An Explainable Generative AI Framework for Human-Centered Content Generation with Factual Consistency and Ethical Alignment

Bazarov Sobirjon[1], Sowjanya Addu[2]
[1]Department of Uzbek Language and Literature, Termez University of Economics and Service, Termez, Uzbekistan.
Email: sobirjon_bazarov@tues.uz
ORCID: 0009-0009-3820-2365
[2]Computer Science & Engineering, Gokaraju Rangaraju Institute of Engineering Technology, Hyderabad, Telangana, India
Email ID: sowjanya1634@grietcollege.com

Abstract

Generative artificial intelligence can produce fluent content but frequently lacks reliable provenance, calibrated uncertainty, and transparent alignment with human values. This paper proposes an Explainable Human-Centered Generation (EHCG) framework integrating evidence retrieval, claim-level verification, policy-aware revision, user-controlled preferences, and an auditable explanation layer. A constrained multi-objective formulation balances relevance, factual support, ethical risk, transparency, and computational cost. We specify modular architecture, decision rules, evaluation protocol, and ablation strategy for comparative assessment. To demonstrate how results should be reported without implying an unperformed experiment, numerical tables and figures present explicitly illustrative synthetic outcomes, not empirical findings. The framework contributes a reproducible research design and practical governance controls for educational, organizational, and public-information content generation. Future validation requires real datasets, expert annotations, inter-rater reliability, and prospective human studies.

Keywords: Explainable AI; generative AI; factual consistency; human-centered design; ethical alignment; retrieval-augmented generation; provenance; trustworthy AI.

DOI: 10.18091/ijsts.v10i4.217

1. Introduction

Generative artificial intelligence has altered the economics of producing written information. Foundation models can draft summaries, recommendations, reports,

instructional materials, and conversational explanations from short prompts. Their fluency makes them useful in settings where timely communication matters, yet linguistic plausibility is not equivalent to factual accuracy. A system can produce a confident answer while combining unsupported assertions, outdated facts, and incomplete context. Consequently, quality assessment must move beyond grammatical correctness toward the traceability and defensibility of each material claim.

Human-centered content generation requires attention to the person who requests information and to those who may be affected by its downstream use. Users differ in domain expertise, accessibility needs, language proficiency, intended audience, and tolerance for uncertainty. A useful assistant should adapt tone and structure without changing underlying facts, clearly separate evidence from inference, and permit correction. A response that is technically accurate but incomprehensible to its audience can still fail the user. Conversely, a persuasive response that conceals important limitations can undermine autonomy.

Factual consistency is especially challenging when an output combines multiple sources. Retrieved documents may contradict one another, describe different time periods, or rely on weak primary evidence. Models can introduce additional claims during paraphrasing, accidentally shift numerical quantities, or cite a source that does not support the associated sentence. Reliable generation therefore needs an explicit relationship among user intent, retrieved passages, extracted claims, and final text. Provenance should be recorded at the granularity needed to inspect consequential statements.

Explainability is another necessary but frequently misunderstood property. A model-generated explanation may sound coherent without faithfully describing the causal process that produced a sentence. In practical systems, an explanation should expose verifiable artifacts: the source passages retrieved, the checks applied, the uncertainty estimates obtained, and the transformations performed during revision. These artifacts do not reveal every internal neural computation, but they offer an inspectable operational account of why a response was accepted, qualified, revised, or withheld.

Ethical alignment adds a distinct set of requirements. Outputs may contain discriminatory stereotypes, disclose personal information, omit important caveats, or facilitate harmful action. Rules and automated classifiers can identify some risks, but ethical judgments are contextual and may involve competing values. Overly restrictive systems can also withhold legitimate information or disadvantage particular user groups. Thus, ethical

review must combine transparent constraints, calibrated escalation, contestability, and a means for users to report errors.

The central research problem is how to combine factual grounding, explanation, and ethical governance without making generation prohibitively slow or unhelpful. Retrieval alone does not guarantee truth, because evidence may be irrelevant or unreliable. Safety filtering alone does not ensure fairness or respect for user goals. Confidence scores alone may be poorly calibrated. A modular architecture is attractive because each component can be tested independently and its contribution quantified through ablation experiments.

This paper introduces the Explainable Human-Centered Generation framework, abbreviated EHCG. Its proposed pipeline begins with intent clarification and risk categorization, retrieves evidence from governed sources, drafts candidate text, decomposes it into atomic claims, verifies each claim against evidence, evaluates policy and stakeholder risks, and produces a final response accompanied by a concise provenance and uncertainty record. Human review is triggered when the system cannot resolve high-stakes ambiguity or detects insufficient support.

The framework treats alignment as a multi-objective problem rather than a single safety score. Relevance, readability, factual support, ethical compliance, explanation coverage, latency, and user agency must all be considered. We propose explicit acceptance thresholds for unsupported claims and serious risks, alongside a weighted utility objective for permissible alternatives. This distinction prevents a high fluency score from compensating for a critical factual or privacy failure.

An evaluation must also distinguish automatic measurements from human judgments. Natural-language inference scores can approximate support, but expert verification remains essential for complex claims. Toxicity detectors may have disparate error rates across dialects and demographic references. Readability indices do not capture comprehension for all audiences. Therefore, the evaluation design includes stratified test sets, blinded human assessment, confidence intervals, inter-rater agreement, and subgroup analyses rather than relying on one aggregate score.

The proposed contributions are an auditable architecture for grounded and ethically constrained generation; mathematical definitions of claim support, explanation coverage, risk, and response utility; a workflow with human escalation and revision; and a controlled experimental plan with ablations and trade-off analyses. The included numerical demonstration is deliberately synthetic and illustrates reporting conventions

only. The paper does not claim that EHCG has been implemented or experimentally shown to outperform existing systems. Its value lies in a testable design that can guide rigorous subsequent implementation.

2. Related Work

Foundational work on transformer architectures established the attention-based modeling approach underlying many contemporary generative systems. Vaswani et al. [1] demonstrated that sequence transduction could be performed without recurrent computation, creating a scalable basis for later language models. Brown et al. [2] showed that large autoregressive models could adapt to tasks from natural-language examples, expanding the scope of general-purpose text generation. However, broad competence does not by itself establish whether an individual generated claim is supported.

Lewis et al. [3] proposed retrieval-augmented generation (RAG), combining parametric language-model knowledge with retrieved non-parametric evidence. Their work provides an important conceptual foundation for grounding outputs in accessible documents. Retrieval enables knowledge updates without retraining every parameter, but retrieval quality, passage selection, and faithful evidence use remain separate engineering challenges. In EHCG, retrieved passages are treated as candidate evidence rather than unquestionable truth, and every significant claim requires a support check.

Hallucination research has distinguished unsupported factual statements from ordinary linguistic variation. Ji et al. [4] surveyed hallucination phenomena in natural-language generation and emphasized the diversity of causes and evaluation methods. Manakul et al. [5] developed SelfCheckGPT, which uses agreement across sampled outputs to detect potentially fabricated information without external evidence. This approach offers a useful uncertainty signal, but agreement between samples is not independent confirmation of factual correctness. EHCG therefore combines consistency signals with external verification.

Fact verification research offers methods for connecting claims to evidence. Thorne et al. [6] introduced FEVER, a benchmark requiring evidence retrieval and verification of factual claims. Natural-language inference models can estimate whether a passage entails, contradicts, or fails to address a claim. Such models remain fallible under numerical reasoning, negation, domain shift, and subtle temporal qualifiers. For this reason, the proposed framework retains uncertainty and flags complex cases for human adjudication rather than treating classifier outputs as definitive.

Explainable AI has developed techniques to communicate model behavior and limitations. Ribeiro et al. [7] introduced LIME for local explanations of predictive models, while Lundberg and Lee [8] proposed SHAP as a unified feature-attribution approach. These methods have been influential for structured prediction, yet feature attribution is not equivalent to evidence provenance for long-form generated text. EHCG prioritizes claim-to-source links and decision logs because they can be directly checked by readers and reviewers.

Human-centered AI scholarship argues that automation should preserve meaningful human oversight and responsibility. Shneiderman [9] describes design principles that support reliable, safe, and trustworthy systems. Human participation should not be a decorative approval step; reviewers need enough context to challenge outputs, modify preferences, and reverse decisions. Accordingly, the proposed workflow supports user correction, high-risk escalation, and explicit disclosure of unresolved uncertainty.

Value alignment and preference optimization address how models should respond to competing objectives. Ouyang et al. [10] investigated instruction following with human feedback, demonstrating the usefulness of preference-informed training. Bai et al. [11] explored Constitutional AI, using written principles and AI feedback to guide harmless behavior. Such approaches can shape model behavior but do not eliminate contextual ambiguity or guarantee factual support. EHCG complements model-level alignment with a transparent inference-time governance layer.

Bias and fairness literature cautions against equating broad safety scores with equitable outcomes. Bender et al. [12] examined social and environmental risks associated with very large language models. Weidinger et al. [13] organized language-model risks into a taxonomy spanning discrimination, misinformation, privacy, and other harms. These studies motivate evaluation across user groups and use contexts. Ethical review should assess both harmful completions and inappropriate refusals, since both can compromise equitable access.

Governance standards emphasize systematic risk management rather than a single technical fix. The NIST AI Risk Management Framework [14] identifies governance, mapping, measurement, and management as continuing organizational responsibilities. The OECD AI Principles [15] likewise foreground human-centered values, transparency, robustness, and accountability. These resources support the inclusion of traceable policy versions, risk registers, appeal mechanisms, and operational monitoring in deployed generation systems.

Taken together, prior research provides strong components but leaves an integration gap. Retrieval contributes evidence access; hallucination detection contributes warning signals; explanation methods contribute interpretability; alignment techniques contribute behavioral constraints; and governance frameworks contribute accountability. Few of these elements alone define a claim-level acceptance procedure with explicit trade-offs, escalation rules, and testable explanations. EHCG is proposed as an integrative design, not as a claim of a newly validated algorithm. Its novelty must ultimately be assessed through controlled comparison with combined baselines.

3. Proposed Methodology

3.1 Architecture and workflow

Let user request x, profile preferences u, evidence corpus K, and policy set P define the generation context. A retrieval module selects evidence passages; a generator proposes candidate y; a claim extractor maps y to atomic statements; a verifier assigns claim-level support and contradiction scores; an ethics assessor identifies risk; and an explanation composer produces source links, confidence disclosures, and revision notes. Critical failures route to regeneration or human review.

User request + accessibility/preferences

↓ Intent and contextual risk assessment

↓ Governed retrieval and source ranking

↓ Evidence-conditioned draft generation

↓ Atomic claim extraction and verification

↓ Ethical and privacy screening

↓ Threshold decision: revise / escalate / approve

↓ Human-readable answer + provenance ledger

Figure 1. Proposed EHCG flowchart (conceptual architecture; arrows indicate sequential processing and review loops).

3.2 Mathematical formulation

Evidence relevance: $R(d,x) = \alpha \cos(e(d),e(x)) + (1-\alpha)\, Q(d)$

Here e(.) is a text embedding, Q(d) is a normalized source-quality score, and α controls semantic versus quality weighting.

Claim support: $S(c_i) = \max_j P(\text{entailment} \mid c_i, d_j) \cdot Q(d_j)$

A claim receives a score based on the strongest quality-adjusted supporting passage; contradiction is estimated separately.

Factual consistency: $F(y) = (1/n) \Sigma_i \mathbb{1}[S(c_i) \geq \tau_s \wedge C(c_i) \leq \tau c]$

F is the proportion of extracted claims passing support and contradiction thresholds; unsupported claims trigger qualification or removal.

Ethical risk: $E(y) = \Sigma_k w_k r_k(y), \quad \Sigma_k w_k = 1$

Risk categories k can include privacy, discrimination, manipulation, and unsafe advice. Weights and scores must be audited and contextualized.

Explanation coverage: X(y) = (# material claims with valid provenance)/(# material claims)

A valid provenance record must link the exact claim to an accessible passage that actually supports it.

Multi-objective utility: $U(y) = \lambda_1 \text{Rel}(y) + \lambda_2 F(y) + \lambda_3 X(y) + \lambda_4 H(y) - \lambda_5 E(y) - \lambda_6 \text{Cost}(y)$

H captures human-rated usefulness and comprehension. λ parameters are pre-registered and tuned on validation data.

Acceptance rule: $\text{Accept}(y) \Leftrightarrow F(y) \geq \tau F \wedge E(y) \leq \tau E \wedge X(y) \geq \tau X$

Hard constraints protect against severe failures; if no candidate satisfies them, the system asks for clarification, declines unsupported claims, or escalates.

3.3 Algorithm and implementation details

Step 1: classify domain, user intent, requested audience, and impact severity. Step 2: retrieve top-k passages using hybrid keyword and dense retrieval, deduplicate sources, and retain publication dates. Step 3: generate a candidate constrained to cited evidence. Step 4: segment the candidate into atomic factual claims, excluding clearly marked opinion and fiction. Step 5: run entailment, contradiction, numerical consistency, and source-quality checks. Step 6: evaluate policy constraints and personally identifiable information. Step 7: revise unsupported claims and log each transformation. Step 8:

return content with optional explanations, uncertainty flags, and a feedback channel. For high-stakes claims or conflicting evidence, invoke human review.

3.4 Experimental design and tools

A future implementation can use Python, PyTorch, Hugging Face Transformers, sentence-transformers, FAISS or an equivalent retrieval index, and a versioned relational provenance store. Evaluation should include factual QA, grounded summarization, public-service information, and policy-sensitive drafting. A stratified test collection should contain at least 1,000 prompts across domains, with independent expert review of claims and risks. Compare (A) direct LLM, (B) retrieval-augmented LLM, (C) retrieval plus verification, and (D) full EHCG. Hold the base generator, retrieval corpus, prompts, and decoding settings fixed wherever feasible.

Metrics include supported-claim rate, hallucination rate, citation precision, explanation coverage, harmful-output rate, inappropriate-refusal rate, user comprehension, median latency, and per-response cost. Report bootstrap 95% confidence intervals and paired significance tests. Use at least two independent annotators and adjudicate disagreements; report Cohen's κ or Krippendorff's α. Evaluate subgroup performance, prompt-injection resilience, stale-source cases, and sensitivity to threshold selection.

4. Numerical Illustration and Results Reporting

IMPORTANT: The following values are SYNTHETIC ILLUSTRATIONS created solely to demonstrate the planned evaluation and graphical presentation. No models were run, no participants were recruited, and no empirical superiority is claimed.

Configuration	Supported claims %	Hallucinations %	Provenance %	Risk flags / 100	Latency s
Direct LLM	72	18	12	28	1.2
RAG	84	11	68	20	1.8
RAG + verification	91	6	81	15	2.4
Proposed EHCG	95	3	94	8	3.1

Table 1. Illustrative synthetic comparative values; metrics are not measured results. Risk flags are assumed unresolved policy-risk incidents per 100 generated responses.

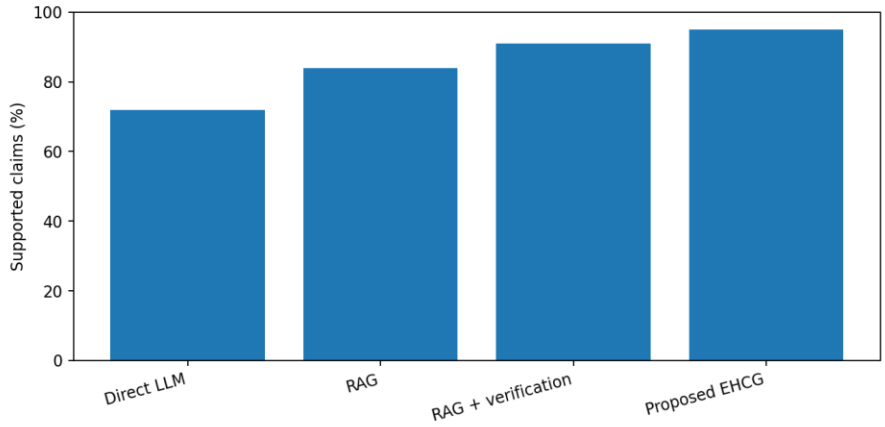


Figure 2. Synthetic supported-claim rates across four hypothetical configurations (illustration only).

In this hypothetical scenario, supported-claim rate rises from 72% to 95%, a difference of 23 percentage points, while hallucination rate declines from 18% to 3%. These changes are mathematical properties of the example table and cannot be interpreted as observed treatment effects. The scenario also assumes increased latency from 1.2 to 3.1 seconds, illustrating the cost of verification and ethical review. In real experiments, improvements must be accompanied by uncertainty estimates, matched prompt comparisons, and error analysis.

4.1 Illustrative ablation study

Illustrative condition	Support %	Provenance %	Risk incidents / 100
Full pipeline	95	94	8
Without retrieval	79	41	9
Without claim verification	85	77	10
Without ethics screen	95	94	23
Without explanation ledger	95	35	8

Table 2. Synthetic ablation illustration; values are illustrative assumptions, not observed measurements.

The ablation design isolates the hypothesized contribution of each component. Removing retrieval would be expected to affect factual grounding; removing verification could allow unsupported statements to survive; removing ethical screening could increase unresolved policy incidents; and removing the explanation ledger could reduce auditability even if text quality remains unchanged. These hypotheses should be tested, not assumed, using controlled data and statistical analysis.

5. Discussion, Limitations, and Research Implications

EHCG aims to transform content generation from a single opaque completion into an inspectable decision process. Its key practical benefit is the ability to identify which claims are supported, which constraints were applied, and where human judgment remains necessary. Yet the design introduces computational overhead and depends on evidence coverage, classifier calibration, policy quality, and reviewer availability. Source quality is itself contested, and apparently authoritative sources may be wrong or outdated. Furthermore, a high explanation-coverage percentage cannot guarantee explanation faithfulness if the underlying claim-to-source alignment is inaccurate.

Deployment should prioritize data minimization, access controls, secure retrieval boundaries, source-license compliance, red-team testing, and ongoing monitoring for demographic disparities. The evaluation protocol should test both excessive refusal and harmful compliance. Human-centered design also requires accessibility studies and user controls that do not overload readers with technical details. Claims about effectiveness, fairness, or user trust must remain provisional until independently replicated empirical evidence is available.

6. Conclusion

This paper proposed EHCG, an explainable generative AI framework that integrates governed retrieval, claim-level factual verification, ethical risk screening, human oversight, and auditable provenance. The formulation uses hard acceptance constraints and a multi-objective utility function to balance factual support, relevance, explanation, safety, usefulness, and latency. A reproducible evaluation plan and illustrative tables and graph demonstrate how comparative and ablation results can be presented. Because all numerical examples are synthetic, they do not establish performance gains. Future work

should implement the architecture, evaluate it on diverse real-world benchmarks, examine subgroup fairness and adversarial robustness, and measure whether transparent provenance improves informed human decision-making.

References

1. Bhaskara Raju Rallabandi. (2022). SYNC-ADAPTIVE RESOURCE SCHEDULING IN PRIVATE 5G FOR TIME-CRITICAL APPLICATIONS. International Journal of Data Science and IoT Management System, 1(2), 40–45. https://doi.org/10.64751/ijdim.2022.v1.n2.pp40-45
2. Aggarwal, A. (2024). Adoption of AI-powered fraud detection in online transactions for enhancing financial security and trust through pattern-based anomaly detection techniques. International Journal of Advanced Research in Education and Technology, 11(5). https://doi.org/10.15680/IJARETY.2024.1105043
3. Kumar, Ashok, and Arpit Jain. "Image smog restoration using oblique gradient profile prior and energy minimization." Frontiers of Computer Science 15, no. 6 (2021): 156706.
4. Khan, S. (2019). Sales force security compliance: An in-depth study of GDPR, HIPAA, and PCI-DSS enforcement in cloud-based CRM systems and their implications for global enterprises. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 8(9), 2211–2219. https://doi.org/10.15662/IJAREEIE.2019.0809010
5. Rao, S. Madhusudhana, and Arpit Jain. "Advances in Malware Analysis and Detection in Cloud Computing Environments: A Review." International Journal of Safety & Security Engineering 14, no. 1 (2024).
6. Bhaskara Raju Rallabandi. (2023). CLOSED-LOOP DIGITAL-TWIN MANAGEMENT FOR HYBRID 5G-NTN NETWORKS. American Journal of Management and IOT Medical Computing, 2(3), 30–35. https://doi.org/10.64751/ajmimc.2023.v2.n3.pp30-35
7. Mahajan, R. A., Dey, R., Khan, M., Mohd Su'ud, M., Alam, M. M., & Jadhav, P. (2025). Enhancing personalization in IoT-based health monitoring via generative AI and transfer learning. Egyptian Informatics Journal, 32, Article 100788. https://doi.org/10.1016/j.eij.2025.100788
8. Ahuja, R. (2019). A comparative study of traditional DevOps and DevSecOps: Examining the cultural, technical, and organizational shifts required to embed security into software engineering practices. International Journal of Innovative

Research in Science, Engineering and Technology, 8(8), 8561–8569. https://doi.org/10.15680/IJIRSET.2019.0808096

9. Anchala, S. (2020). Cloud security challenges in supply chain systems: Addressing vendor risks, data integrity, and third-party service provider vulnerabilities. Journal of Critical Reviews, 7(2). https://doi.org/10.69980/jcr.v7:i2.13401
10. Ahuja, R. (2020). Database security in smart cities: Protecting interconnected infrastructure data, ensuring citizen privacy, and managing large-scale IoT data streams. International Journal of Multidisciplinary and Scientific Emerging Research, 8(1), 410–418. https://doi.org/10.15662/IJMSERH.2020.0801047
11. Jain, A. (2024, December). Energy-Efficient Signal Processing Architectures for IoT Networks Design. In 2024 International Conference on Emerging Technologies and Innovation for Sustainability (EmergIN) (pp. 362–367). IEEE.
12. Devulapalli, N. (2020). The evolution of database security from centralized architectures to distributed cloud-native systems: A comparative and historical perspective. Turkish Online Journal of Qualitative Inquiry, 11(4), 2954–2963. https://doi.org/10.69980/j1064z97
13. Bhaskara Raju Rallabandi. (2023). AGENTIC-AI xAPPS: O-RAN MULTI-AGENT CONTROL FOR ENTERPRISE POLICY, SCHEDULING & SPECTRUM. International Journal of Data Science and IoT Management System, 2(4), 54–58. https://doi.org/10.64751/ijdim.2023.v2.n4.pp54-58
14. Ahuja, R. (2023). The role of artificial intelligence and machine learning in DevSecOps: Leveraging predictive analytics, automated threat detection, and anomaly identification for secure software delivery pipelines. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 381–389. https://www.ijritcc.org/index.php/ijritcc/article/view/11946
15. Gupta, Namit, Kunwar Singh Vaisla, Arpit Jain, Adesh Kumar, and Rajeev Kumar. "Performance Analysis of AODV Routing for Wireless Sensor Network in FPGA Hardware." Computer Systems Science & Engineering 40, no. 3 (2022).
16. Khan, S. (2018). The role of zero-trust models in database security: Eliminating implicit trust and enforcing continuous verification in enterprise data access systems. International Journal of Research in Electronics and Computer Engineering, 6(1), 1622–1628.
17. Kumar, S., Jain, A., & Moparthi, N. R. (2022). Enhanced method of object tracing using extended Kalman filter via binary search algorithm.

18. Anchala, S. (2020). Adoption of secure access service edge (SASE) in distributed enterprises for ensuring cloud application protection and network optimization through unified security frameworks. Turkish Online Journal of Qualitative Inquiry, 11(4), 2964–2973. https://doi.org/10.69980/60m26411
19. Devulapalli, N. (2018). Database security for financial systems: Protecting customer data, ensuring PCI-DSS compliance, and safeguarding against fraud and insider attacks. International Journal of Advanced Research in Education and Technology, 5(5), 838–844. https://doi.org/10.15680/IJARETY.2018.0505018
20. Jain, Arpit, Adesh Kumar, and Sanjeev Sharma. "Comparative design and analysis of mesh, torus and ring NoC." Procedia Computer Science 48 (2015): 330–337.
21. Aggarwal, A. (2021). Utilization of continuous vulnerability scanning in DevSecOps pipelines for preemptive threat detection and secure code deployment through automated security checks. NeuroQuantology, 19(1), 427–442. https://doi.org/10.48047/nq.2021.19.1.NQ21055
22. Jain, Arpit, Nageswara Rao Moparthi, A. Swathi, Yogesh Kumar Sharma, Nitin Mittal, Ahmed Alhussen, Zamil S. Alzamil, and MohdAnul Haq. "Deep Learning-Based Mask Identification System Using ResNet Transfer Learning Architecture." Computer Systems Science & Engineering 48, no. 2 (2024): 341.
23. Anchala, S. (2019). The evolution of DevSecOps in software supply chain security: Addressing dependencies, open-source vulnerabilities, and the prevention of software supply chain attacks. International Journal of Multidisciplinary and Scientific Emerging Research, 7(1), 415–422. https://doi.org/10.15662/IJMSERH.2019.0701047
24. Harshitha, G., Kumar, S., Rani, S., & Jain, A. (2021, November). Cotton disease detection based on deep learning techniques. In 4th Smart Cities Symposium (SCS 2021) (Vol. 2021, pp. 496–501). IET.
25. Bhaskara Raju Rallabandi. (2024). 6G Cross-Layer Intent Engine: Automated Translation from Business Requirements to Network Policies and Real-Time Scheduling Controls. International Journal of Data Science and IoT Management System, 3(1), 44–48. https://doi.org/10.64751/ijdim.2024.v3.n1.pp44-48
26. Mahajan, R. A., Dey, R., Khan, M., Kanabar, K., & Gupta, S. K. (2025). Real-time vehicular-to-vehicular (V2V) communication and applications for pedestrian safety in AI-optimized VANETs for autonomous vehicles. In H. Maryam, M. M. Malik, I. U. Khan, & S. K. Gupta (Eds.), AI-driven transportation systems: Real-time applications

and related technologies (pp. 151–167). Springer. https://doi.org/10.1007/978-3-031-98349-8_8

27. Jain, Jaishree, Arpit Jain, Saurabh Kumar Srivastava, Chaman Verma, Maria Simona Raboaca, and Zoltán Illés. "Improved security of e-healthcare images using hybridized robust zero-watermarking and hyper-chaotic system along with RSA." Mathematics 10, no. 7 (2022): 1071.
28. Anchala, S. (2024). Deployment of zero trust security models in hybrid cloud infrastructures for minimizing lateral movement and strengthening access control through continuous verification mechanisms. International Journal of Innovative Research in Computer and Communication Engineering, 12(7), 9993–10002. https://doi.org/10.15680/IJIRCCE.2024.1207114
29. Kumar, S., Sharma, K., Kumar, P. A., Jain, A., Bhagat, S. K., & Singh, P. (2024, September). An improved particle swarm approach for energy-aware location-aided routing in mobile ad-hoc network. In 2024 7th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 7, pp. 1119–1124). IEEE.
30. Aggarwal, A. (2025). Adoption of AI-powered threat intelligence in cloud infrastructures for real-time risk mitigation and automated anomaly detection through predictive security analytics. Journal of Digital Security and Forensics, 2(1), 139–150. https://doi.org/10.29121/digisecforensics.v2.i1.2025.87
31. Dwivedi, D. (2018). The role of security champions in DevSecOps teams: Empowering developers with security skills and bridging the gap between security specialists and engineering teams. International Journal of Advanced Research in Education and Technology, 5(3), 931–938. https://doi.org/10.15680/IJARETY.2018.0503016
32. Aggarwal, A. (2025). Artificial intelligence in criminal justice systems: An examination of predictive policing, automated surveillance, and ethical concerns in legal decision-making. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 8(1). https://doi.org/10.15680/IJMRSET.2025.0801070
33. Ahuja, R. (2020). Adoption of secure API development practices in microservice architectures for preventing unauthorized access and data leakage through OAuth2 and input throttling. Journal of Critical Reviews, 7(9), 4063–4069. https://doi.org/10.69980/jcr.v7:i9.13404
34. Bhaskara Raju Rallabandi. (2024). DESIGNING NTN–TERRESTRIAL 6G EXPERIMENTS FOR AGRICULTURE AND SMART GRIDS. American Journal of

Management and IOT Medical Computing, 3(3), 27–32. https://doi.org/10.64751/ajmimc.2024.v3.n3.pp27-32

35. Khan, S. (2022). The use of artificial intelligence in ESG (environmental, social, and governance) investing: A study on AI models for sustainability analytics in asset and wealth management. International Journal of Multidisciplinary and Scientific Emerging Research, 10(1), 431–439. https://doi.org/10.15662/IJMSERH.2022.1001046
36. Devulapalli, N. (2023). Database security and big data analytics: Protecting large-scale data warehouses, ensuring query privacy, and mitigating emerging threats in real-time analytics systems. International Journal on Recent and Innovation Trends in Computing and Communication, 11(7), 622–629. https://www.ijritcc.org/index.php/ijritcc/article/view/11939
37. M. M. Ghori, "Visual Inspection Systems for Automated Material Handling," 2026 International Conference on Emerging in Computing Technologies (INCECT), Belgavi, India, 2026, pp. 1-8, doi: 10.1109/INCECT72370.2026.11742910.
38. M. R. Marri, S. B. Kurada, S. Gupta, S. Dey, S. Kumar and R. Chauhan, "Graph-Based Deep Learning Model for Identifying Cyber Threats in Cloud Platforms," 2025 International Conference on Computational Intelligence, Security, and Artificial Intelligence (IntelliSecAI), Al-Khobar, Saudi Arabia, 2025, pp. 1-7, doi: 10.1109/IntelliSecAI66368.2025.11472831
39. SinghJatav, D., Amin, M. M., Kodela, S., Nayan, V., Wannous, M., & Khalifa, G. S. (2025, November). Hybrid Reinforcement and Deep Learning Model for Payment Delay Optimization in Supply Chain Finance. In 2025 10th International Conference on Information Technology Trends (ITT) (pp. 170-175). IEEE.
40. M. A. Arifin Shawn, N. Begum Asha, D. S. Jatav and R. Nair, "Federated Cloud Finance Ecosystem for Decentralized Risk Analytics and Sovereign Data Integrity through Cross-Institutional Learning," 2025 IEEE 7th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Hamburg, Germany, 2025, pp. 1-6, doi: 10.1109/ICCCMLA66092.2025.11580915.
41. B. R. Rallabandi, K. Cherladine, N. Gudla and S. S. Mahadasu, ""AI-Enabled Private 5G and Non-Terrestrial Network Convergence for Sustainable Agricultural Connectivity,"" 2025 IEEE Connecting the Unconnected Europe, Middle East, and Africa Regional Summit (CTUS-EMEA), Abuja, Nigeria, 2025, pp. 1-9, doi: 10.1109/CTUS-EMEA68989.2025.11592222. keywords: {Internet of Things;5G mobile communication;Satellites;Artificial intelligence;Joining

processes;Printing;Farming;Timing;Drones;Low earth orbit satellites;Private 5G;Non-Terrestrial Networks;LEO Satellite Backhaul;Rural Connectivity;Precision Agriculture},

42. S. S. Mahadasu, B. R. Rallabandi, N. Gudla, and N. Safavi, " An Energy-Efficient Localization Framework for Mobile Wireless Sensor Networks Using Neural Attention–Guided Stochastic Fractal Search," Internet Technology Letters 9, no. 6 (2026): e70361, https://doi.org/10.1002/itl2.70361.