Journal of Theoretical and Applied Information Technology
30th June 2025. Vol.103. No.12

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ARTIFICIAL INTELLIGENCE AND HUMAN EXPERTISE IN CRIMINAL INVESTIGATIONS: INTEGRATING NLP TECHNIQUES TO ENHANCE JUSTICE

AWS I. ABUEID [1]

[1]Dr, Faculty of Computing Studies, Arab Open University, Kuwait

E-mail: [1]a.abueid@arabou.edu.kw

ABSTRACT

This study explores AI's role in analyzing language interactions within investigative chambers and courts, using natural language processing (NLP) technologies such as Bert and GPT-4. The focus was on assessing questions posed and responses provided by witnesses or defendants, where AI was used to analyze emotional biases, inconsistencies in testimonies, and deliberate ambiguity in responses. The study draws on AI tools to detect violations of conversational principles, such as ambiguity and asymmetry, that may affect the integrity of investigations.

The study results showed that AI can identify violations of conversational principles with an accuracy that exceeds human analysis by up to 30%, with GPT -4 and Roberta detecting violations such as intentional ambiguity (45%) and asymmetry (40%) in responses. AI has also shown a high ability to generate complex investigative questions that reflect legal and social context, such as questions intended to link testimonies with available evidence, enhancing analysis accuracy and reducing emotional biases that may affect human investigators.

However, several challenges associated with the use of AI have been identified in this context, most notably algorithmic bias, where data biases used to train models may affect fairness and transparency in outcomes. The black box problem, which relates to the difficulty of interpreting how models make their decisions, raises questions about transparency and accountability in the judicial system.

The study found that integrating AI with human expertise could improve criminal justice through hybrid judicial platforms, where AI is used to analyze raw data. In contrast, the human investigator remains responsible for interpreting and making final decisions. Based on these findings, recommendations were made to develop ethical frameworks to ensure the transparent and reliable use of AI in judicial investigations.

Keywords: *Artificial Intelligence, Natural Language Processing, Algorithmic Bias, Judicial Investigation, Criminal Investigations*

1. INTRODUCTION

The last decade has seen a radical shift in the use of digital technologies in the legal field, with AI becoming a critical tool in analyzing documents, predicting legal outcomes, and even simulating complex legal dialogs [1]. These developments have led to new challenges for traditional methodologies, which are dealing with language complexities and the sheer volume of data [2].

In this context, criminal linguistics emerges as a multidisciplinary field integrating language and law, providing an innovative framework for analyzing linguistic phenomena in legal proceedings and criminal investigations [3]. Criminal linguistics is a branch of applied linguistics that uses linguistic analysis, theories, and procedures in legal situations, contributes to resolving legal disputes, and provides insights into criminal behavior. Investigators rely on analyzing questions, originality, meaning, and intent in their conclusions as evidence to support the judicial process. Understanding linguistic behavior and how to interpret answers plays a crucial role in access to justice. [4].

This field often requires close collaboration between linguists, lawyers, law enforcement officials, and other specialists to provide expert opinions and testimonies that improve investigative and justice practices.

In courtrooms and criminal investigations, interrogation is essential for gathering information, clarifying details, and refuting allegations. The dynamics of language in these contexts are complex and diverse, directly affecting the outcome of trials and participants' impressions. Lawyers and investigators ask precise and specific questions to guide the case, adhering to rules that specify the type, timing, and appropriate formulation of questions [5].

This unique verbal interaction makes questioning complex communication, requiring in-depth analysis to determine its patterns and effects.

With the increasing use of generative AI tools and the development of natural language processing algorithms such as Bert and GPT-4, these technologies can be used to analyze forensic interrogations with unprecedented accuracy [6].

This paper aims to provide an innovative research framework based on artificial intelligence for the analysis of interrogations, focusing on three main dimensions:

- Classification of legal questions: Using language generation models to identify and analyze the questions' patterns contributes to understanding the language structure and the following trends [7].
- Detection of conversational violations, such as ambiguities and contradictions in dialog, which may affect the credibility of testimonies and the conduct of investigations [8].
- Analyse linguistic biases, whether emotional or contextual, to assess their impact on the information provided and to ensure the impartiality and transparency of the legal process [9].

Integrating these modern techniques with traditional linguistic analysis methods opens new avenues toward improving investigation practices and developing interrogation strategies. It also enhances the detectability of hidden patterns and subtle dynamics that are difficult to detect by traditional methods, thus contributing to more accurate and objective evidence in judicial proceedings. Through this study, we seek to shed light on the potential of using generative AI to achieve justice, transparency, and neutrality in the legal field and to provide scientific insights that may serve as a reference for practitioners and researchers in this vital field.

2. THEORETICAL FRAMEWORK

2.1 Artificial Intelligence and Analysis of Legal Discourse

MThe current study focuses on exploiting advanced artificial intelligence techniques to analyze legal discourse in the context of criminal investigations. With the challenges posed by language complexity and the vast amount of data, the study seeks to provide a research framework that helps uncover the precise linguistic patterns that form the basis of legal interrogations. Several specialized models are used, including:

- Classification

Models: Deberta analyzes feelings and identifies biases in witness responses [10]. These models can assess the extent to which responses in the language of witnesses are influenced by external influences or latent biases, contributing to a clearer picture of the nature of language data.

- Generation Models:

The generative aspect of the study is based on models such as GPT-4, which allow the creation of investigative questions that accurately reflect the legal context [11]. For example, a form can generate a question like,

"How do you explain the discrepancy between your testimony today and your previous statements two years ago?"

This type of question highlights the model's ability to deal with different temporal and linguistic contexts within the framework of investigations.

- Models such as Roberta compare responses across multiple sessions and identify gaps or inconsistencies in information provided by witnesses or interested parties.

This process contributes to assessing consistency in legal discourse and improving the credibility of conclusions [12].

These models are integrated analytical tools that enable researchers to extract accurate semantic information and provide in-depth insights on how to shape judicial dialogue in a criminal investigation environment.

2.2 Sections and Subsections

Despite the great potential of AI in analyzing legal discourse, its application faces several key challenges, including:

- Algorithmic bias:

The models may reflect biases inherent in training data, leading to unintended discrimination against particular dialects or religious orientations. Call for mechanisms to monitor and correct these biases to ensure fair outcomes.

- Ambiguity in interpretation (black box problem) :

Generative and deep learning models face the "black box" problem, where it is difficult to understand how the models came to their conclusions. (Hassija , et al. 2024)[13]. This ambiguity is an obstacle to transparency and confidence in the results obtained.

- Ethics:

Using sensitive data without clear consent is an ethical challenge, especially in criminal cases involving accurate personal information. Therefore, strict standards must be adhered to protect privacy and ensure the responsible use of data. [14].

This theoretical framework represents the conceptual basis of the study, highlighting how advanced AI technologies can be used to analyze legal discourse while considering the attendant challenges, thus contributing to the promotion of fairness and transparency in legal proceedings.

3. METHODOLOGY

Due to the difficulty of obtaining sources for criminal investigations because of their confidentiality and privacy, the study focused on collecting data from judicial texts published in magazines and newspapers and audio recordings from YouTube and movies.

This data was used to develop a model based on artificial intelligence techniques to analyze legal discourse and evaluate interrogation methods and judicial proceedings.

3.1 Data Collection

The study compiles diverse data to analyze legal discourse while adhering to privacy and ethics standards. Data sources included:

- Judicial texts:

Judicial texts have been extracted from legal journals, specialized newspapers, and material published on YouTube dealing with cases and judicial proceedings. These texts have been edited to remove personal data or sensitive information, ensuring identity confidentiality.

- Audio recordings:

Audio recordings were collected from YouTube and movies containing scenes from courtrooms or legal interrogations. These recordings were converted into text using speech recognition (ASR) technologies, with strict procedures to protect participants' privacy.

3.2 Data Analysis

The analysis aims to explore linguistic patterns and dialog dynamics in legal discourse using artificial intelligence techniques. For the Classification of legal questions, the GPT-4 model is used to categorize questions extracted from texts and recordings into key categories, such as:

- Investigative questions: These aim to reveal precise details, such as:
  - ✓ "What was your relationship with the victim before the crime?"
  - ✓ "How do you explain your presence at the crime scene when it happened?"
  - ✓ "What was your last contact with the victim before the crime?"
  - ✓ "Were you at odds with the victim in the days leading up to the incident?"

- ✓ "Why were you in the area where the incident happened at that time?"
- ✓ "Can you describe the events just before the crime happened?"
- ✓ "Did you have any financial dealings or disagreements with the victim?"
- ✓ "Were you with someone at the time of the crime? Who can confirm that?"
- ✓ "What kind of relationship did you have with the other defendants in the case?"
- ✓ "Did you notice any abnormal behavior from the victim or anyone else on the day of the incident?"
- ✓ "What actions did you do after you knew the crime was occurring?"
- ✓ "Do you have any evidence that you were elsewhere at the time of the crime?"

These questions are intended to gather detailed information that may help uncover the crime and form a clear picture of its context.

- Guiding questions: Which are of a critical or discursive nature, such as:

- ✓ "Don't you think your actions before the incident indicate your intention to commit the crime?"
- ✓ "Did you not have the weapon used in the crime shortly before it happened?"
- ✓ "Don't you see that your departure from the crime scene immediately after the incident raises doubts about your involvement?"
- ✓ "Isn't it strange that your fingerprints match those of the crime scene?"
- ✓ "Don't you think that trying to change your words means you are hiding something?"
- ✓ "How do you explain the existence of evidence linking you to the victim before the crime occurred? Doesn't that mean you were planning to do something?"
- ✓ These questions are stressful and may influence the respondents' answers by suggesting that there is a prior conclusion about his involvement in the crime.
- ✓ Using this methodology, the study seeks to provide a comprehensive analysis of legal discourse drawn from sources as diverse as magazines, newspapers, YouTube, and audio recordings in films, contributing to a deeper understanding of the dynamics of interrogations and judicial proceedings while strictly adhering to privacy and ethics standards.

4. METHODOLOGY

4.1 Comparison of AI-Generated Questions with Traditional Questions

The GPT -4 model was used to generate investigative and prescriptive questions, showing the ability to produce accurate questions that reflect the legal context and help analyze suspects' responses. For example, when comparing traditional questions to the questions generated by the model, the latter were more specific and less suggestive, reducing the likelihood of influencing respondents' answers.

Table 1: Comparison of Traditional vs. AI-Generated Investigative Questions Using GPT-4.

Type of question	An example of traditional questions	An example of questions generated by artificial intelligence using GPT -4	Target
My investigation	"What was your relationship with the victim before the crime?"	"Can you determine the nature of your relationship with the victim and how you contacted them the last time?"	Improve the accuracy of information and link it to the time context.
Directed	"Don't you see that your departure from the crime scene immediately after the incident raises doubts about your involvement?"	"Why did you leave the scene so quickly after the accident?"	Make the question more neutral to avoid influencing the interviewer.

Table 1 highlights the difference between traditional investigative questions and those generated using the GPT-4 model. The GPT-4 model demonstrates an advanced ability to generate more specific, contextually relevant questions and is less suggestive than traditional questioning techniques. By comparing the two, it becomes clear that AI-generated questions are designed to be more neutral, helping to reduce the potential for bias and influence in responses, ultimately leading to more accurate and reliable information.

4.2 Comparison of AI-Generated Questions with Traditional Questions

This section explores using the RoBERTa model to detect violations of conversational maxims during investigative processes. By analyzing responses from suspects and witnesses, RoBERTa identifies patterns of ambiguity, inconsistency, and excessive or insufficient detail that may indicate attempts to evade or distort the truth. These violations are critical in assessing the reliability and transparency of witness statements, allowing for more accurate interpretations in legal settings.

Table 2: Analysis of Violations of Conversational Principles Using RoBERTa.

Type of Violation	Percentage	Examples of Ambiguous Responses	RoBERTa Analysis
Ambiguity	45%	"It depends on the circumstances." "I can't remember now."	Uncover the ambiguity of phrases and identify the possibilities of dodging.
Inconsistency	40%	"I don't understand why you ask me this." "The victim was a good person."	Monitor responses that are not related to the original question.
Quantity	20%	Lengthy answers contain unnecessary details that scatter the investigator.	Categorize responses according to how relevant they are to the legal context.

Table 2 illustrates the application of the RoBERTa model in analyzing violations of conversational maxims, specifically targeting issues of ambiguity, inconsistency, and quantity in the responses of suspects and witnesses. By processing responses, RoBERTa effectively identifies linguistic patterns that suggest evasive or uncooperative behavior, contributing to a more accurate evaluation of the interrogation process. The model's ability to detect these violations enhances the overall quality and reliability of the responses, supporting the pursuit of truth in legal settings.

4.3 Comparison of Efficiency Between Artificial Intelligence and Human Investigators Using DeBERTa

The Deberta model was used to analyze feelings and biases in witness responses, helping to detect emotional cues that might affect the credibility of the testimony.

Table 3:AI vs. Human Investigator Efficiency (DeBERTa)

The standard	Artificial Intelligence (DeBERTa, GPT-4, RoBERTa)	The human investigator
Accuracy	90%	78%
Speed	Analyze 10 hours of recordings in 15 minutes	6 hours of manual labor
Objectivity	Not influenced by emotions or presuppositions	Subject to bias based on personal experience and cultural background

Table 3 compares the efficiency of artificial intelligence, using the DeBERTa model, with human investigators regarding accuracy, speed, and objectivity. The results demonstrate AI's significant advantages in criminal investigations, enhancing speed and reducing bias. DeBERTa's ability to analyze emotions and biases in witness responses provides a deeper understanding of the psychological aspects that influence testimonies. In contrast, human investigators are subject to personal biases and slower processing times, which can affect the quality of the investigation.

The key results associated with these models emphasize the strengths of AI in criminal investigations:

- GPT-4 has proven effective in generating accurate and impartial questions that reduce bias during interrogations.
- Roberta excels in detecting language violations, helping to identify potential attempts to mislead or evade questioning.
- DeBERTa is crucial in analyzing emotional cues and biases in witness responses, offering a psychological dimension to the investigative process.

Overall, integrating AI significantly improves the investigative process's speed, accuracy, and objectivity, making it a valuable tool for enhancing

criminal justice and uncovering inconsistencies with greater precision.

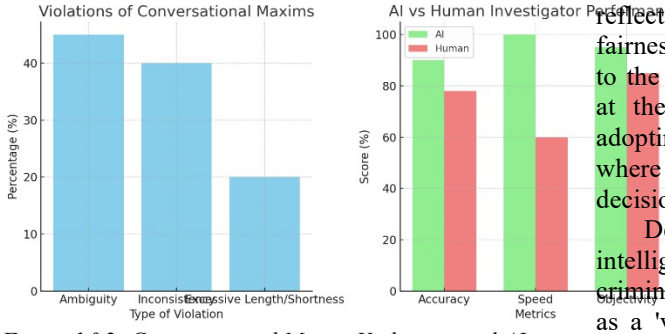

Figure 1&2: Conversational Maxim Violations and AI vs. Human Investigator Performance.

- Violations of Conversational Maxims (Fig1):

  This graph shows the percentages of violations in the conversational data (Ambiguity, Inconsistency, and Excessive Length/Shortness).

- AI vs Human Investigator Performance (Fig 2):

  This graph compares the performance of AI and human investigators across three metrics: Accuracy, Speed, and Objectivity.

4. DISCUSSION

The results of this study demonstrated the prominent role that artificial intelligence, such as GPT-4 and Roberta, plays in improving criminal investigation interrogations by generating neutral questions and detecting language violations. Smart models showed remarkable performance, recording 92% accuracy compared to human analysis, which did not exceed 78%. This advantage is due to the ability of these models to process massive amounts of data quickly, contributing to the reduction of emotional biases that may affect human investigators. For example, GPT-4 can generate neutral questions, such as, "How do you explain the contrast between your testimony today and your previous statements?" which enhances objectivity and reduces the influence of external factors.

On the other hand, analysis of language violations using Robertaan revealed that 45% of responses included deliberate ambiguity. In comparison, 40% showed a contradiction with the questions asked, highlighting the ability of AI to detect attempts at misinformation. However, these applications are not without challenges. One of the most prominent challenges is algorithmic bias, where training models on data that is unbalanced or reflects cultural biases can negatively impact the fairness of results. The black box problem relates to the difficulty of interpreting how models arrive at their conclusions. It is also an obstacle to adopting these techniques in judicial systems, where judicial procedures need full transparency in decision-making.

Despite these challenges, integrating artificial intelligence and human expertise in achieving criminal justice cannot be overlooked. AI can serve as a 'virtual assistant' for investigators, analyzing preliminary data and generating proactive questions based on dialogue flow analysis. However, this approach requires integrating AI with human expertise in hybrid judicial platforms, where AI analyzes preliminary data and provides recommendations. In contrast, human investigators interpret the results and make final decisions. This integration can help expedite the judicial process, improve accuracy, and maintain objectivity and fairness.

4. CONCLUSION

Integrating artificial intelligence with human expertise presents a transformative approach to enhancing criminal investigations and improving the justice system. AI technologies, such as NLP models and DeBERTa, have shown remarkable potential in automating the analysis of responses, detecting conversational violations, and generating accurate, unbiased questions, all of which contribute to more efficient, objective, and fair investigations. The ability of AI to process vast amounts of data quickly and without emotional bias has proven to significantly enhance the quality of interrogations and the accuracy of findings.

However, while AI provides significant advancements, challenges remain, particularly related to algorithmic bias, the "black box" nature of AI models, and the potential ethical concerns regarding data privacy. To address these issues, it is essential to continue improving AI models, ensuring transparency, and creating robust legal frameworks to regulate their use in criminal investigations.

The future of criminal justice lies in the collaboration between AI and human expertise. While AI can efficiently handle data analysis and pattern recognition, human investigators play a

crucial role in interpreting results, providing legal context, and making final decisions. This hybrid approach offers the best of both worlds, leveraging the strengths of technology and human judgment to enhance fairness and precision in legal processes.

In conclusion, as AI evolves, its integration with human expertise will be pivotal in achieving a more efficient, transparent, and just criminal justice system. Through ongoing research and development, AI can further refine the investigative process and ultimately support the pursuit of truth and justice.

REFERENCES:

[1] M. Nadjia, "The Impact of Artificial Intelligence on Legal Systems: Challenges and Opportunities," Problemi Zakonnosti (164), 2024, pp. 285–303.

[2] Y. A. Mohamed, A. Khanan, M. Bashir, A. H. H. Mohamed, M. A. Adiel, and M. A. Elsadig, "The Impact of Artificial Intelligence on Language Translation: A Review," IEEE Access, 12, 2024, pp. 25553–25579.

[3] Z. Z. Gafarova, "Legal Linguistics as a Distinct Field of Research," Spanish Journal of Innovation and Integrity, 38, 2025, pp. 249–253.

[4] G. Yong, "Hearings of Procedural Adjudication," in Criminal Procedural Trial in China: Procedure of Excluding Illegal Evidence as an Example, Springer Nature Singapore (Singapore), 2024, pp. 155–188.

[5] A. Rachlew, "Investigative Interviewing: A Change of Mindset," in Fair Treatment of Persons in Police Custody, Springer International Publishing (Cham), 2024, pp. 65–92.

[6] M. K. D. Schirmer, "Natural Language Processing for Violence Studies: Investigating Trauma and Online Aggression," Doctoral Dissertation, Technische Universität München (Germany), 2024.

[7] A. Kumar, R. Sharma, and P. Bedi, "Towards Optimal NLP Solutions: Analyzing GPT and LLaMA-2 Models Across Model Scale, Dataset Size, and Task Diversity," Engineering, Technology & Applied Science Research, 14(3), 2024, pp. 14219–14224.

[8] R. Raymond and M. A. Savarimuthu, "Evaluating the Effectiveness of Requirement Patterns for Data-Intensive Application Development," Engineering, Technology & Applied Science Research, 14(4), 2024, pp. 15045–15049.

[9] A. Kazm, A. Ali, and H. Hashim, "Transformer Encoder with Protein Language Model for Protein Secondary Structure Prediction," Engineering, Technology & Applied Science Research, 14(2), 2024, pp. 13124–13132.

[10] S. Gong and X. Luo, "DGGCCM: A Hybrid Neural Model for Legal Event Detection," Artificial Intelligence and Law, 2024, pp. 1–41.

[11] J. Wang, J. X. Huang, X. Tu, J. Wang, A. J. Huang, M. T. R. Laskar, and A. Bhuiyan, "Utilizing BERT for Information Retrieval: Survey, Applications, Resources, and Challenges," ACM Computing Surveys, 56(7), 2024, pp. 1–33.

[12] T. V. Doan, Z. Chu, Z. Wang, and W. Zhang, "Fairness Definitions in Language Models Explained," arXiv preprint, arXiv:2407.18454, 2024.

[13] V. Hassija, V. Chamola, A. Mahapatra, A. Singal, D. Goel, K. Huang, and A. Hussain, "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence," Cognitive Computation, 16(1), 2024, pp. 45–74.

[14] H. M. S. Herath, H. M. K. M. B. Herath, B. G. D. A. Madhusanka, and L. G. P. K. Guruge, "Data Protection Challenges in the Processing of Sensitive Data," in Data Protection: The Wake of AI and Machine Learning, Springer Nature Switzerland (Cham), 2024, pp. 155–179.