Integritas: Jurnal Antikorupsi
Vol 11, No. 2, 2025, pp. 255-264

https://jurnal.kpk.go.id/index.php/integritas
©Komisi Pemberantasan Korupsi

Integrating the DeepSeek-R1 model as an analytical assistant in digital forensic investigations of corruption cases

Hafni Ferdian [a]*, Ruki Harwahyu [b]
Universitas Indonesia. Depok 16424, Indonesia
[a] hafni.ferdian@ui.ac.id; [b] ruki.h@ui.ac.id
* Corresponding Author

Abstract*:* This article examines the integration of the DeepSeek-R1 large language model as an analytical assistant in digital forensic investigations, particularly in corruption cases. The growing volume of digital evidence often leads to substantial analysis backlogs that can extend for months or even years, thereby hindering law enforcement efforts (Kadage et al., 2024). The use of Artificial Intelligence (AI) and Large Language Models (LLMs) offers the potential to improve investigative efficiency by reducing the burden of processing massive datasets. DeepSeek-R1 is a newly released open-source LLM with advanced reasoning capabilities, achieving performance comparable to state-of-the-art models developed by OpenAI. This study outlines the role of DeepSeek-R1 in supporting digital forensic workflows—from tracing electronic evidence and analyzing data relationships to generating preliminary investigative reports. The methodology includes simulated corruption case experiments comparing investigations assisted by DeepSeek-R1 with conventional manual analysis. The results show that DeepSeek-R1 can accelerate the retrieval of relevant information and produce concise summaries rapidly, significantly reducing analysis time. However, the model remains prone to factual errors (hallucinations) and biases, making human validation by investigators essential. With proper risk mitigation and oversight, integrating DeepSeek-R1 has the potential to significantly enhance the effectiveness of digital forensic investigations in combating corruption.
Keywords: Digital forensics; Artificial intelligence; Large language model; Investigative assistant; Anti-corruption enforcement

How to Cite: Ferdian, H., & Harwahyu, R. (2025). Integrating the DeepSeek-R1 model as an analytical assistant in digital forensic investigations of corruption cases. *Integritas: Jurnal Antikorupsi*, *11*(2), 255-264. https://doi.org/10.32697/integritas.v11i2.1557

Introduction

The advancement of information technology has led to an increasing presence and volume of digital evidence involved in criminal investigations, including corruption cases. Today, corruption investigations often require the examination of thousands of electronic files, emails, instant messages, and financial records. This massive volume of data presents a serious challenge for law enforcement agencies; digital forensic laboratories experience months-long case backlogs that negatively impact judicial processes (Kadage et al., 2024). The rapid growth of digital evidence has resulted in significant delays in forensic processing, hindering investigations and court proceedings. Reports have even indicated instances of cases being dismissed because digital evidence could not be analyzed within statutory detention limits, or because suspects passed away before the investigation was completed (Kadage et al., 2024). These conditions highlight the urgent need for new strategies to accelerate digital forensic investigations without compromising accuracy or the integrity of law enforcement.

In practical settings, similar challenges are observed in corruption cases involving public procurement. Investigators may receive thousands of digital artefacts, including instant messages between public officials and private contractors, emails discussing specification changes, and financial documents indicating suspicious fund flows. All these artefacts must be reviewed manually under tight time constraints, increasing the risk of overlooking critical information. This illustrates the pressing need for technological assistance to support faster and more structured early-stage evidence identification.

One potential solution to this challenge is the adoption of artificial intelligence (AI) and machine learning in digital forensic workflows. AI has been widely recognized as a transformative

technology capable of improving efficiency and analytical precision in forensic investigations. Prior research has explored AI-based automation for various forensic tasks, such as data mining, pattern recognition, and big-data forensic analysis. More recently, the emergence of Generative AI and Large Language Models (LLMs) such as GPT-4 and ChatGPT has opened new opportunities. LLMs are capable of understanding and generating natural language, making them suitable for assisting investigators with information filtering, text interpretation, language translation, and report drafting. Recent literature suggests that the adoption of LLMs in digital forensics—when implemented with proper controls and safeguards—has the potential to enhance investigative efficiency, strengthen audit trails, and reduce both technical and procedural barriers in law enforcement.

Despite these promises, the use of LLMs in digital forensic investigations raises several concerns. Previous studies indicate that although LLMs such as GPT-4 and ChatGPT can assist in certain forensic tasks, including artefact interpretation, evidence searching, and code-script generation, these models also present serious limitations (Scanlon et al., 2023). In particular, they are prone to generating linguistically plausible but factually inaccurate responses. OpenAI itself acknowledges that GPT-4 is "not fully reliable," often hallucinating facts or making reasoning errors, and advises caution against its use in high-stakes contexts. Risks associated with the use of LLMs in digital forensics include training-data bias that may affect outputs, hallucinated results that could mislead investigations, insufficient explainability, potential over-reliance by investigators, and legal-ethical concerns related to data confidentiality—especially when cloud-based services require uploading sensitive evidence. Previous studies also find that while LLMs can assist in drafting forensic reports, their outputs require thorough expert review and correction to ensure accuracy and admissibility (Michelet & Breitinger, 2024). In other words, LLMs should serve as supportive tools for experts—not as replacements for human investigators.

This study focuses on integrating the DeepSeek-R1 LLM as an analytical assistant for digital forensics in corruption investigations. DeepSeek-R1 was developed by DeepSeek-AI with the objective of enhancing reasoning ability in LLMs through large-scale reinforcement learning without initial dependence on supervised fine-tuning (Guo et al., 2025). As a result, DeepSeek-R1 demonstrates reasoning performance comparable to OpenAI's proprietary models (OpenAI-o1-1217, equivalent to GPT-4) on several problem-solving benchmarks. These capabilities position DeepSeek-R1 as a promising AI engine capable of understanding complex patterns in corruption-case data. Unlike GPT-4, which is commercial and cloud-based, DeepSeek-R1 is open-source and can be deployed locally. The model is available in multiple parameter sizes ranging from approximately 1.5 billion to 70 billion via distillation techniques, enabling law-enforcement institutions to tailor its use to their computational resources. Local deployment ensures that sensitive data—such as investigative documents from the Corruption Eradication Commission (KPK)—does not need to be uploaded to third-party services, thus preserving confidentiality. With strong reasoning capabilities and full operational control, DeepSeek-R1 has the potential to serve as a virtual assistant that helps investigators process digital evidence more efficiently, accurately, and securely.

Based on this context, the study formulates three central research questions: (1) How can DeepSeek-R1 be integrated as an analytical assistant in digital forensic investigations of corruption cases? (2) To what extent can the model improve the efficiency and quality of evidence assessment compared with manual investigative analysis? (3) What legal, ethical, and procedural risks must be considered before implementing AI in corruption investigations? The objectives of this study are: (1) To design an integration scenario for incorporating DeepSeek-R1 into digital forensic investigation stages in corruption cases; (2) To evaluate its performance and impact on investigative effectiveness through experimental simulations; and (3) To discuss the findings, benefits, and limitations of DeepSeek-R1 as an investigative assistant in the context of law enforcement and anti-corruption work.

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Methods

This study employs an experimental approach using a simulated digital forensic investigation scenario based on corruption cases. The integration framework for DeepSeek-R1 was designed by adapting standard digital forensic process models, such as the DFRWS model or the NIST forensic framework, consisting of the collection, examination, analysis, and reporting stages. DeepSeek-R1 was implemented as an assistant during the examination and analysis phases and also supported the preparation of the initial investigation report. In this setup, DeepSeek-R1 receives a collection of digital artefacts (e.g., email dumps, chat logs, financial documents) derived from the simulated case. The model then produces outputs in the form of extracted key information, document summaries, identification of patterns or anomalies, and draft investigative narratives.

Case Scenario and Test Data

We constructed a corruption-case scenario involving multi-source digital evidence. The scenario includes elements such as alleged bribery supported by email exchanges between involved parties, instant-message logs that obscure illegal transactions, and spreadsheets containing hidden financial records. The test dataset was partly derived from curated public forensic datasets commonly used for research purposes, and partly synthesized to resemble realistic corruption-case patterns. Each artefact was assigned a ground-truth label indicating its relevance and whether it contains indications of corruption, which served as an evaluation reference.

Application of DeepSeek-R1

The DeepSeek-R1 model used in this study has 14 billion parameters, chosen as a balance between capability and computational efficiency. The model was run locally on a server equipped with an adequate GPU to ensure that no case data left the testing environment. Prompt engineering was applied to instruct the model to complete various forensic tasks: (1) Emails: "Summarize the following email and identify whether it contains any indication of bribery or solicitation of illicit payments."; (2) Chat logs: "Analyze this chat log and identify any messages related to illegal projects or payments."; (3) Spreadsheets: "Here are the transaction records. Identify suspicious patterns (e.g., unusually large amounts or suspicious descriptions)." The model then generated textual outputs for each of these prompts. In addition, DeepSeek-R1 was used in an interactive mode: investigators could ask follow-up questions through prompts (e.g., "Is there any connection between the chat message on date X and the email on date Y?"), and the model provided reasoning-based responses grounded in the supplied dataset.

Evaluation Experiments

To assess the contribution of DeepSeek-R1, we compared three analysis scenarios: (a) Manual Analysis — human investigators analyzed the dataset without AI assistance; (b) Analysis with DeepSeek-R1 — investigators used the model's outputs and recommendations as part of their analysis; (c) Analysis with Another LLM (baseline) — for comparison, ChatGPT (GPT-4) was used via API, where allowable portions of the dataset were uploaded and analyzed in a similar manner. For scenario (c), data security considerations were strictly applied: the dataset was anonymized and partitioned to avoid exposing sensitive information since cloud upload was required. All scenarios were performed on identical datasets.

Performance Metrics

Multiple quantitative and qualitative metrics were measured: (1) Time Efficiency: total duration required to identify all key evidence; (2) Completeness of Findings: measured using recall, i.e., the percentage of relevant evidence items successfully identified in each scenario; (3) Accuracy of Findings: measured using precision, i.e., the percentage of identified items that were truly relevant, including detection of false alarms; (4) Report Quality: two digital forensic experts evaluated the initial investigative reports generated by the model (case summary and findings) on a 1–5 scale, assessing clarity, accuracy, and usefulness. Manual reports produced by investigators were also evaluated as a baseline; (5) Error Analysis: each significant model error (e.g.,

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

severe factual hallucinations or failure to detect crucial evidence) was documented and categorized (e.g., factual error, loss of contextual meaning, etc.). Interaction logs with the model were reviewed to understand how DeepSeek-R1 arrived at its conclusions to the extent observable, given the model's black-box nature.

Qualitative Analysis

In addition to the above metrics, structured interviews were conducted with three digital forensic investigators from law enforcement agencies, who tested DeepSeek-R1 on the simulated dataset with anonymized case identifiers. They were asked to provide feedback on usability, relevance of outputs, level of trust in the model's findings, and recommendations for real-world implementation. These insights were used to complement the quantitative results and to consider practical field implications.

This methodological approach was designed to provide a comprehensive understanding of DeepSeek-R1's impact as an investigative assistant. By evaluating the model under controlled simulation conditions, we aimed to identify both its strengths and limitations before recommending its application to real-world corruption cases.

Results and Discussion

Experimental Results

Efficiency and Accuracy. From the perspective of time efficiency, integrating DeepSeek-R1 clearly accelerated the analysis process. In the manual analysis scenario, investigators required an average of 9 working hours to filter the entire dataset and identify key evidence. With the assistance of DeepSeek-R1, this time was reduced by approximately 40%, to around 5.5 hours. The efficiency gain was most evident in the initial screening phase: the model automatically generated summaries for each document within minutes, allowing investigators to quickly discard irrelevant items and focus on potential evidence. For example, DeepSeek-R1 was able to highlight 8 out of 10 emails containing bribery-related keywords (such as "transfer" or "project commission"), along with their contextual information, in a single query. This aligns with the findings of previous studies, who reported that AI can help analysts process large volumes of data rapidly without sacrificing rigor in the detection of critical evidence (Dunsin et al., 2024).

Table 1. Comparison of analysis time and artifact-identification accuracy between manual analysis, DeepSeek-R1, and ChatGPT

Method	Average Analysis Time per Case	Recall (Relevant Items Retrieved)	Precision (Relevant Items / Total Flags)
Manual (without DeepSeek-R1)	9 hours	75%	95%
With DeepSeek-R1	5.5 hours (↑ ≈ +40% more efficient)	92% (↑ higher)	85% (↓ Slightly lower)
With ChatGPT	4 hours	80%	70%

(Recall and precision are computed against the ground-truth dataset; ↑ and ↓ indicate improvement or decline relative to the manual method.)

As shown in Table 1, the use of DeepSeek-R1 significantly reduced analysis time while also improving the coverage of identified evidence. Higher recall indicates that the model was able to detect more relevant items that might otherwise be missed in manual analysis, whereas the slightly lower precision indicates the need for verification to remove a small number of false positives. From a practical standpoint, this trade-off is acceptable because the time saved is considerably greater than the additional effort required to validate a limited number of less accurate detections. Thus, integrating an LLM provides a net positive benefit to the investigative process.

In terms of completeness of findings, the DeepSeek-R1 scenario achieved a recall of approximately 90%, meaning that the vast majority of relevant evidence items were successfully identified. This is higher than manual analysis (≈75%) and the ChatGPT-assisted scenario (≈80%). This suggests that DeepSeek-R1 is more effective in capturing diverse evidence patterns—likely due

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

to its reasoning-oriented training and multi-step, chain-of-thought-like capabilities. For instance, the model successfully linked an obfuscated chat message ("the package has been delivered") with a transaction in the spreadsheet on the same date, which human investigators initially overlooked on their first pass. The model's precision was also reasonably high (≈85%), slightly below manual analysis (≈95%). This indicates that DeepSeek-R1 produced some false positives—for example, flagging a routine internal email as suspicious merely because it contained the word "invoice," even though its context was unrelated to bribery. Nevertheless, a precision level of 85% is still strong for an automated system, and the false alerts can be filtered by investigators. Compared with ChatGPT, DeepSeek-R1 exhibited higher precision (ChatGPT ≈70%), suggesting that this open-source model is better trained to avoid misleading outputs in the tested forensic scenario, possibly due to its architecture and training process being more explicitly focused on correct reasoning. These findings support the claim that DeepSeek-R1 offers reasoning and contextual understanding approaching top-tier models such as GPT-4.

Quality of Analytical Assistance and Reporting

DeepSeek-R1 not only accelerated evidence identification but also assisted in constructing the narrative of the case. The model produced a draft case summary of roughly three pages, outlining the actors involved, a chronology of events based on email and chat timestamps, and key suspicious transactions. Experts rated this draft as helpful, giving it an average score of 4 out of 5 for having a systematic structure and logical flow. They particularly appreciated the model's ability to weave relationships between pieces of evidence into a coherent story task that ordinarily consumes substantial time when done manually. By comparison, the initial reports generated by ChatGPT, although linguistically fluent, contained several inaccuracies such as misinterpreting chat context and introducing assumptions not supported by the data, and thus received an average score of 3 out of 5. This reinforces the view that open-source models that can be tailored or customized, such as DeepSeek-R1, are better suited to domain-specific tasks than general-purpose LLMs that lack forensic-domain adaptation. However, experts also identified limitations in DeepSeek-R1's draft reports. Some legal terminology was imprecise; for example, the model occasionally described "bribery" as a "gift" in its summaries, requiring legal experts to correct these terms. In addition, the model tended not to reference the precise sources of evidence in a manner consistent with forensic standards (e.g., explicit links to evidence items or file hashes), meaning investigators had to manually add these details to ensure that reports complied with evidentiary requirements in court. This limitation is understandable, as the language model does not have an intrinsic concept of chain of custody—that responsibility remains with investigators, who must ensure that every finding can be traced back to the original evidence.

Error Analysis and Risk Considerations

Several model errors were documented during the experiments. In one instance of hallucination, DeepSeek-R1 stated that "there was a direct order from a superior to secure project X," even though no explicit statement to that effect appeared in the dataset—the model appeared to have inferred this from multiple subtle cues. Although such inference might seem logically plausible, under evidentiary standards, conclusions of that nature must be drawn by investigators based on documented facts, not presented as direct claims generated by AI. This finding is consistent with OpenAI's warning that GPT-4 and other advanced LLMs can hallucinate facts and commit reasoning errors when left to generate conclusions autonomously (Achiam et al., 2023).

Accordingly, every output from DeepSeek-R1 must be verified. In the case above, investigators flagged the statement and searched for supporting evidence. It turned out that the model had combined information from two separate emails to synthesize its conclusion. After review, investigators revised the report by reframing the statement as a hypothesis requiring further corroboration, rather than as an established fact.

Bias-related issues were also observed. For example, the model was more likely to flag English-language communications as "suspicious" compared with informal Indonesian-language exchanges. This may stem from the composition of its training data, which likely contains a higher proportion of English content and global corruption patterns, making the model more attuned to those

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

contexts than to local nuances. In one Indonesian chat log, a slang phrase such as "oke sip, nanti malam dikirim" ("okay, it will be sent tonight") was not highlighted as significant, even though the broader conversation suggested a probable money transfer. This lack of sensitivity to local language and context could be mitigated by fine-tuning the model on Indonesian corruption-case data, though this lies beyond the scope of the present study. What is clear is that local investigators must be aware of such potential biases and avoid relying solely on AI when dealing with Indonesian-language content. Instead, they can enrich prompts with contextual explanations or employ bilingual configurations.

From a legal standpoint, all stakeholders agreed that AI tools such as DeepSeek-R1 should be treated solely as analytical aids, not as decision-makers. AI-generated outputs cannot be presented as evidence in court, but may be used as preliminary intelligence or investigative leads. Therefore, standard operating procedures (SOPs) are needed to regulate AI use in digital forensics within institutions such as the Corruption Eradication Commission (KPK) or the police. These SOPs should cover, among others: ensuring that only copies of evidence, not the original artefacts, are processed by AI to preserve evidentiary integrity; recording all AI-assisted analysis steps for auditing purposes (e.g., documenting prompts and outputs separately in case files); and requiring that all final conclusions be verified and formally endorsed by certified human investigators. Such approach helps maintain chain of custody and legal accountability even when AI is used as an assistant.

Investigators' Challenges and Researcher Positionality

The use of AI in investigations is strongly influenced by differences in digital literacy and professional experience among investigators. Those accustomed to digital analysis tools tend to interpret model recommendations more quickly, while those from non-technical backgrounds require more time to assess the relevance and reliability of outputs. This affects the consistency of analytical quality when AI is employed as a supporting tool. In this study, the researcher had direct experience handling digital evidence in corruption cases, providing contextual understanding of investigative needs. However, this position may also introduce optimism bias toward AI-based solutions. To mitigate such bias, the research design combines quantitative evaluation with analytical discussions grounded in actual model outputs from the experiments.

Legal, Ethical, and Privacy Implications

The use of AI in investigations must comply with confidentiality requirements and personal data protection principles, as mandated by Indonesia's Personal Data Protection Law. Digital evidence in corruption cases frequently contains sensitive information that may only be accessed by authorized parties. Consequently, all interactions with the model must be conducted locally and logged for internal audit. AI outputs cannot be treated as standalone evidence, but rather as supporting analytical information. Each finding generated by the model must be traceable back to its originating digital artefact to preserve the integrity of the chain of custody. Therefore, AI must operate within a strict forensic governance framework.

Risk of Over-Reliance on AI

The convenience offered by AI systems carries the risk that investigators may become overly dependent on model recommendations. Without appropriate safeguards, investigative reasoning could shift toward merely validating AI output rather than conducting an independent, critical assessment. This is particularly dangerous in corruption cases, which require multi-layered evidentiary scrutiny. Human-in-the-loop mechanisms and procedures mandating manual verification of all key conclusions are therefore essential. Digital literacy training is likewise a prerequisite, enabling investigators to understand the model's limitations, prevent over-reliance, and maintain analytical autonomy throughout the investigative process.

Comparison with Related Literature

The findings of this study are consistent with earlier literature indicating that LLMs can serve as useful supporting tools in the hands of knowledgeable users. Previous studies evaluating GPT-

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

4 in forensic contexts found that the model can assist with various tasks such as artefact search, carving script generation, and log anomaly detection, but still requires users to recognize when the AI is wrong (Scanlon et al., 2023). We observed a similar pattern: DeepSeek-R1 effectively accelerates routine tasks and reveals patterns, yet investigator expertise remains crucial for validating results.

Other studies have reported that while LLMs can draft forensic reports, they cannot fully replace human report authors due to accuracy issues (Michelet & Breitinger, 2024). Our findings support this conclusion—DeepSeek-R1's drafts are useful but require manual refinement. The observed advantages of DeepSeek-R1 over ChatGPT, particularly in terms of data privacy and customizability, echo the argument that local or open-source models mitigate the need to upload sensitive evidence to the cloud. Prior work emphasizes the importance of deploying LLMs in controlled environments when using them as investigative copilots (Henseler & van Beek, 2023). This aligns closely with the on-premises deployment approach adopted for DeepSeek-R1 in this study.

Interestingly, multi-agent AI approaches have also begun to be explored in the forensic domain. Recent research has introduced intelligent agent frameworks in which multiple agents collaborate in digital investigations (Wickramasekara & Scanlon, 2024). In the future, DeepSeek-R1 could potentially be integrated into a multi-agent ecosystem, for instance, working alongside task-specific AI agents used for image forensics or network analysis, to deliver a more holistic investigation. In the context of anti-corruption enforcement, this could mean one LLM agent focusing on financial documents while DeepSeek-R1 concentrates on textual communication, with consistent information exchange to ensure that no evidence is overlooked. However, such multi-agent coordination would require robust control mechanisms to maintain accountability and prevent compounding errors.

Discussion of Implications

Overall, the integration of DeepSeek-R1 demonstrates clear added value in digital forensic investigations of corruption cases. The speed and breadth of AI-assisted analysis can reduce case backlogs and help investigators find the proverbial "needle in a haystack" within vast datasets, thereby accelerating anti-corruption law enforcement. This can enhance the performance of institutions such as the KPK in uncovering cases more proactively.

At the same time, our findings underscore that AI is not a magic solution that can operate autonomously without oversight. The principle of human-in-the-loop remains non-negotiable. From a legal standpoint, regulations governing AI use in pro-justitia processes need to be developed—for example, determining whether AI-assisted analytic steps must be disclosed in court as part of investigative methodology for transparency purposes. Investigators' competencies in IT and AI also need to be strengthened so they can both leverage and critically assess AI-generated results.

Ethical considerations are equally important. DeepSeek-R1 and similar LLMs are trained on internet-scale data that may embed particular cultural biases. When applied to Indonesian corruption cases, the model may exhibit bias—for instance, being more suspicious of terminology associated with certain business sectors due to skewed training data dominated by Western cases. This should be anticipated through local fine-tuning and periodic evaluation of AI output bias. On the positive side, DeepSeek-R1's open nature allows the broader community to contribute to model improvements, in contrast to closed commercial systems.

Conlusion

This study demonstrates that the DeepSeek-R1 model can significantly accelerate the early identification of digital evidence in corruption investigations. The model assists investigators in detecting patterns, interpreting the contextual meaning of communications, and identifying relevant documents. Compared with manual analysis, DeepSeek-R1 expedites document filtering and reduces the risk of overlooking critical information during early investigative stages. Relative to traditional analytical methods, the use of DeepSeek-R1 shortens evidence-identification time and

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

increases recall. In other words, AI can automate time-consuming routine tasks, allowing human investigators to focus on higher-level reasoning and strategic decision-making. This aligns with global trends showing that the adoption of LLMs in digital forensics can enhance investigative efficiency and reduce both technical and legal-procedural barriers.

Nevertheless, the use of DeepSeek-R1 must be accompanied by clear oversight, responsibility, and caution due to risks of bias, misinterpretation, potential privacy violations, and the possibility of investigator over-reliance on model recommendations. This study underscores the necessity of human investigators to verify AI outputs and provide the appropriate legal context. DeepSeek-R1 is not a replacement for investigators but rather an intelligent supporting tool. Every AI-generated output should be treated as supplementary information requiring confirmation. We recommend that law enforcement agencies, including the Corruption Eradication Commission (KPK), develop formal guidelines and training programs before deploying such AI systems. Data security is also crucial; while the open-source nature of DeepSeek-R1 enables secure local deployment, data-handling procedures must continue to follow forensic principles to preserve the integrity of evidence. Our findings also highlight the need to develop standard operating procedures (SOPs) for AI use in investigations, strengthen investigator capacity through digital literacy training, and conduct further studies regarding AI integration within Indonesian legal practice. AI usage must remain aligned with legal and ethical frameworks of investigation and must not compromise confidentiality or accountability in law enforcement.

From a development perspective, the results of this study open several avenues for future research. First, fine-tuning DeepSeek-R1 or similar models specifically on corruption-domain and Indonesian-language datasets would likely improve the model's sensitivity to local context. Second, research on explainable AI in forensics should be advanced—for example, mechanisms enabling LLMs to explain the basis of their outputs or highlight supporting evidence, thereby simplifying verification and increasing trust among investigators and judges. Third, controlled field trials on real-world cases are necessary to evaluate AI performance in operational environments, including its impact on investigative team workflows. Fourth, multidisciplinary collaboration among digital forensic experts, legal scholars, and data scientists is essential to develop regulatory and ethical frameworks for AI use in law enforcement.

In conclusion, integrating AI systems such as DeepSeek-R1 reflects an inevitable technological progression in the fight against corruption. With responsible and evidence-based implementation, this technology can become a powerful ally in uncovering increasingly sophisticated white-collar crimes. Combating corruption in the digital era requires innovation, and DeepSeek-R1 demonstrates how artificial intelligence can enhance the investigative capacity of human investigators.

Referensi

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., & Anadkat, S. (2023). Gpt-4 technical report. In *arXiv preprint arXiv:2303.08774*. Cornell University. https://doi.org/10.48550/arXiv.2303.08774

Dunsin, D., Ghanem, M. C., Ouazzane, K., & Vassilev, V. (2024). A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response. *Forensic Science International: Digital Investigation*, *48*, 301675. https://doi.org/10.1016/j.fsidi.2023.301675

Guo, D., Yang, D., Zhang, H., Song, J., Zhang, R., Xu, R., Zhu, Q., Ma, S., Wang, P., & Bi, X. (2025). Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. In *arXiv preprint arXiv:2501.12948*. https://doi.org/10.48550/arXiv.2501.12948

Henseler, H., & van Beek, H. (2023). ChatGPT as a copilot for investigating digital evidence. *Proceedings of the Third International Workshop on Artificial Intelligence and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2023)*, 58–69.

Kadage, A. D., Nayak, B. M., Vishal Sharad Hingmire, Wanjale, K., Bogiri, N., & Mandale, P. L. (2024). AI-enhanced digital forensics: Automated techniques for efficient investigation and evidence collection. *Journal of Electrical Systems*, *20*(1s), 211–229.

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

https://doi.org/10.52783/jes.766

Michelet, G., & Breitinger, F. (2024). ChatGPT, Llama, can you write my report? An experiment on assisted digital forensics reports written using (local) large language models. *Forensic Science International: Digital Investigation*, *48*, 301683. https://doi.org/10.1016/j.fsidi.2023.301683

Scanlon, M., Breitinger, F., Hargreaves, C., Hilgert, J.-N., & Sheppard, J. (2023). ChatGPT for digital forensic investigation: The good, the bad, and the unknown. *Forensic Science International: Digital Investigation*, *46*, 301609. https://doi.org/10.1016/j.fsidi.2023.301609

Wickramasekara, A., & Scanlon, M. (2024). A framework for integrated digital forensic investigation employing autogen AI agents. *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*, 1–6. https://doi.org/10.1109/ISDFS60797.2024.10527235

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Integritas: Jurnal Antikorupsi
Vol 11, No. 2, 2025, pp. 255-264

https://jurnal.kpk.go.id/index.php/integritas
©Komisi Pemberantasan Korupsi

Integrasi model *DeepSeek*-R1 sebagai asisten analisis dalam investigasi forensik digital pada kasus korupsi

Hafni Ferdian [a] *, Ruki Harwahyu [b]

Universitas Indonesia. Depok 16424, Indonesia
[a] hafni.ferdian@ui.ac.id; [b] ruki.h@ui.ac.id
* Corresponding Author

Abstrak*:* Artikel ini membahas integrasi model bahasa besar DeepSeek-R1 sebagai asisten analisis dalam investigasi forensik digital, khususnya untuk kasus korupsi. Meningkatnya volume data yang berasal dari bukti digital kerap menimbulkan penumpukan analisis bukti yang dapat berlangsung berbulan-bulan hingga bertahun-tahun, menghambat penegakan hukum (Kadage et al., 2024). Penggunaan *Artificial Intelligence* (AI) dan *Large Language Model* (LLM) menawarkan potensi untuk meningkatkan efisiensi investigasi dengan meringankan beban analisis data yang masif. *DeepSeek*-R1 merupakan model LLM terbaru berkemampuan penalaran tinggi yang tersedia secara *open-source*, dengan kinerja penalaran sebanding model mutakhir milik *OpenAI*. Penelitian ini menguraikan peran *DeepSeek*-R1 dalam membantu proses forensik digital-mulai dari penelusuran bukti elektronik, analisis keterkaitan data, hingga penyusunan laporan awal investigasi. Metode yang digunakan mencakup eksperimen simulasi kasus korupsi dengan membandingkan investigasi yang dibantu *DeepSeek*-R1 versus analisis konvensional. Hasil menunjukkan bahwa DeepSeek-R1 dapat mempercepat penemuan informasi relevan dan menghasilkan ringkasan yang ringkas dengan cepat, sehingga secara signifikan mengurangi waktu analisis. Namun, model masih rentan terhadap kesalahan faktual (*hallucinations*) dan bias, sehingga validasi manusia oleh penyidik tetap esensial. Dengan mitigasi risiko dan pengawasan yang tepat, integrasi DeepSeek-R1 berpotensi meningkatkan efektivitas investigasi forensik digital dalam upaya pemberantasan korupsi.
Kata kunci: Digital forensik; Kecerdasan buatan; *Large language model*; Asisten investigasi; Pemberantasan korupsi

How to Cite: Ferdian, H., & Harwahyu, R. (2025). Integrating the DeepSeek-R1 model as an analytical assistant in digital forensic investigations of corruption cases. *Integritas: Jurnal Antikorupsi*, *11*(2), 255-264. https://doi.org/10.32697/integritas.v11i2.1557

Pendahuluan

Kemajuan teknologi informasi telah menyebabkan meningkatnya keberadaan dan volume bukti digital dalam penyidikan tindak pidana, termasuk kasus korupsi. Saat ini, penyidikan korupsi kerap memerlukan pemeriksaan ribuan berkas elektronik, email, pesan instan, dan catatan keuangan. Volume data yang masif ini menjadi tantangan serius bagi aparat penegak hukum; laboratorium forensik digital sering mengalami penumpukan perkara selama berbulan-bulan yang berdampak negatif terhadap proses peradilan (Kadage et al., 2024). Pertumbuhan cepat bukti digital telah menimbulkan keterlambatan signifikan dalam pemrosesan forensik, sehingga menghambat penyidikan dan proses persidangan. Laporan bahkan menunjukkan adanya kejadian perkara dihentikan atau gugur karena bukti digital tidak dapat dianalisis dalam batas waktu penahanan yang ditentukan undang-undang, atau karena tersangka meninggal sebelum penyidikan selesai (Kadage et al., 2024). Kondisi ini menegaskan kebutuhan mendesak akan strategi baru untuk mempercepat investigasi forensik digital tanpa mengorbankan akurasi maupun integritas penegakan hukum.

Dalam praktik, tantangan serupa juga terlihat pada perkara korupsi yang berkaitan dengan pengadaan barang dan jasa. Penyidik dapat menerima ribuan artefak digital, termasuk pesan instan antara pejabat publik dan kontraktor swasta, email yang membahas perubahan spesifikasi, serta dokumen keuangan yang mengindikasikan aliran dana mencurigakan. Seluruh artefak tersebut harus ditelaah secara manual dalam tekanan waktu yang ketat, sehingga meningkatkan risiko terlewatnya informasi yang krusial. Hal ini menunjukkan perlunya bantuan teknologi untuk mendukung identifikasi bukti tahap awal secara lebih cepat dan terstruktur.

Salah satu upaya untuk mengatasi tantangan tersebut adalah mengadopsi teknologi AI dan machine learning dalam proses forensik digital. AI telah diakui sebagai teknologi transformatif yang dapat meningkatkan efisiensi dan ketelitian investigasi forensik digital. Berbagai penelitian terdahulu telah mengeksplorasi penggunaan AI untuk otomatisasi tugas-tugas forensik, misalnya data mining, pattern recognition, hingga analisis big data forensik. Secara khusus, kemunculan Generative AI dan Large Language Models (LLM) seperti GPT-4 dan ChatGPT membuka peluang baru dalam konteks ini. LLM mampu memahami dan menghasilkan bahasa alami, sehingga berpotensi membantu penyidik dalam menyaring informasi, memahami konten teks, menerjemahkan bahasa, hingga menyusun laporan. Studi literatur terbaru menyimpulkan bahwa adopsi LLM dalam forensik digital, dengan kontrol dan batasan yang tepat, berpotensi meningkatkan efisiensi investigasi, memperkuat jejak audit, dan mengurangi hambatan teknis maupun prosedural dalam penegakan hukum.

Meski menjanjikan, penggunaan LLM dalam investigasi forensik digital juga menimbulkan sejumlah kekhawatiran. Studi-studi sebelumnya menunjukkan bahwa meskipun LLM seperti GPT-4 dan ChatGPT dapat membantu tugas forensik tertentu (misalnya tugas forensik dalam interpretasi artefak, pencarian bukti, dan pembuatan skrip kode), model-model tersebut tetap memiliki keterbatasan serius (Scanlon et al., 2023). Secara khusus, model cenderung menghasilkan jawaban yang tampak meyakinkan secara bahasa namun keliru secara faktual. OpenAI sendiri menyatakan model GPT-4 "tidak sepenuhnya andal," sering berhalusinasi fakta dan melakukan kesalahan penalaran dan memperingatkan agar berhati-hati menggunakannya dalam konteks yang berisiko tinggi dan menuntut keandalan. Risiko penggunaan LLM dalam forensik digital antara lain meliputi: bias atau kesalahan pada data latih yang dapat memengaruhi output, kecenderungan halusinasi yang dapat menyesatkan penyidikan, isu legal-etis terkait kerahasiaan data (misalnya jika harus mengunggah bukti ke layanan *cloud*), kurangnya *explainability*, hingga potensi *over-reliance* (ketergantungan berlebihan) oleh penyidik. Penelitian terdahulu menemukan bah-wa LLM mampu membantu penulisan laporan forensik, namun memerlukan peninjauan dan koreksi mendalam oleh pakar untuk memastikan kebenaran dan keabsahan konten (Michelet & Breitinger, 2024). Dengan kata lain, LLM sebaiknya diposisikan sebagai asisten yang mendukung ahli, bukan sebagai pengganti penyidik manusia.

Penelitian ini berfokus pada integrasi model LLM DeepSeek-R1 sebagai asisten analisis forensik digital di lingkup investigasi korupsi. DeepSeek-R1 dikembangkan oleh DeepSeek-AI dengan tujuan meningkatkan kemampuan penalaran LLM melalui reinforcement learning skala besar tanpa ketergantungan awal pada *supervised fine-tuning* (Guo et al., 2025). Hasilnya, Deep-Seek-R1 menunjukkan performa penalaran setara dengan model LLM tertutup milik OpenAI (OpenAI-o1-1217, sekelas GPT-4) pada berbagai *benchmark* pemecahan masalah. Keunggulan ini menjadikan DeepSeek-R1 kandidat menarik sebagai *engine AI* yang dapat memahami pola kom-pleks dalam data kasus korupsi. Selain itu, berbeda dengan GPT-4 yang komersial dan berjalan di *server cloud*, DeepSeek-R1 tersedia secara terbuka (*open-source*) dan dapat di-*deploy* secara lokal. Model ini dirilis dalam berbagai ukuran parameter (mulai1,5 miliar hingga 70 miliar) melalui teknik distilasi, memungkinkan instansi penegak hukum menyesuaikan penggunaan model sesuai sumber daya komputasi mereka sendiri. Ketersediaan model secara lokal memastikan data sensitif — seperti dokumen penyidikan Komisi Pemberantasan Korupsi (KPK)— tidak perlu diunggah ke layanan pihak ketiga, sehingga kerahasiaan dapat terjaga. Dengan kemampuan penalaran yang kuat dan kendali operasional penuh, DeepSeek-R1 berpotensi menjadi asisten virtual yang membantu penyidik memproses bukti digital secara lebih efisien, akurat, dan aman.

Berdasarkan konteks tersebut, penelitian ini merumuskan tiga pertanyaan utama: (1) Bagaimana model DeepSeek-R1 dapat diintegrasikan sebagai asisten analisis dalam proses investigasi forensik digital pada kasus korupsi? (2) Sejauh mana model tersebut dapat meningkatkan efisiensi dan kualitas penelaahan bukti dibandingkan analisis manual oleh penyidik? (3) Apa saja implikasi hukum, etika, dan risiko prosesual yang perlu dipertimbangkan sebelum pemanfaatan kecerdasan buatan diterapkan dalam penyidikan korupsi? Tujuan penelitian ini adalah untuk: (1) Merancang skenario integrasi DeepSeek-R1 dalam tahapan investigasi forensik digital kasus korupsi; (2) Mengevaluasi kinerja dan dampaknya terhadap efektivitas investigasi melalui

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

eksperimen; dan (3) Membahas temuan, manfaat, serta keterbatasan penggunaan DeepSeek-R1 sebagai asisten penyidik dalam konteks penegakan hukum dan pemberantasan korupsi.

Metode

Penelitian ini menggunakan pendekatan eksperimen dengan skenario simulasi investigasi forensik digital berbasis kasus korupsi. Kerangka integrasi DeepSeek-R1 dirancang dengan mengadaptasi model proses forensik digital standar, seperti model DFRWS atau kerangka NIST, yang mencakup tahap koleksi, pemeriksaan, analisis, dan pelaporan. DeepSeek-R1 diimplementasikan sebagai asisten pada fase pemeriksaan dan analisis, serta mendukung penyusunan laporan investigasi awal. Dalam pengaturan ini, DeepSeek-R1 menerima kumpulan artefak digital (misalnya dump email, *log chat*, dokumen keuangan) yang berasal dari kasus simulasi. Model kemudian menghasilkan keluaran berupa ekstraksi informasi kunci, ringkasan dokumen, identifikasi pola/anomali, serta draf narasi investigatif.

Skenario Kasus dan Data Uji

Kami menyusun skenario kasus korupsi yang melibatkan bukti digital dari berbagai sumber. Skenario mencakup unsur dugaan suap pejabat dengan bukti berupa e-mail antara pihak-pihak terkait, log pesan instan yang menyamarkan transaksi ilegal, dan file spreadsheet berisi catatan transaksi keuangan tersembunyi. Dataset uji sebagian berasal dari dataset forensik publik yang umum digunakan untuk penelitian, dan sebagian disintesis agar menyerupai pola realistis kasus korupsi. Setiap artefak diberi label ground-truth mengenai relevansi dan indikasi korupsi sebagai acuan evaluasi.

Penerapan DeepSeek-R1

Model DeepSeek-R1 yang digunakan berukuran 14 miliar parameter, dipilih sebagai kompromi antara kapabilitas dan efisiensi komputasi. Model dijalankan secara lokal pada server dengan GPU memadai agar tidak ada data kasus yang keluar dari lingkungan pengujian. *Prompt engineering* digunakan untuk mengarahkan model menyelesaikan tugas forensik: (1) Email: "Ringkas email berikut dan identifikasi apakah ada indikasi suap atau permintaan pembayaran ilegal."; (2) *Log chat*: "Analisis *log chat* ini dan identifikasi pesan terkait proyek atau pembayaran ilegal."; (3) *Spreadsheet*: "Berikut catatan transaksi. Identifikasi pola mencurigakan (misalnya jumlah tidak wajar atau deskripsi mencurigakan)." Model lalu menghasilkan keluaran teks untuk setiap *prompt*. Selain itu, DeepSeek-R1 digunakan secara interaktif: penyidik dapat mengajukan pertanyaan lanjutan melalui *prompt* (misalnya, "Apakah ada hubungan antara pesan *chat* tanggal X dan email tanggal Y?"), dan model memberikan jawaban berbasis penalaran yang berlandaskan dataset yang diberikan.

Eksperimen Evaluasi

Untuk menilai kontribusi DeepSeek-R1, kami membandingkan tiga skenario analisis: (a) Analisis manual—penyidik menganalisis dataset tanpa bantuan AI. (b) Analisis dengan DeepSeek-R1—penyidik menggunakan keluaran dan rekomendasi model sebagai bagian analisis. (c) Analisis dengan LLM lain (*baseline*)—sebagai pembanding, ChatGPT (GPT-4) digunakan melalui API, dengan bagian dataset yang diizinkan diunggah dan dianalisis secara serupa. Untuk skenario (c), pertimbangan keamanan data diterapkan secara ketat: dataset dianonimkan dan dipartisi untuk menghindari paparan informasi sensitif karena diperlukan unggah ke *cloud*. Ketiga skenario dilakukan pada dataset yang sama.

Pengukuran Kinerja

Kami mengukur metrik kuantitatif dan kualitatif: (1) Efisiensi waktu: total durasi untuk mengidentifikasi seluruh bukti kunci. (2) Kelengkapan temuan: diukur dengan recall, yaitu persentase item bukti relevan yang berhasil teridentifikasi pada tiap skenario. (3) Akurasi temuan: diukur dengan precision, yaitu persentase item yang ditandai relevan dan benar-benar relevan, termasuk deteksi false alarm. (4) Kualitas laporan: dua pakar forensik digital menilai laporan investigasi awal yang dihasilkan model (ringkasan kasus dan temuan) pada skala 1–5 berdasarkan kejelasan,

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

akurasi, dan kegunaan. Laporan manual penyidik juga dinilai sebagai *baseline*. (5) Analisis kesalahan: setiap kesalahan model yang signifikan (misalnya halusinasi faktual berat atau kegagalan mendeteksi bukti krusial) didokumentasikan dan dikategorikan (misalnya kesalahan faktual, hilangnya makna konteks, dan sebagainya). Log interaksi dengan model ditelaah untuk memahami bagaimana DeepSeek-R1 sampai pada kesimpulannya sejauh dapat diamati, mengingat sifat black-box model.

Analisis Kualitatif

Selain metrik tersebut, dilakukan wawancara terstruktur dengan tiga penyidik forensik digital dari institusi penegak hukum yang menguji DeepSeek-R1 pada dataset simulasi dengan *identifier* kasus yang dianonimkan. Mereka diminta memberi umpan balik tentang *usability*, relevansi keluaran, tingkat kepercayaan pada temuan model, dan rekomendasi implementasi dunia nyata. Temuan ini melengkapi hasil kuantitatif dan mempertimbangkan implikasi praktis di lapangan.

Pendekatan metodologis ini dirancang untuk memberikan pemahaman komprehensif tentang dampak DeepSeek-R1 sebagai asisten investigasi. Dengan mengevaluasi model pada kondisi simulasi yang terkontrol, kami bertujuan mengidentifikasi kekuatan dan keterbatasannya sebelum merekomendasikan penerapan pada kasus korupsi nyata.

Hasil dan Pembahasan

Hasil Eksperimen: Efisiensi dan Akurasi

Dari perspektif efisiensi waktu, integrasi DeepSeek-R1 secara jelas mempercepat proses analisis. Dalam skenario analisis manual, penyidik memerlukan rata-rata 9 jam kerja untuk menyaring seluruh dataset dan mengidentifikasi bukti kunci. Dengan bantuan DeepSeek-R1, waktu tersebut berkurang sekitar 40%, menjadi sekitar 5,5 jam. Peningkatan efisiensi paling terlihat pada tahap penyaringan awal: model secara otomatis menghasilkan ringkasan untuk setiap dokumen dalam hitungan menit, sehingga memungkinkan penyidik dengan cepat menyingkirkan item yang tidak relevan dan memfokuskan perhatian pada bukti potensial. Sebagai contoh, DeepSeek-R1 mampu menyoroti 8 dari 10 email yang mengandung kata kunci terkait suap (seperti “transfer”, “komisi proyek”) beserta informasi kontekstualnya, hanya dalam satu kueri. Hal ini sejalan dengan temuan studi sebelumnya yang melaporkan bahwa AI dapat membantu analis memproses volume data yang besar dengan cepat tanpa mengorbankan ketelitian dalam mendeteksi bukti kritis (Dunsin et al., 2024).

Tabel 1. Perbandingan durasi analisis dan akurasi identifikasi artefakantara metode manual dan dengan DeepSeek-R1 dan ChatGPT

Metode	Rata-rata Waktu Analisis per Kasus	Recall (Temuan Relevan Terambil)	Precision (Temuan Relevan/Total)
Manual (tanpa *DeepSeek-R1*)	9 jam	75%	95%
Dengan *DeepSeek-R1*	5.5 jam (↑ efisien ~+40%)	92% (↑ lebih tinggi)	85% (↓ sedikit lebih rendah)
Dengan ChatGPT	4 Jam	80%	70%

(*Recall* dan *precision* dihitung terhadap *ground truth dataset*; panah ↑↓ menunjukkan peningkatan/penurunan relatif terhadap metode manual.)

Sebagaimana ditunjukkan pada Tabel 1, penggunaan DeepSeek-R1 secara signifikan mengurangi waktu analisis sekaligus meningkatkan cakupan bukti yang teridentifikasi. Nilai *recall* yang lebih tinggi menunjukkan bahwa model mampu mendeteksi lebih banyak item relevan yang berpotensi terlewat dalam analisis manual, sedangkan *precision* yang sedikit lebih rendah menunjukkan perlunya verifikasi untuk mengeliminasi sejumlah kecil *false positives*. Dari sudut pandang praktis, *trade-off* ini dapat diterima karena waktu yang dihemat jauh lebih besar dibandingkan upaya tambahan yang diperlukan untuk memvalidasi sejumlah kecil deteksi yang kurang akurat. Dengan demikian, integrasi LLM memberikan manfaat bersih yang positif bagi proses investigasi.

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Dari sisi kelengkapan temuan, skenario DeepSeek-R1 mencapai recall sekitar 90%, yang berarti sebagian besar item bukti relevan berhasil diidentifikasi. Nilai ini lebih tinggi dibandingkan analisis manual (≈75%) maupun skenario berbantuan ChatGPT (≈80%). Hal ini menunjukkan bahwa DeepSeek-R1 lebih efektif dalam menangkap pola bukti yang beragam—kemungkinan besar karena pelatihan yang berorientasi pada penalaran dan kemampuan multi-langkah yang menyerupai *chain-of-thought*. Sebagai contoh, model berhasil mengaitkan pesan *chat* tersamar ("kode paket sudah diantar") dengan sebuah transaksi pada *spreadsheet* di tanggal yang sama, yang pada tahap awal terlewat oleh penyidik manusia. *Precision model* juga tergolong tinggi (≈85%), meskipun sedikit di bawah analisis manual (≈95%). Ini menunjukkan bahwa DeepSeek-R1 menghasilkan beberapa *false positives*—misalnya menandai email internal rutin sebagai mencurigakan hanya karena mengandung kata "*invoice*", meskipun konteksnya tidak terkait suap. Meskipun demikian, tingkat *precision* sebesar 85% tetap kuat untuk sistem otomatis, dan peringatan palsu tersebut dapat disaring oleh penyidik. Dibandingkan ChatGPT, DeepSeek-R1 menunjukkan *precision* yang lebih tinggi (ChatGPT ≈70%), yang mengindikasikan bahwa model *open-source* ini lebih terlatih untuk menghindari keluaran menyesatkan dalam skenario forensik yang diuji, kemungkinan karena arsitektur dan proses pelatihannya yang lebih berfokus pada penalaran yang benar. Temuan ini mendukung klaim bahwa DeepSeek-R1 menawarkan kemampuan penalaran dan pemahaman konteks yang mendekati model tingkat atas seperti GPT-4.

Kualitas Bantuan Analisis dan Laporan

DeepSeek-R1 tidak hanya mempercepat identifikasi bukti, tetapi juga membantu dalam penyusunan narasi perkara. Model menghasilkan draf ringkasan kasus sepanjang sekitar tiga halaman yang menguraikan para aktor yang terlibat, kronologi peristiwa berdasarkan timestamps email dan chat, serta transaksi mencurigakan utama. Para ahli menilai draf ini bermanfaat dan memberikan skor rata-rata 4 dari 5 karena struktur yang sistematis dan alur logis. Mereka secara khusus mengapresiasi kemampuan model dalam merangkai hubungan antar bukti menjadi satu alur cerita yang koheren—sebuah tugas yang biasanya memakan waktu signifikan jika dilakukan secara manual. Sebagai perbandingan, laporan awal yang dihasilkan oleh ChatGPT, meskipun lancar secara linguistik, mengandung beberapa ketidakakuratan seperti salah menafsirkan konteks percakapan dan memasukkan asumsi yang tidak didukung data, sehingga hanya memperoleh skor rata-rata 3 dari 5. Hal ini memperkuat pandangan bahwa model *open-source* yang dapat disesuaikan, seperti DeepSeek-R1, lebih sesuai untuk tugas spesifik domain dibandingkan LLM tujuan umum yang tidak diadaptasi untuk domain forensik. Namun demikian, para ahli juga mengidentifikasi keterbatasan dalam draf laporan DeepSeek-R1. Beberapa terminologi hukum digunakan secara kurang tepat; sebagai contoh, model sesekali menyebut "penyuapan" sebagai "hadiah" dalam ringkasannya, sehingga memerlukan koreksi oleh ahli hukum. Selain itu, model cenderung tidak merujuk sumber bukti secara presisi sesuai standar forensik (misalnya tautan eksplisit ke item bukti atau *file hash*), sehingga penyidik harus menambahkan detail tersebut secara manual agar laporan memenuhi persyaratan pembuktian di pengadilan. Keterbatasan ini dapat dipahami, karena model bahasa tidak memiliki konsep bawaan mengenai *chain of custody*—tanggung jawab tersebut tetap berada pada penyidik, yang harus memastikan bahwa setiap temuan dapat ditelusuri kembali ke bukti asli.

Analisis Kesalahan dan Risiko

Beberapa kesalahan model terdokumentasi selama eksperimen. Dalam satu contoh halusinasi, DeepSeek-R1 menyatakan bahwa "terdapat perintah langsung dari atasan untuk mengamankan proyek X," padahal tidak ada pernyataan eksplisit seperti itu dalam dataset—model tampaknya menyimpulkan hal tersebut dari beberapa petunjuk implisit. Meskipun inferensi semacam ini mungkin tampak masuk akal secara logis, dalam standar pembuktian, kesimpulan tersebut harus ditarik oleh penyidik berdasarkan fakta terdokumentasi, bukan disajikan sebagai klaim langsung yang dihasilkan AI. Temuan ini konsisten dengan peringatan OpenAI bahwa GPT-4 dan LLM canggih lainnya dapat berhalusinasi fakta dan melakukan kesalahan penalaran ketika dibiarkan menghasilkan kesimpulan secara otonom (Achiam et al., 2023).

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Oleh karena itu, setiap keluaran dari DeepSeek-R1 harus diverifikasi. Pada kasus tersebut, penyidik menandai pernyataan tersebut dan mencari bukti pendukung. Setelah ditelaah, diketahui bahwa model menggabungkan informasi dari dua email terpisah untuk mensintesis kesimpulannya. Setelah peninjauan, penyidik merevisi laporan dengan membingkai ulang pernyataan tersebut sebagai hipotesis yang memerlukan konfirmasi lebih lanjut, bukan sebagai fakta yang telah ditetapkan.

Isu terkait bias juga teramati. Sebagai contoh, model lebih cenderung menandai komunikasi berbahasa Inggris sebagai "mencurigakan" dibandingkan percakapan informal berbahasa Indonesia. Hal ini kemungkinan berasal dari komposisi data latihnya, yang diduga mengandung proporsi konten berbahasa Inggris dan pola korupsi global yang lebih besar, sehingga model lebih peka terhadap konteks tersebut dibandingkan nuansa lokal. Dalam satu *log chat* berbahasa Indonesia, frasa slang seperti "oke sip, nanti malam dikirim" tidak disorot sebagai signifikan, meskipun percakapan yang lebih luas mengindikasikan kemungkinan transfer uang. Kurangnya sensitivitas terhadap bahasa dan konteks lokal ini dapat dimitigasi melalui *fine-tuning model* pada data kasus korupsi Indonesia, meskipun hal tersebut berada di luar cakupan penelitian ini. Yang jelas, penyidik lokal harus menyadari potensi bias tersebut dan menghindari ketergantungan penuh pada AI ketika menangani konten berbahasa Indonesia. Sebagai gantinya, mereka dapat memperkaya *prompt* dengan penjelasan kontekstual atau menggunakan konfigurasi dwibahasa.

Dari sudut pandang hukum, seluruh pemangku kepentingan sepakat bahwa alat AI seperti DeepSeek-R1 harus diperlakukan semata-mata sebagai alat bantu analitis, bukan pengambil keputusan. Keluaran yang dihasilkan AI tidak dapat diajukan sebagai alat bukti di pengadilan, tetapi dapat digunakan sebagai intelijen awal atau petunjuk investigatif. Oleh karena itu, diperlukan prosedur operasional standar (SOP) untuk mengatur penggunaan AI dalam forensik digital di institusi seperti Komisi Pemberantasan Korupsi (KPK) atau kepolisian. SOP tersebut antara lain harus mencakup: memastikan bahwa hanya salinan bukti, bukan artefak asli, yang diproses oleh AI untuk menjaga integritas pembuktian; pencatatan seluruh langkah analisis berbantuan AI untuk keperluan audit (misalnya mendokumentasikan *prompt* dan keluaran secara terpisah dalam berkas perkara); serta kewajiban bahwa seluruh kesimpulan akhir diverifikasi dan disahkan secara formal oleh penyidik manusia tersertifikasi. Pendekatan ini membantu menjaga chain of custody dan akuntabilitas hukum meskipun AI digunakan sebagai asisten.

Tantangan Penyidik dan Positionalitas Peneliti

Penggunaan AI dalam investigasi sangat dipengaruhi oleh perbedaan literasi digital dan pengalaman profesional antar penyidik. Mereka yang terbiasa dengan alat analisis digital cenderung lebih cepat menafsirkan rekomendasi model, sementara penyidik dengan latar belakang nonteknis memerlukan waktu lebih lama untuk menilai relevansi dan keandalan keluaran. Hal ini memengaruhi konsistensi kualitas analisis ketika AI digunakan sebagai alat pendukung. Dalam penelitian ini, peneliti memiliki pengalaman langsung menangani bukti digital pada kasus korupsi, sehingga memiliki pemahaman kontekstual atas kebutuhan investigatif. Namun, posisi ini juga berpotensi menimbulkan optimism bias terhadap solusi berbasis AI. Untuk memitigasi bias tersebut, desain penelitian mengombinasikan evaluasi kuantitatif dengan diskusi analitis yang berlandaskan keluaran model aktual dari eksperimen.

Implikasi Hukum, Etika, dan Privasi Data

Penggunaan AI dalam investigasi harus mematuhi persyaratan kerahasiaan dan prinsip perlindungan data pribadi sebagaimana diatur dalam Undang-Undang Perlindungan Data Pribadi Indonesia. Bukti digital dalam kasus korupsi sering kali mengandung informasi sensitif yang hanya boleh diakses oleh pihak berwenang. Oleh karena itu, seluruh interaksi dengan model harus dilakukan secara lokal dan dicatat untuk keperluan audit internal. Keluaran AI tidak dapat diperlakukan sebagai bukti mandiri, melainkan sebagai informasi analitis pendukung. Setiap temuan yang dihasilkan model harus dapat ditelusuri kembali ke artefak digital asalnya untuk menjaga integritas *chain of custody*. Dengan demikian, AI harus beroperasi dalam kerangka tata kelola forensik yang ketat.

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Risiko Ketergantungan Berlebihan pada AI

Kemudahan yang ditawarkan sistem AI membawa risiko bahwa penyidik menjadi terlalu bergantung pada rekomendasi model. Tanpa pengamanan yang memadai, penalaran investigatif dapat bergeser menjadi sekadar memvalidasi keluaran AI alih-alih melakukan penilaian kritis secara independen. Hal ini sangat berbahaya dalam kasus korupsi yang menuntut penelaahan bukti berlapis. Oleh karena itu, mekanisme *human-in-the-loop* dan prosedur yang mewajibkan verifikasi manual atas seluruh kesimpulan kunci menjadi esensial. Pelatihan literasi digital juga merupakan prasyarat agar penyidik memahami keterbatasan model, mencegah ketergantungan berlebihan, dan mempertahankan otonomi analitis sepanjang proses investigasi.

Perbandingan dengan Pustaka Terkait

Temuan penelitian ini konsisten dengan literatur sebelumnya yang menunjukkan bahwa LLM dapat berfungsi sebagai alat pendukung yang berguna di tangan pengguna yang kompeten. Studi terdahulu yang mengevaluasi GPT-4 dalam konteks forensik menemukan bahwa model tersebut dapat membantu berbagai tugas seperti pencarian artefak, pembuatan *skrip carving*, dan deteksi anomali log, namun tetap mengharuskan pengguna menyadari kapan AI melakukan kesalahan (Scanlon et al., 2023). Pola serupa juga diamati dalam penelitian ini: DeepSeek-R1 secara efektif mempercepat tugas rutin dan mengungkap pola, tetapi keahlian penyidik tetap krusial untuk memvalidasi hasil.

Penelitian lain melaporkan bahwa meskipun LLM dapat menyusun laporan forensik, model tersebut tidak dapat sepenuhnya menggantikan penulis laporan manusia karena masalah akurasi (Michelet & Breitinger, 2024). Temuan kami mendukung kesimpulan ini—draf DeepSeek-R1 bermanfaat tetapi tetap memerlukan penyempurnaan manual. Keunggulan DeepSeek-R1 dibandingkan ChatGPT, khususnya dalam hal privasi data dan kemampuan kustomisasi, sejalan dengan argumen bahwa model lokal atau *open-source* mengurangi kebutuhan untuk mengunggah bukti sensitif ke *cloud*. Penelitian sebelumnya menekankan pentingnya menjalankan LLM dalam lingkungan terkontrol ketika digunakan sebagai *investigative copilot* (Henseler & van Beek, 2023). Hal ini selaras dengan pendekatan *on-premise* yang diadopsi untuk DeepSeek-R1 dalam studi ini.

Menariknya, pendekatan AI multi-agen juga mulai dieksplorasi dalam domain forensik. Penelitian terbaru memperkenalkan kerangka *intelligent agent* dimana beberapa agen berkolaborasi dalam investigasi digital (Wickramasekara & Scanlon, 2024). Ke depan, DeepSeek-R1 berpotensi diintegrasikan ke dalam ekosistem multi-agen, misalnya bekerja bersama agen AI khusus untuk forensik citra atau analisis jaringan, guna menghasilkan investigasi yang lebih holistik. Dalam konteks penegakan hukum antikorupsi, hal ini dapat berarti satu agen LLM berfokus pada dokumen keuangan sementara DeepSeek-R1 memusatkan perhatian pada komunikasi tekstual, dengan pertukaran informasi yang konsisten untuk memastikan tidak ada bukti yang terlewat. Namun, koordinasi multi-agen semacam ini memerlukan mekanisme kontrol yang kuat untuk menjaga akuntabilitas dan mencegah akumulasi kesalahan.

Diskusi Implikasi

Secara keseluruhan, integrasi DeepSeek-R1 menunjukkan nilai tambah yang jelas dalam investigasi forensik digital pada kasus korupsi. Kecepatan dan cakupan analisis berbantuan AI dapat mengurangi penumpukan perkara dan membantu penyidik menemukan "jarum dalam tumpukan jerami" di tengah dataset yang sangat besar, sehingga mempercepat penegakan hukum antikorupsi. Hal ini berpotensi meningkatkan kinerja institusi seperti KPK dalam mengungkap perkara secara lebih proaktif.

Namun demikian, temuan penelitian ini juga menegaskan bahwa AI bukanlah solusi ajaib yang dapat beroperasi secara otonom tanpa pengawasan. Prinsip *human-in-the-loop* tetap tidak dapat ditawar. Dari perspektif hukum, regulasi yang mengatur penggunaan AI dalam proses pro justitia perlu dikembangkan—misalnya untuk menentukan apakah langkah analisis berbantuan AI harus diungkapkan di pengadilan sebagai bagian dari metodologi investigasi demi transparansi. Kompetensi penyidik dalam bidang TI dan AI juga perlu diperkuat agar mereka mampu memanfaatkan sekaligus menilai secara kritis hasil yang dihasilkan AI.

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Pertimbangan etis juga sama pentingnya. DeepSeek-R1 dan LLM sejenis dilatih pada data skala internet yang dapat mengandung bias budaya tertentu. Ketika diterapkan pada kasus korupsi di Indonesia, model dapat menunjukkan bias—misalnya lebih mencurigai terminologi yang terkait sektor bisnis tertentu karena data latih yang didominasi kasus Barat. Hal ini perlu diantisipasi melalui *fine-tuning* lokal dan evaluasi berkala terhadap bias keluaran AI. Di sisi positif, sifat *open-source* DeepSeek-R1 memungkinkan komunitas yang lebih luas berkontribusi pada perbaikan model, berbeda dengan sistem komersial tertutup.

Simpulan

Penelitian ini menunjukkan bahwa model DeepSeek-R1 dapat secara signifikan mempercepat identifikasi awal bukti digital dalam penyidikan kasus korupsi. Model ini membantu penyidik dalam mendeteksi pola, menafsirkan makna kontekstual komunikasi, dan mengidentifikasi dokumen yang relevan. Dibandingkan dengan analisis manual, DeepSeek-R1 mempercepat proses penyaringan dokumen dan mengurangi risiko terlewatnya informasi penting pada tahap awal penyidikan. Dibandingkan metode analitis tradisional, penggunaan DeepSeek-R1 memperpendek waktu identifikasi bukti dan meningkatkan *recall*. Dengan kata lain, AI dapat mengotomatisasi tugas rutin yang memakan waktu, sehingga penyidik manusia dapat berfokus pada penalaran tingkat tinggi dan pengambilan keputusan strategis. Temuan ini sejalan dengan tren global yang menunjukkan bahwa adopsi *large language models* dalam forensik digital dapat meningkatkan efisiensi investigasi serta mengurangi hambatan teknis dan prosedural hukum.

Meskipun demikian, penggunaan DeepSeek-R1 harus disertai dengan pengawasan, tanggung jawab, dan kehati-hatian yang jelas karena adanya risiko bias, salah tafsir, potensi pelanggaran privasi, serta kemungkinan ketergantungan berlebihan penyidik terhadap rekomendasi model. Penelitian ini menegaskan perlunya penyidik manusia untuk memverifikasi keluaran AI dan memberikan konteks hukum yang tepat. DeepSeek-R1 bukanlah pengganti penyidik, melainkan alat pendukung cerdas. Setiap keluaran yang dihasilkan AI harus diperlakukan sebagai informasi tambahan yang memerlukan konfirmasi. Kami merekomendasikan agar aparat penegak hukum, termasuk Komisi Pemberantasan Korupsi (KPK), menyusun pedoman formal dan program pelatihan sebelum menerapkan sistem AI semacam ini. Keamanan data juga merupakan aspek krusial; meskipun sifat *open-source* DeepSeek-R1 memungkinkan *deployment* lokal yang aman, prosedur penanganan data harus tetap mengikuti prinsip forensik untuk menjaga integritas bukti. Temuan penelitian ini juga menyoroti perlunya pengembangan prosedur operasional standar (SOP) penggunaan AI dalam investigasi, penguatan kapasitas penyidik melalui pelatihan literasi digital, serta penelitian lanjutan mengenai integrasi AI dalam praktik hukum di Indonesia. Penggunaan AI harus tetap selaras dengan kerangka hukum dan etika penyidikan serta tidak boleh mengorbankan kerahasiaan maupun akuntabilitas penegakan hukum.

Dari perspektif pengembangan, hasil penelitian ini membuka beberapa peluang untuk penelitian selanjutnya. Pertama, *fine-tuning* DeepSeek-R1 atau model sejenis secara khusus pada dataset kasus korupsi dan berbahasa Indonesia berpotensi meningkatkan sensitivitas model terhadap konteks lokal. Kedua, penelitian mengenai *explainable AI* dalam forensik perlu dikembangkan, misalnya melalui mekanisme yang memungkinkan LLM menjelaskan dasar dari keluarannya atau menyoroti bukti pendukung, sehingga mempermudah proses verifikasi dan meningkatkan kepercayaan penyidik serta hakim. Ketiga, uji lapangan terkontrol pada kasus nyata diperlukan untuk mengevaluasi kinerja AI dalam lingkungan operasional, termasuk dampaknya terhadap alur kerja tim investigasi. Keempat, kolaborasi multidisipliner antara ahli forensik digital, pakar hukum, dan ilmuwan data menjadi penting untuk mengembangkan kerangka regulasi dan etika penggunaan AI dalam penegakan hukum.

Sebagai penutup, integrasi sistem AI seperti DeepSeek-R1 mencerminkan perkembangan teknologi yang tidak terelakkan dalam upaya pemberantasan korupsi. Dengan implementasi yang bertanggung jawab dan berbasis bukti, teknologi ini berpotensi menjadi sekutu yang kuat dalam mengungkap kejahatan kerah putih yang semakin kompleks. Pemberantasan korupsi di era digital menuntut inovasi, dan DeepSeek-R1 menunjukkan bagaimana kecerdasan buatan dapat meningkatkan kapasitas investigatif penyidik manusia.

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Referensi

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., & Anadkat, S. (2023). Gpt-4 technical report. In *arXiv preprint arXiv:2303.08774*. Cornell University. https://doi.org/10.48550/arXiv.2303.08774

Dunsin, D., Ghanem, M. C., Ouazzane, K., & Vassilev, V. (2024). A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response. *Forensic Science International: Digital Investigation*, *48*, 301675. https://doi.org/10.1016/j.fsidi.2023.301675

Guo, D., Yang, D., Zhang, H., Song, J., Zhang, R., Xu, R., Zhu, Q., Ma, S., Wang, P., & Bi, X. (2025). Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. In *arXiv preprint arXiv:2501.12948*. https://doi.org/10.48550/arXiv.2501.12948

Henseler, H., & van Beek, H. (2023). ChatGPT as a copilot for investigating digital evidence. *Proceedings of the Third International Workshop on Artificial Intelligence and Intelligent Assistance for Legal Professionals in the Digital Workplace (LegalAIIA 2023)*, 58–69.

Kadage, A. D., Nayak, B. M., Vishal Sharad Hingmire, Wanjale, K., Bogiri, N., & Mandale, P. L. (2024). AI-enhanced digital forensics: Automated techniques for efficient investigation and evidence collection. *Journal of Electrical Systems*, *20*(1s), 211–229. https://doi.org/10.52783/jes.766

Michelet, G., & Breitinger, F. (2024). ChatGPT, Llama, can you write my report? An experiment on assisted digital forensics reports written using (local) large language models. *Forensic Science International: Digital Investigation*, *48*, 301683. https://doi.org/10.1016/j.fsidi.2023.301683

Scanlon, M., Breitinger, F., Hargreaves, C., Hilgert, J.-N., & Sheppard, J. (2023). ChatGPT for digital forensic investigation: The good, the bad, and the unknown. *Forensic Science International: Digital Investigation*, *46*, 301609. https://doi.org/10.1016/j.fsidi.2023.301609

Wickramasekara, A., & Scanlon, M. (2024). A framework for integrated digital forensic investigation employing autogen AI agents. *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*, 1–6. https://doi.org/10.1109/ISDFS60797.2024.10527235

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)

Copyright © 2025, Integritas: Jurnal Antikorupsi
2615-7977 (ISSN Online) | 2477-118X (ISSN Print)