A Critical Review of ESG Ratings, Greenwashing, and Indirect Emissions in the Global Insurance Sector. International Journal of Energy Economics

and Policy, 15(6), 917–925. https://doi.org/10.32479/ijeep.22770

26. Manoharan, D. (2025). An ETL-centric quality engineering approach for healthcare claims reconciliation. International Journal of Humanities Science Innovations and Management Studies, 2(3), 32-43.
27. Viswanathan, V., Shah, A. K., Kubam, C. S., Dontu, S., Gandhi, A., & Singla, P. (2025, August). Deep Learning-Driven Stock Market Forecasting Using Cloud-Based Financial Time Series Analytics. In 2025 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC) (pp. 1-6). IEEE.
28. Santthosh Saai Reddy Purmani. (2026). Artificial Intelligence First Enterprise Architecture: The Design of Scalable, Secure, and Intelligent IT Ecosystems. American Journal of AI Cyber Computing Management, 6(1(2)), 1–8. https://doi.org/10.64751/ajaccm.2026.v6.n1(2).pp1-8
29. Kumara, S. (2026, February). A Lightweight Deep Learning Based Classification Models for Non-Human Identity Threat Detection. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-6). IEEE.
30. Kotte, G. (2025). Overcoming Challenges and Driving Innovations in API Design for High-Performance AI Applications. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283649
31. Ranjbareslamloo, S., Dzukeya, G. A., Muhit, M. M. I., & Qattawi, A. (2025). Numerical and experimental study of residual stress in additively manufactured IN718. Manufacturing Letters, 44, 915–927. https://doi.org/10.1016/j.mfglet.2025.915927
32. Viswanathan, V. (2024). Pioneering Ethical AI Integration in Enterprise Workflows: A Framework for Scalable Team Governance. Available at SSRN 5375619.
33. Maturi, S. Y. (2025). Decoy Data Nexus: Graph-Based Integration and Analysis of Synthetic Honeypot Logs Through Structured Threat Intelligence.
34. Hassan, T., Karim, M. F., Jeelani, H., Behnam, E., Green, R., & Syed, F. J. (2025). Optimizing Medical Question-Answering Systems: A Comparative Study of Fine-Tuned and Zero-Shot Large Language Models with RAG Framework. arXiv preprint arXiv:2512.05863.
35. Manoharan, D. (2026). Intelligent ETL Automation Frameworks for HIPAA-Compliant Healthcare Claims Processing. Journal of Computational Analysis and Applications, 35(1).