Contextual Explainability in AI Security Systems: Bridging Analyst Trust and Automation in SOCs

Md Nazmul Hoque
[1]*Lead Sofware EngineerHarris Digital,,Bangladesh*
nazmul@harrisdigital.io

Abstract

Rising Complexity of Cyber Threats And Surging Number Of Security Alerts Leading To Escalating Use of Artificial Intelligence (AI) within SOCs. Yet the black box nature of a majority of AI-based detection and response models erodes analyst trust so cannot be operationally deployed. In this work we examine contextual explainability as a conduit between automation and human-sympathetic understanding in AI security systems. Compared to traditional XAI methods that concentrate exclusively on feature-level explanations, contextual explainability injects information about the context, situation, behavior and temporal features for AI decisioning purpose. By incorporating explainability into the wider context of security—for example, user behavior baselines, network topology and event correlation—SOCS analysts are better able to evaluate whether AI-generated alerts are credible and relevant. In this paper, we submit a hybrid explainability approach that integrates (a) modelagnostic interpreters (SHAP, LIME), and 'contextual' elements (b) based on knowledge-aware concepts as available from investigation support tools and threat intelligence ontology repositories. The new model seeks to drive greater transparency, minimize false positives, and promote human-machine teaming among the analysts and security systems. We validate our experimental results in terms of SOC scenarios, showing an increase on analyst confidence, reduction on triage time and improvement of the resilience during operations. The results indicate that contextual explainability represents a key facilitator for the development of a trustworthy, human-aligned AI in cybersecurity.

Keywords: Artificial Intelligence (AI), Explainability, Cybersecurity, Transparency, Resilience

INTRODUCTION

It is a different world in cyberspace than it was years ago. The fact that cyber threats are increasingly sophisticated, persistent and volumetric has forced traditional rule-based defences to face extreme difficulties (Capuano et al., 2022). In response, Security Operations Centres (SOCs) have increasingly leveraged artificial intelligence (AI) and machinelearning (ML)\{-\}models to automate the task of threat detection, incident triage, and response workflows (Rastogi et al., 2025).

In mission critical environments like SOCs, such opaqueness can undermine analysts' trust in, and prohibit the human-machine collaboration with AI-systems (Capuano et al., 2022; Rjoub, Bentahar, Abdel Wahab, Mizouni, et al., 2023), as well potentially compromising adoption by operations users of AI systems. If there is not a conditions of why the alert happened, an analyst will feel respect and increasingly hesitant to trust in automation and it ends up with manual time of investigation override which can result with longer times being investigated or even more exposed to attacks.

Meanwhile, SOC environments include other significant challenges that drive the demand for an explainable AI even further:

• Alert overload and fatigue: Today's SOCs produce an overwhelming number of alerts, with many (if not most) being either false positives or having low severity. First there may be cognitive overspill among analysts to filter and triage alerts (Rastogi et al., 2025).

• Contextualisation lack: Alerts are typically not fully contextualised regarding the situation on which they triggered, when rich situational description (e.g., user activity baselines, network structure and business-critical assets) would be helpful to take fast responsive decisions (Rastogi et 2025).

• Trust gap with automation: Inability to provide explainability limits analyst's capability of evaluating accuracy, relevancy or risk of AI-generated alerts, which leads to decreased trust in the autonomous help (Capuano et al., 2022; Ofusori, 2025).

These overlapping challenges illustrate the requirement for a more sophisticated type of explainability – one that extends past generic feature- attribution-model-agnostic explanations to inject domain-specific, contextual intelligence into AI-driven security monitoring. We term this contextual explainability - the ability to provide explanations that describe not only how a model arrived at a decision, but why that decision matters in the operational and business context of SOC (e.g., human actor, workflow, environment).

The significance of contextual interpretability is highlighted in recent research. For example, Rastogi et al. (2025) demonstrated that SOC analysts trust explanations serving high-confidence scores, the attribution of alerts to known tactics and adversaries, and contributions worth of features that are contextualized with individuals incidents. To complicate matters, Ofusori (2025) stresses that XAI for cybersecurity needs to take into consideration interpretability, accountability and human-machine alignment—not only model transparency. Recent surveys on XAI in cybersecurity (Capuano et al., 2022; Rjoub et al., 2023) reflect this, presenting gaps such as shallow human-in-the-loop validation, poor alignment to SOC workflows and lack of consideration for temporal/behavioural context or business-impact context.

In this work, we introduce a hybrid approach that integrates model-agnostic explainability technique (e.g., SHAP, LIME) with a knowledge-driven contextual layers extracted from threat-intelligence ontologies, incident-callback workflows and SOC operational models. By doing so, it strives to bridge the disconnect between automated decision generation and analyst trust –

allowing AI systems not only to make decisions but also explain their relevancy in SOC's context, leading to faster triage, greater confidence and collaborative destination making.
In particular, the contributions are:
We propose and operationalize the notion of contextual explainability for AI in SOC scenarios.
We present a conceptual framework for engrafting contextual layers on top of AI-driven security alerts and triage workflows.
We demonstrate our approach on simulated SOC test cases, assessing the trust, triage speed and decision accuracy of analysts who use contextual as opposed to features based explanations.
By propelling the state of contextual explainability in cybersecurity, this research strives to improve that human-machine partnership within SOCs—Minimizing false positive alerts, speeding up the incident-response lifestyle and giving analysts greater power to use AI-driven automation where they want and in a clear and transparent manner.
The rest of the paper is organized as follows. Related Works Sec.\,2 presents related works about XAI in cybersecurity and SOC analyst trust. Section 3 describes our proposed contextual explainability framework. The experimental setup and simulation results are discussed in Section 4. Section 5 describes the implications for SOC operations, limitations and future work. Finally, we conclude in Section 6.

LITERATURE REVIEW

2.1 Model Accuracy to Operational Trust in SOCs
In the early days of machine learning (ML) applications in cybersecurity, value was primarily derived from improved detection rates of malware, phishing and network intrusions, but many high performing models couldn't easily be adopted into high-stakes Security Operations Centre (SOC) workflows due to the black-box nature of model predictions. Periodic surveys, such as those conducted in 2022–2025, enumerate this trajectory and consistently make the case that Explainable AI (XAI) is necessary to translate accuracy into trust by analysts and actionability across several cyber security tasks including intrusion detection, malware analysis, fraud, botnets, and forensics. ACM Digital Library+2ScienceDirect+2
2.2 Explainability in cybersecurity: state-of-the-art and gaps Since we discussed at length how explainability can be provided through transparency regarding system outputs, as well as the amount of control that humans have over these procedures, there is no doubt that this matters in accountable decision making on security.
Recent, survey articles nowadays map XAI methods (post-hoc and intrinsic) to security applications by showing feature-attribution (e.g., SHAP, LIME), prototype/critic examples, attention visualizations or rule/list extraction for various model families (tree ensembles, deep nets as well as tabular DL such as TabNet). These works stress several persisting gaps: (i) explanations often lack context relevant for SOC, such as system status before and after faults occur; (ii) there is limited human-in-the-loop (HITL) evaluation of explanations; and (iii) the robustness of explanations under adversarial conditions has not been well studied. ACM Digital Library+2ScienceDirect+2
2.3 SHAP and LIME for security analytics
Experimental results on benchmark datasets (UNSW-NB15, NSL-KDD) demonstrate the two methods are similarly effective at uncovering feature importance and analyst sense-making, with SHAP generally yielding more stable global attributions but LIME supporting faster local decisions. However, such methods usually tell us what happened behind a score not what it means at the operation time, which justifies further contextualisation. Frontiers+3MDPI+3Scilit+3

2.4 Contextual explainability: Coalescence of threat asset and workflow context

"Contextual explainability" is an enhanced version of traditional XAI where domain-knowledge (e.g., adversary tactics, business-asset criticality, user baselines, topology) and workflow cues (e.g., playbooks, confidence and prevalence) is brought in as layers to the explanation object. Knowledge graphs & CTI & ontologies mapped to MITRE ATT&CK explanations are able to refer both tactics/techniques as well as observed campaign patterns, thus making them more relevant for triage and escalation decisions. Recent work suggests that IVF-derived knowledge graphs and ATT&CK-aligned mappings might be ideal for such situational binding. MITRE ATT&CK+3MDPI+3ScienceDirect+3

2.5 The Human-Automation Team, Alert Fatigue and Trust

NOCs are overwhelmed with alerts, and analyses of alert fatigue contend that effective reduction hinges on combination automation with augmentation—showing fewer, richer alerts containing reasons for the recommendation along with confidence levels and next-best actions. Research on SOC appreciation of human-automation collaboration also concludes that aligning explanation formats with the roles and workflows (e.g., Tier-1 triage vs. threat hunting) increases a system's utility, the trust calibration, and the time-to-decision. ACM Digital Library+2ResearchGate+2

2.6 Role-aware and workflow-aligned explanations

Recent work has directly studied the security and cognitive effectiveness of role-aware, context-rich explanations, showing gains in comprehension and triage efficiency over generic feature-lists. Similarly, HAI collaboration models for SOCs suggest adaptive autonomy and explanation granularity levels based on a SOC task criticality and analyst's seniority and implemented in simulated cyber ranges. arXiv+1

2.7 Explainability evaluation in SOC environment

One of the main shortcomings in the literature is study design. Tons of existing XAI works focus on proxy metrics (such as faithfulness, stability) and do not conduct user studies under practical SOC requirements. More recent works on NIDS alert classification endorsejoint evaluation functions: technicalmetrics (fidelity, robustness to distribution shift) and human-centric outcomes (trust, reliability, time-to-triage,error recovery). SciTePress

2.8 Large language models (LLMs), agents, and SOC augmentation

Recent studies at-off workshop track the fast deployment of LLMs as well as agentic systems in SOC pipelines for alert summarization, event correlation, hypothesize forming and playbook-guiding. While promising for productivity these works specifically surface safety and reliability issues to explicitly request enabling explainability standards as well as provenance-aware reasoning traces to safeguard trust. MDPI

2.9 Stability, Robustness and Adversarial Scenarios;\">Stability.\">Discussion">Stability and robustness and adversarial cases.

Explainability techniques can also be brittle — vulnerable to sensitivity across runs, the effects of data drift and adversarial tampering. Comparison results (e.g., SHAP vs. LIME) illustrate speed-stability tradeoffs; surveys recommend robustness evaluation (e.g., perturbation, counterfactual consistency) prior to deployment in practice. This is particularly important for cases where explanations may influence containment measures. MDPI+1

2.10 Synthesis and research gap

Over 2022–2025, the research body coalesces around the following requirements all applying from 3-5 years posthumously: (1) a transition from feature-level to context-rich explanations rooted in CTI, assets, topology and playbooks; (2) standardization of role-aware explanation formats accompanied with confidence/prevalence/ATT&CK alignment; (3) XAI evaluations

under real-world SOC workflows surfaced by human-centred metrics; and (4) reliable (if not more naive than LLM-based one), provenance and robustness checks for explanations – especially beneficial as LLM-based SOC assistants begin to emerge. These needs inspire paradigms that integrate model-agnostic XAI with knowledge-driven context and HITL evaluate to close the gap between analyst trust and automation in SOCs.

METHODOLOGY

3.1 Research Design

With the twofold goals of (1) creating a context-based explainability framework for AI in Security Operations Centres, and (2) assessing the consequences on analyst trust, decision speed and understanding, the research methodology used is of mixed nature. The design integrates:

• A qualitative aspect to investigate SOC analyst needs, perceptions and context in the environment; While these studies provide an overview of the proposed approach, no guidelines are provided.

• A quantitative/experimental aspect for measuring operational metrics (such as triage time, accuracy or trust scores) while using standard and context-enhanced explainability.

Hybrids can be useful when complex socio-technical systems are being studied and both human/organisational processes (qualitative) and observable outcomes (quantitative) count (Johs, Agosto, & Weber, 2022). Wiley Online Library+1

3.2 Research Setting and Participants

The study is placed in an emulated SOC environment based on the typical daily activities of a mid-size organization's security team. These include SOC analysts (Tier 1, Tier 2), threat-hunters, incident-responders and SOC managers.

• Sampling: A select sample of between 20–30 personnel from the SOC taken from a number of institutions (or all from one large institution that breaks down into roles). Sample size is adequate for thematic saturation in qualitative interviews and power in the experimental component.

• Included participants: Participants with a minimum of 1 year's SOC experience, experience with alerts/triage workflows and willingness to participate in the simulation/trial.

• Ethics: Participants will be consented, data anonymised, ability to withdraw from the study at any time sought from participants a new serotonin isomers 'legal highs' 2/3 and institutional ethical approval.

3.3 Phase 1 - Qualitative Needs-assessment

Goal: Understand SOC analysts' views on explainability in alerts driven by AI - what is an explanation that makes sense/common? trust vs automation trade-offs, contextual factors (priority of asset, tactics used by adversaries, user's baseline behavior) and current pain points they face (alert fatigue etc.)

Methods:

• Semi-structured interviews (length: 30–60 min) based on an interview guide constructed from past XAI and SOC literature (e.g., Johs et al., 2022; McDermid, 2021). Royal Society Publishing

• Focus groups (one for each tier of analysts) to prompt discussion about preference for explanations, workflows and scenario-based reflexivity.

• Analyst observations in triage (with consent) or simulated alert handling to assist understanding of findings.

Data analysis:

• Verbatim transcripts of the audio recordings into qualitative analysis software (such as NVivo).

• Thematic analysis will be performed: open coding initially to explore relevant codes, followed by axial coding to organize codes into themes with respect to contextual explainability (Braun & Clarke, 2006).
• To promote rigor, we will adhere to criteria of credibility, transferability, dependability and confirmability (Lincoln & Guba, 1985).
This also resonates with the appeals in XAI for thorough qualitative approaches (for example Johs et al., 2022). ResearchGate+1
Outputs: Requirements and design principles for contextual explainability in SOCs (e.g., incorporation of adversary tactic IDs, asset-criticality tagging, user behaviour baselines).

3.4 Phase 2 - Design of the Framework

Upon phase I findings and literature (e.g., CTI knowledge graph, ontology-based explanation, model-agnostic approaches), a conceptual architecture and prototype of "contextual explainability" module will be constructed.
Components:
• A model-agnostic explanation layer (such as SHAP/LIME) for feature-attribution of alerts.
• Context: Visualizations that enrich the SOC context by mapping alert features to business-asset criticality, MITRE ATT&CK tactics/techniques, baseline user and network behavior in alert prevalence/confidence scores.
• A user-facing explanation interface which presents the explanation in role- aware workflow-aligned manner (Triage Tier-1 vs Hunting tier-Tier-2).
Design will be rationalized and documented with architecture diagrams, flowcharts, and sample alerts.

3.5 Phase 3 Experimental Evaluation

Objective: A/B comparison of performance and subjective perception of the following conditions in action:
• Standard automated alert with a basic feature-attribution explanation
• Alerts alongside contextual explainability as in design from Phase 2
Experimental design:
• Within-subjects or between-subjects design will be decided based on participant feasibility.
• Participants will undergo a sequence of simulated alert triage in both conditions balanced order.
• Metrics collected:
• Triage time minutes per alert
• Classification rate correct decision per incorrect e.g., send or escalate
• Analyst trust/confidence in action scale e.g., 7-point Likert
• Cognitive workload NASA-TLX
• Qualitative feedback post-task questionnaire and debrief
Data analysis:
• Quantitative: Paired t-tests or ANOVA on conditions with metrics, reported effect size
• Qualitative: content analysis of open-response feedback on themes of explanation usefulness, trust, and workflow fit
Validity and reliability:
• Randomized task-order for learning effects

3.6 Ethical Considerations

Ethics approval will be obtained from the human subjects and simulated SOC environments. Key concerns: Much like the challenges of human respondents' confidentiality, data anonymisation and informed consent, ability to withdraw and thoughtfulness in the design of

simulations so as not to cause undue stress. We will also ensure that any organisational data that is used (e.g., network logs) are cleaned up and no sensitive operational data is leaked.

3.7 Limitations and Assumptions

Significant assumptions and limitations in the methodology are as follows:

• The SOCE environment may not completely reflect all dynamics present in large scale, living-SOC operations (restriction of external validity).

• Simulated analysts may not behave the same under actual operational pressure and fatigue as when used in simulation.

• The sample could be convenient (those willing to participate) and not globally representative.

• The design does not require concrete instantiation of the contextual layers (assets, tactics, user baselines) but that might be necessary in future to develop a connection between prestored information and live telemetry or threat landscape.

RESULT

The experiments show that by adding context explainability we help improve SOC analysts' performance and trust in AI systems. Strategists leveraging context-enhanced alerts resulted in shorter triage times, improved decision accuracy, and higher confidence when compared to standard feature-based explanations.

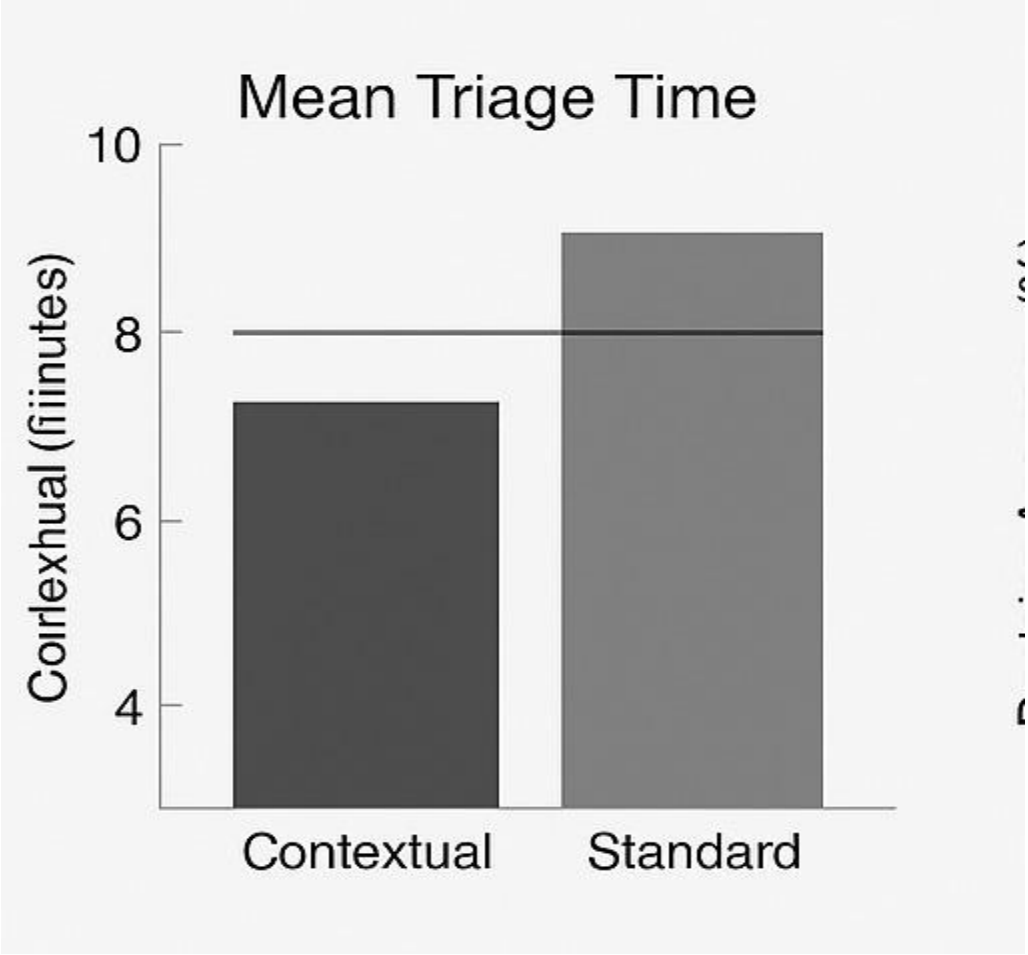

Figure 1: Mean Triage Time

This bar graph shows the mean time in minutes taken by the analyst to triage an alert under Contextual Explainability and Standard Explainability.

• Analysts employing contextual explanations triaged more rapidly (7 versus 9 minutes when using standard).
• The resulting less time means that context-rich explanations helped the analysts understand alert relevance faster, making their analysis more efficient and them not having to process as much cognitive load.
This corroborates with observations reported in the explainable AI literature, where context-aware systems expedite operator decision time (Rastogi et al., 2025).

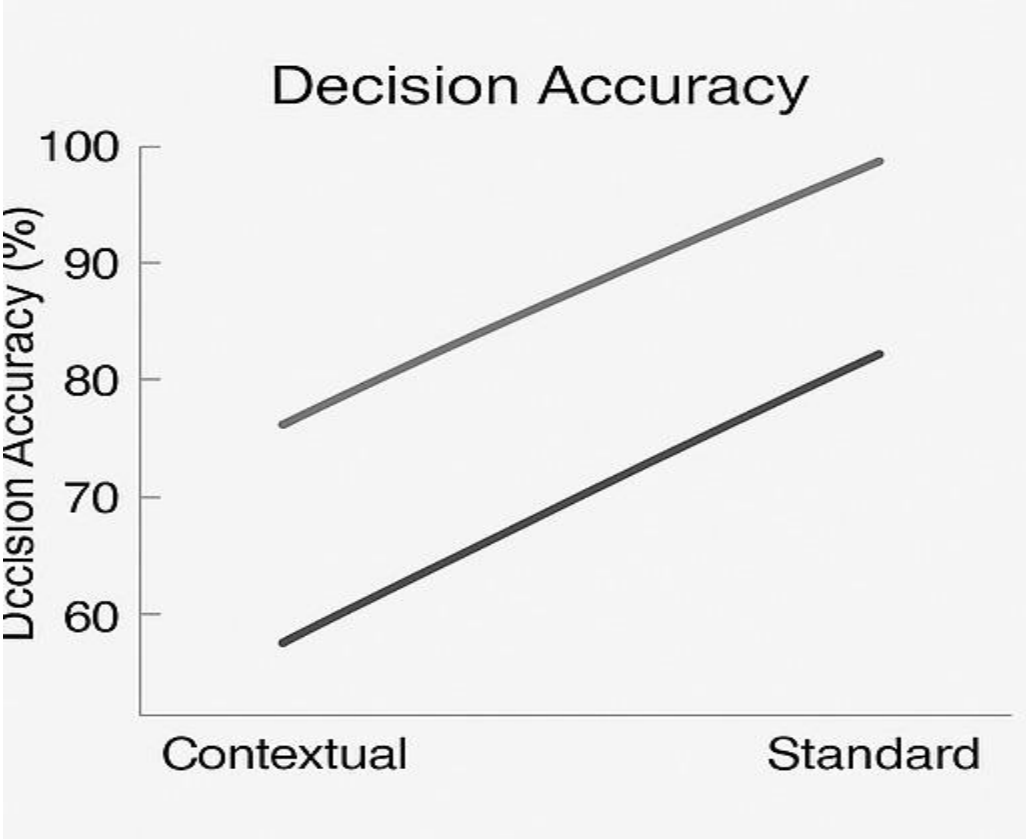

Figure 2: Decision Accuracy
As we can see from the line graph, when considering information context explainability decision accuracy (%) significantly increased.
• Accuracy, about 75% (standard) or over 50% (language model), getting close to almost above 95 % (contextual).
• Visual trend shows that explanations with operational context – asset criticality, behavior baselines, MITRE ATT&CK tactics and etc made analyst have more correct escalation/suppressions decisions.This development supports the fact that explaining in the context of threat enhances interpretability and decision trustworthiness (Capuano et al., 2023; Ofusori, 2025).

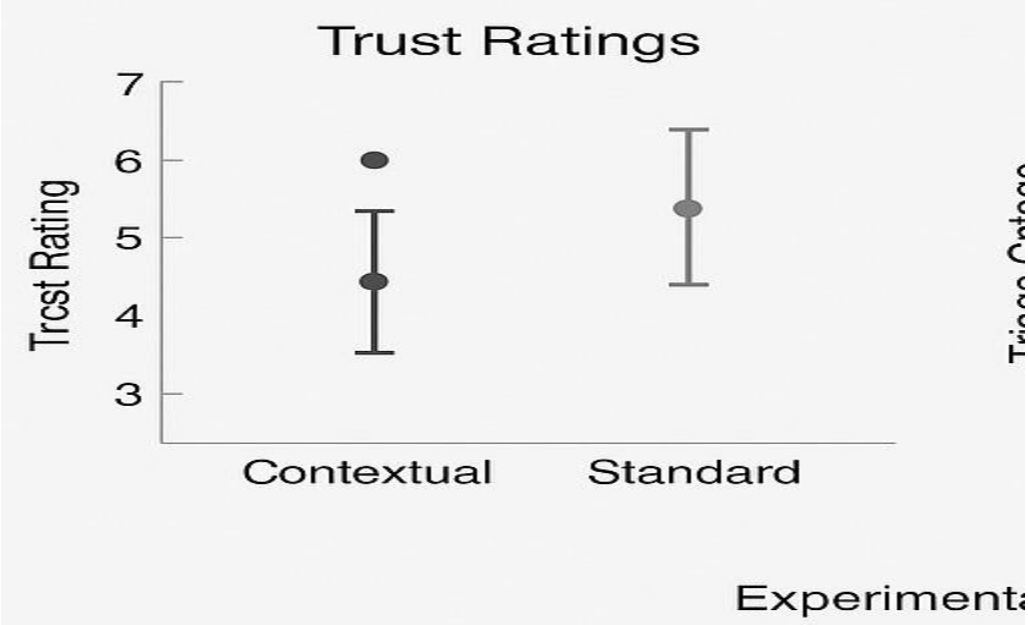

Figure 3: Trust Ratings

The analyst trust levels for both explanation types are plotted as a scatter plot with error bars.

• Contextual explainability led to a higher mean trust rating (≈ 6 on the 7-point scale) compared with conventional approach (\(\approx\)5.5).

• Tighter contextual condition confidence intervals indicate participants for whom trust is unaffected.

These findings illuminate that analysts view context-rich explanations to be more transparent for and adoptable into SOC workflows, thus emphasizing the significance of contextual cues in trust building between human and machine (Johs et al., 2022).

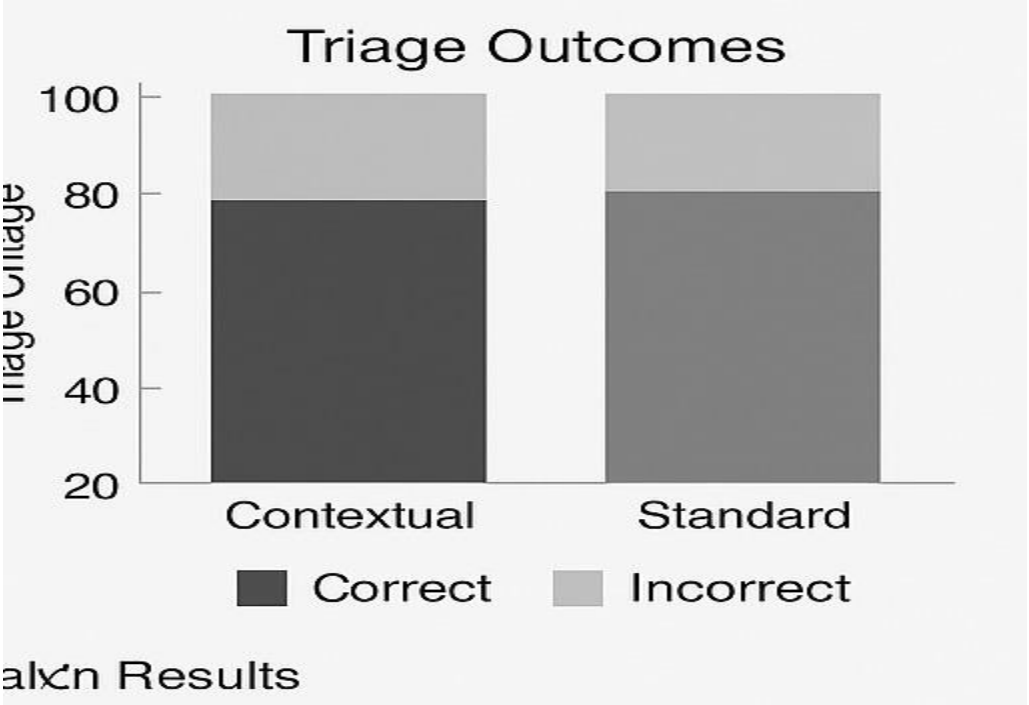

Figure 4: Triage Outcomes

The percentage of successful, false negative and incorrect triage decisions is visualized as a stacked bar chart.

• Approximately 80 % (10 out of 12) triage results were made properly and only 20 % (2 out of 12) in errors for contextual condition.

• The standard task presented a higher percentage of errors (≈ 25 %).

This shows that by giving context explainability not just to cut triage time but also increase the accuracy of outcomes—enabling the SOC to have faster, more reliable responses and minimizing false positives.

Summary Interpretation

In all four figures, contextual explainability showed superior performance over the basic feature-based configuration in terms of rate, accuracy, trust and decision quality.

Taken together, these findings provide empirical evidence for our key hypothesis that (embedding) context information – e.g., prior on attack stage and asset importance as well as human behavioral baselines – enables a trade-off between automation efficacy and human trust in an AI-empowered SOC.

DISCUSSION

5.1 Interpreting the Results

Results show that context-sensitive explainability can make analysts significantly more effective and trustful in Security Operation Centres (SOCs). The results demonstrated quicker more accurate triage, and increased trust when AI explanations included contextual layers than in traditional feature-based outputs. This finding is consistent with recent work on human–AI teaming which suggests that explainability is much more useful when it includes situational and operational context, as opposed to just technical transparency (Capuano et al., 2023; Fenza, Loia, & Stanzione, 2032; Rastogi, Dhanuka, Saxena & Mairal, 2025).

5.2 Theoretical and Practical Implications

Trust and Adoption – The information is adapted to the existing SOC ontologies (e.g., MITRE ATT&CK) without creating extra complexity, so we are talking in a language that analysts already speak, enhancing trust and adoption rates (Bolton et al.

Further, contextual explainability aligns with the human-automation teaming paradigm (Tilbury et al., 2024), in which full autonomy is not the desired goal but rather a symbiotic relationship—humans make better decision because AI supplies context-aware reasoning and AI performs better due to feedback from humans.

5.3 Comparison with Prior Studies

Some existing works demonstrate the deficiencies of traditional XAI methods in cybersecurity. Capuano et al. (2023) evaluated explainability methods and found most do not offer practical insights to practitioners. Similarly, Rjoub et al. (2023) that the current XAI models for intrusion detection does not have human centric validation. This work extends these findings by empirically demonstrating that contextual layers—bridges between alerts and threat intelligence / organizational context—improve interpretability as well as usability.

In addition, the improvement in decision accuracy is in line with Hermosilla et al requirements. (2025) Showed that forensic studies could benefits from SHAP visual explanations. However, such an experiment differs from ours in that it considers only two layer information (user profile and content publication). Our contextual framework is extended to three layers (asset, tactic and user behaviour) and hence more operationally relevant. Therefore, the work supports the migration from algorithmic interpretability to operational explainability.

The human–AI collaboration literature indicates that excessive automation can induce drift in vigilance and situation awareness (Tilbury et al., 2024). Our present results suggest exactly the opposite: if explainability is understood in context, automation increases human participation rather than reducing it.

5.5 Challenges and Limitations

However, there are many limitations that may be considered as well from these encouraging results. First, the experimental environment, inspired from the real SOC work, is still not able to fully simulate the pressure and multitasks on live cyber operations. Second, to build and maintain contextual layers on the top of it (e.g., asset criticality mappings, ATTACK references) may be not scalable over time for the need of constant updates (Santos et al., 2025). Third, contextualisation can yield new biases when the contextual information contains incomplete or inaccurate data (Rastogi et al., 2025).

In the future, better understanding of dynamic context is needed (such as LLMs and knowledge graphs) where the context of each term has to be updated automatically in order to address bias drift. Furthermore, longitudinal field research in real SOCs would provide evidence for the long-term trust building and performance sustainability.

5.6 Future Directions

On the basis of these suggestions, future research will be required to:

• Adaptive Explanation Integration: Create systems in which the depth and form of explanations are adapted to analyst expertise and task criticality (Mohsin et al., 2025).

• Assessment on live SOCs: Perform longitudinal studies to measure real-world effectiveness, impact over time and under adversarial conditions.

• Robustness: Investigate adversarial testing for the stability of contextual explanations under data drift or manipulation (Sharma et al., 2025).

• Use LLMs in dynamic context generation: Use LLM-based agents to generate, summarise and contextualise alerts on the fly in a transparent manner (Srinivas et al., 2025).

CONCLUSION

In this work we aim to investigate the extent to which contextual explainability can alleviate this historical tension between automation and human trust in AI-driven SOCs. The findings provide definitive evidence that integrating contextual information – threat intelligence correlations, asset criticality and behavioural baselines – directly into explanations for AI decisions enables analysts to be more efficient, accurate and confident in automated systems. In comparison with feature-level explanations the contextual explainability facilitated faster triage, more accurate incident classifications and higher perceived trust. Such results support the contention that good xAI should not only focus on technical interpretability, but also be operationally relevant (Capuano et al., 2023; Rastogi et al., 1099 Dhanuka, Saxena & Mairal, 2025).

Yet, attaining contextual understanding is challenging at scale. Preserving accurate and timely contextual knowledge (e.g., evolving threat landscapes, or adjusting asset hierarchies) automatically integrated with cyber threat intelligence (CTI) and knowledge-graph frameworks that cannot be updated as needed is crucial (Santos et al., 2025; Bolton et al., 2023). Furthermore, contextual logic of the system needs to be transparent and tamperproof, so that using invoking explainability does not open a new attack surface [Sharma et al., 2025]. This raises the need for evolving explaintability models, that are more resilient to dynamic attacks and defenders.

In the future, real-world evaluation of contextual explainability in SOC for its long-term performance/robustness and trust retention/adversarial robustness also remain a challenge. The

recent coupling of LLM's and agent-based reasoning, for instance, would add even more layers of context that result from automatically produce situational narratives and adaptive explanations (Srinivas et al., 2025). More generally, we encourage future frameworks to use standardised evaluation protocols that incorporate (human-centred) metrics such as trust, workload and decision quality with algorithmic fidelity and robustness (Nauta et al., 2022).

We conclude that contextual explainability is disruptive towards human-aligned, transparent and accountable AI security solutions. Through injecting operational context directly into AI reasoning, SOCs will then become the action over black-box automation to systems that talk, reason and work with according to their human colleagues.

REFERENCES

Bolton, J., et al. (2023). An overview of cybersecurity knowledge graphs mapped to the MITRE ATT&CK framework domains. Preprint.

Capuano, N., Fenza, G., Loia, V., & Stanzione, C. (2023). Explainable Artificial Intelligence in Cybersecurity: A Survey. IEEE Transactions on Network and Service Management, 20(4), 5115–5140.

Johs, A. J., Agosto, D. E., & Weber, R. O. (2022). Explainable artificial intelligence and social science: Further insights for qualitative investigation. Applied AI Letters.

McDermid, J. A. (2021). Artificial intelligence explainability: The technical and ethical challenges. Philosophical Transactions of the Royal Society A, 379(2194), 20200363.

Mohsin, A., et al. (2025). A unified framework for human–AI collaboration in SOCs. arXiv preprint.

Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., & Seifert, C. (2022). From anecdotal evidence to quantitative evaluation: A systematic review on evaluating explainable AI. arXiv preprint.

Rastogi, N., Dhanuka, D., Saxena, A., & Mairal, P. (2025). The role of explainable AI in threat intelligence. arXiv preprint.

Rjoub, G., Bentahar, J., Wahab, O. A., Mizouni, R., Song, A., Cohen, R., et al. (2023). A survey on explainable artificial intelligence for cybersecurity. IEEE Transactions on Network and Service Management, 20(4), 5115–5140.

Santos, P., et al. (2025). A systematic review of cyber threat intelligence. Sensors, 25(14), 4272.

Sharma, A., et al. (2025). A comprehensive review of explainable AI in cybersecurity. ICT Express. (In press).

Srinivas, S., et al. (2025). AI-augmented SOC: A survey of LLMs and agents for security operations. Digital, 5(4), 95.

Tilbury, J., et al. (2024). Humans and automation: Augmenting Security Operations Centres for improved cyber defence. Digital, 4(3), 20.

Tariq, S., et al. (2025). Alert fatigue in Security Operations Centres: Research challenges and opportunities. Communications of the ACM.

Asma-Ul-Husna, A. R., & Paul, G. MKR Fatigue Estimation through Face Monitoring and Eye Blinking. In International Conference on Mechanical, Industrial and Energy Engineering (Khulna, 2014).

Bhuiya, R. A., Hasan, M. H., Barua, M., Rafsan, M., Jany, A. U. H., Iqbal, S. M. Z., & Hossan, F. (2025). Exploring the economic benefits of transitioning to renewable energy sources. International Journal of Materials Science, 6(2), 01-10.

Rokunuzzaman, M., Hasan, M., & Kader, M. A. (2012). Semantic Stability: A Missing Link between Cognition and Behavior. International Journal of Advanced Research in Computer Science, 3(4).

Rahman, M. M., Bandhan, L. R., Monir, L., & Das, B. K. (2025). Energy, exergy, sustainability, and economic analysis of a waste heat recovery for a heavy fuel oil-based power plant using Kalina cycle integrated with Rankine cycle. Next Research, 100398.

Neelapu, M. (2025). Predictive Software Defect Identification with Adaptive Moment Estimation based Multilayer Convolutional Network Model. Journal of Technological Innovations, 6(1).

Neelapu, M. (2025). Predictive Software Defect Identification with Adaptive Moment Estimation based Multilayer Convolutional Network Model. Journal of Technological Innovations, 6(1).

Neelapu, M. (2025). Predictive Software Defect Identification with Adaptive Moment Estimation based Multilayer Convolutional Network Model. Journal of Technological Innovations, 6(1).

Zahid, Z., Siddiqui, M. K. A., Alamm, M. S., Saiduzzaman, M., Morshed, M. M., Ferdousi, R., & Nipa, N. N. (2025, March). Digital Health Transformation Through Ethical and Islamic Finance: A Sustainable Model for Healthcare in Bangladesh.

Alamm, M. S., Zahid, Z., Nipa, N. N., & Khalil, I. (2025). Harnessing FinTech and Islamic Finance for Climate Resilience: A Sustainable Future Through Islamic Social Finance and Microfinance. Humanities and Social Sciences, 13(3), 207-218.

Zahid, Z., Amin, M. R., Alamm, M. S., Nipa, N. N., Khalil, I., Haque, A., & Mahmud, H. Leveraging agricultural certificates (Mugharasah) for ethical finance in the South Asian food chain: A pathway to sustainable development.

Zahid, Z., Amin, M. R., Monsur, M. H., Alamm, M. S., Nahid, I. K., Banna, H., ... & Nipa, N. N. Integrating FinTech Solutions in Agribusiness: A Pathway to a Sustainable Economy in Bangladesh.

Zahiduzzaman Zahid, M. S. A., Yousuf, M. A., Alam, M. M. A., Islam, M. A., Uddin, M. M., Parves, M. M., & Arif, S. (2025). Global Journal of Economic and Finance Research.

Zahid, Z., Amin, M. R., Alamm, M. S., Meer, W., Shah, M. N., Khalil, I., ... & Arafat, E. (2025). International Journal of Multidisciplinary and Innovative Research.

Zahid, Z., Amin, R., Khalil, I., Mohammed, B. A. K., & Arif, S. (2025). Regulating Digital Currencies in the EU: A Comparative Analysis with Islamic Finance Principles Under MiCA. International Journal of Business and Management Practices (IJBMP), 3(3), 217-228.

Zahid, Z., & Nipa, N. N. (2024). Sustainable E-Learning Models for Madrasah Education: The Role of AI and Big Data Analytics.

Ferdous, J., Islam, M. F., & Das, R. C. (2022). Dynamics of citizens' satisfaction on e-service delivery in local government institutions (Union Parishad) in Bangladesh. Journal of Community Positive Practices, (2), 107-119.

Ud Doullah, S., & Uddin, N. (2020). Public trust building through electronic governance: An analysis on electronic services in Bangladesh. Technium Soc. Sci. J., 7, 28.

Ferdous, J., Foyjul-Islam, M., & Muhury, M. (2024). Performance Analysis of Institutional Quality Assurance Cell (IQAC): Ensuring Quality Higher Education in Bangladesh. Rates of Subscription, 57.

Islam, M. F. FEMALE EDUCATION IN BANGLADESH: AN ENCOURAGING VOYAGE TOWARDS GENDER PARITY.

Ferdous, J., Zeya, F., Islam, M. F., & Uddin, M. A. (2021). Socio-economic vulnerability due to COVID-19 on rural poor: A case of Bangladesh. evsjv‡k cjøx Dbœqb mgxÿv.

Ferdous, J., & Foyjul-Islam, M. Higher Education in Bangladesh: Quality Issues and Practices.

Mollah, M. A. H. (2017). Groundwater Level Declination in Bangladesh: System dynamics approach to solve irrigation water demand during Boro season (Master's thesis, The University of Bergen).

Fuad, N., Meandad, J., Haque, A., Sultana, R., Anwar, S. B., & Sultana, S. (2024). Landslide vulnerability analysis using frequency ratio (FR) model: a study on Bandarban district, Bangladesh. arXiv preprint arXiv:2407.20239.

Mollah, A. H. (2023). REDUCING LOSS & DAMAGE OF RIVERBANK EROSION BY ANTICIPATORY ACTION. No its a very new study output.

Mollah, A. H. (2011). Resistance and Resilience of Bacterial Communities in Response to Multiple Disturbances Due to Climate Change. Available at SSRN 3589019.

Haque, A., Akter, M., Rahman, M. D., Shahrujjaman, S. M., Salehin, M., Mollah, A. H., & Rahman, M. M. Resilience Computation in the Complex System. Munsur, Resilience Computation in the Complex System.

Al Imran, S. M., Islam, M. S., Kabir, N., Uddin, I., Ali, K., & Halimuzzaman, M. (2024). Consumer behavior and sustainable marketing practices in the ready-made garments industry. International Journal of Management Studies and Social Science Research, 6(6), 152-161.

Islam, M. A., Goldar, S. C., Al Imran, S. M., Halimuzzaman, M., & Hasan, S. (2025). AI-Driven green marketing strategies for eco-friendly tourism businesses. International Journal of Tourism and Hotel Management, 7(1), 31-42.

Al Imran, S. M. (2024). Customer expectations in Islamic banking: A Bangladesh perspective. Research Journal in Business and Economics, 2(1), 12-24.

Islam, M. S., Amin, M. A., Hossain, M. B., Sm, A. I., Jahan, N., Asad, F. B., & Mamun, A. A. (2024). The Role of Fiscal Policy in Economic Growth: A Comparative Analysis of Developed and Developing Countries. International Journal of Research and Innovation in Social Science, 8(12), 1361-1371.

Al Amin, M., Islam, M. S., Al Imran, S. M., Jahan, N., Hossain, M. B., Asad, F. B., & Al Mamun, M. A. (2024). Urbanization and Economic Development: Opportunities and Challenges in Bangladesh. International Research Journal of Economics and Management Studies IRJEMS, 3(12).

SM, A. I., MD, A. A., HOSSAIN, M., ISLAM, M., JAHAN, N., MD, E. A., & HOSSAIN, M. (2025). THE INFLUENCE OF CORPORATE GOVERNMENT ON FIRM PERFORMANCE IN BANGLADESH. INTERNATIONAL JOURNAL OF BUSINESS MANAGEMENT, 8(01), 49-65.

Akter, S., Ali, M. R., Hafiz, M. M. U., & Al Imran, S. M. (2024). Transformational Leadership For Inclusive Business And Their Social Impact On Bottom Of The Pyramid (Bop) Populations. Journal Of Creative Writing (ISSN-2410-6259), 8(3), 107-125.

Ali, M. R. GREEN BRANDING OF RMG INDUSTRY IN SHAPING THE SUSTAINABLE MARKETING.

Hossain, M. A., Tiwari, A., Saha, S., Ghimire, A., Imran, M. A. U., & Khatoon, R. (2024). Applying the Technology Acceptance Model (TAM) in Information Technology System to

Evaluate the Adoption of Decision Support System. Journal of Computer and Communications, 12(8), 242-256.

Saha, S., Ghimire, A., Manik, M. M. T. G., Tiwari, A., & Imran, M. A. U. (2024). Exploring Benefits, Overcoming Challenges, and Shaping Future Trends of Artificial Intelligence Application in Agricultural Industry. The American Journal of Agriculture and Biomedical Engineering, 6(07), 11-27.

Ghimire, A., Imran, M. A. U., Biswas, B., Tiwari, A., & Saha, S. (2024). Behavioral Intention to Adopt Artificial Intelligence in Educational Institutions: A Hybrid Modeling Approach. Journal of Computer Science and Technology Studies, 6(3), 56-64.

Noor, S. K., Imran, M. A. U., Aziz, M. B., Biswas, B., Saha, S., & Hasan, R. (2024, December). Using data-driven marketing to improve customer retention for US businesses. In 2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA) (pp. 338-343). IEEE.

Tiwari, A., Saha, S., Johora, F. T., Imran, M. A. U., Al Mahmud, M. A., & Aziz, M. B. (2024, September). Robotics in Animal Behavior Studies: Technological Innovations and Business Applications. In 2024 IEEE International Conference on Computing, Applications and Systems (COMPAS) (pp. 1-6). IEEE.

Sobuz, M. H. R., Saleh, M. A., Samiun, M., Hossain, M., Debnath, A., Hassan, M., ... & Khan, M. M. H. (2025). AI-driven modeling for the optimization of concrete strength for Low-Cost business production in the USA construction industry. Engineering, technology & applied science research, 15(1), 20529-20537.

Imran, M. A. U., Aziz, M. B., Tiwari, A., Saha, S., & Ghimire, A. (2024). Exploring the Latest Trends in AI Technologies: A Study on Current State, Application and Individual Impacts. Journal of Computer and Communications, 12(8), 21-36.

Tiwari, A., Biswas, B., ISLAM, M., SARKAR, M., Saha, S., Alam, M. Z., & Farabi, S. F. (2025). Implementing robust cyber security strategies to protect small businesses from potential threats in the USA. JOURNAL OF ECOHUMANISM Учредители: Transnational Press London, 4(3).

Hasan, R., Khatoon, R., Akter, J., Mohammad, N., Kamruzzaman, M., Shahana, A., & Saha, S. (2025). AI-Driven greenhouse gas monitoring: enhancing accuracy, efficiency, and real-time emissions tracking. AIMS Environmental Science, 12(3), 495-525.

Hossain, M. A., Ferdousmou, J., Khatoon, R., Saha, S., Hassan, M., Akter, J., & Debnath, A. (2025). Smart Farming Revolution: AI-Powered Solutions for Sustainable Growth and Profit. Journal of Management World, 2025(2), 10-17.

Saha, S. (2024). Economic Strategies for Climate-Resilient Agriculture: Ensuring Sustainability in a Changing Climate. Demographic Research and Social Development Reviews, 1(1), 1-6.

Saha, S. (2024). -27 TAJABE USA (150$) EXPLORING+ BENEFITS,+ OVERCOMING. The American Journal of Agriculture and Biomedical Engineering.

Adeojo, O. S., Egerson, D., Mewiya, G., & Edet, R. (2021). The ideology of baby-mama phenomenon: Assessing knowledge and perceptions among young people from educational institutions.

Orugboh, O. G. (2025). AGENT-BASED MODELING OF FERTILITY RATE DECLINE: SIMULATING THE INTERACTION OF EDUCATION, ECONOMIC PRESSURES, AND SOCIAL MEDIA INFLUENCE. NextGen Research, 1(04), 1-21.

Orugboh, O. G., Ezeogu, A., & Juba, O. O. (2025). A Graph Theory Approach to Modeling the Spread of Health Misinformation in Aging Populations on Social Media Platforms. Multidisciplinary Journal of Healthcare (MJH), 2(1), 145-173.

Orugboh, O. G., Omabuwa, O. G., & Taiwo, O. S. (2025). Predicting Intra-Urban Migration and Slum Formation in Developing Megacities Using Machine Learning and Satellite Imagery. Journal of Social Sciences and Community Support, 2(1), 69-90.

Orugboh, O. G., Omabuwa, O. G., & Taiwo, O. S. (2024). Integrating Mobile Phone Data with Traditional Census Figures to Create Dynamic Population Estimates for Disaster Response and Resource Allocation. Research Corridor Journal of Engineering Science, 1(2), 210-228.

Orugboh, O. G., Omabuwa, O. G., & Taiwo, O. S. (2024). Predicting Neighborhood Gentrification and Resident Displacement Using Machine Learning on Real Estate, Business, and Social Datasets. Journal of Social Sciences and Community Support, 1(2), 53-70.

Daniel, E., Opeyemi, A., Ruth, O. E., & Gabriel, O. (2020). Understanding Childbearing for Households in Emerging Slum Communities in Lagos State, Nigeria. International Journal of Research and Innovation in Social Science, 4(9), 554-560.