OPEN ACCESS

Kashf Journal of Multidisciplinary Research
Vol: 03 - Issue 03 (2026)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

DECENTRALIZED PROVENANCE LAYER FOR FOUNDATION MODELS: A FRAMEWORK FOR QUANTIFYING AND PENALIZING SYNTHETIC DATA CONTAMINATION IN RECURSIVE LLM TRAINING

[1]SyedTalib Zaheer Zaidi, [2]Muhammad Zamin Ali Khan, [3]Muhammad Usama Khan,[4]Amad Asif, [5]Khalid Bin Muhammad,[6]Faigha Karim, [7]Ammad Mallick

[1] HBL, Karachi, Pakistan

[2] Department of Computer Science, Iqra University, Karachi, Pakistan

[3]UHF Solutions Pvt Ltd, Karachi, Pakistan

[4]Graphica Pro Artistry (Australia, Broadmeadows, Victoria)

[5] COCSE, Ziauddin University, Karachi, Pakistan

[6]Department of Computer Science, Iqra University, Karachi, Pakistan

[7]Department of Computer Science, Cardiff Metropolitan University, London, UK

[*]Corresponding Author: ([2]muhammad.zamin@iqra.edu.pk)

DOI:(*https://doi.org/10.71146/kjmr944*)

Article Info

Abstract

The rapid development of Large Language Models (LLMs) and other Foundation Models requires substantial amounts of high-quality, human-generated training data [1]. However, the global collection of original human-generated content, known as the 'internet corpus,' is experiencing a marked decline [1]. To train successive generations of models, increasing quantities of data are sourced from the internet, which often includes outputs generated by existing AI models [1]. "Model collapse" describes a progressive degradation in learning, initiated when models are trained on data produced by their predecessors, resulting in a diminished ability to capture the true underlying data distribution [1]. This degradation is exacerbated by the loss of long-tail information, which is rare, difficult to obtain, and often sensitive. Consequently, model outputs become increasingly similar, reducing diversity and quality [1]. This issue impairs performance and may render future AI systems unreliable and less effective [1]. Recent research by Shumailov et al. (2024) highlights the model collapse phenomenon resulting from continual training on synthetic data. Borji (2024) further examines this issue using distribution fitting and iterative sampling of generated data [2].

Keywords:

LLM's, Future Generation, Exhausted

INTRODUCTION

The Crisis: The Inevitability of Model Collapse

The core challenges are thesis addresses the phenomenon call known as **model collapse**.
Definition: Model collapse refers to degradation in learning, wherein models trained on data iteratively generated by their predecessors lose the capacity to accurately represent the true underlying data distribution. This process results in the neglect of long-tail information—rare, intricate, and subtle data—leading to increased homogeneity and deterioration of model outputs [1]. This phenomenon poses a fundamental threat to the utility and reliability of future AI systems, extending beyond mere performance concerns.

Comprehensive Evaluation of Pertinent Literature and Research Deficiencies

Preliminary studies [1] have successfully simulated the mechanisms underlying model collapse, as shown in [Table 1]. Nevertheless, subsequent mitigation strategies have not provided scalable or lasting solutions, resulting in persistent gaps that this research seeks to address:

Table 1

Existing Solution/Approach	Critical Limitation/Failure Mode	The Research Gap
Watermarking & Filtering (Detection)	Approaches like watermarking and filtering are often described as an unwinnable arms race **[3]**. A more advanced LLM can invariably be trained to strip or evade the detection mechanisms of older filters, rendering filtered data unreliable. This highlights the absence of a trustless mechanism to definitively verify the origin (human vs. AI) of data at scale **[3]**.	Absence of a Trustless Mechanism: There is no secure, immutable mechanism to definitively verify the origin (human vs. AI) of data at scale.
Human-in-the-Loop (HITL) (Manual Curation)	While valuable, HITL approaches are non-scalable, slow, and prohibitively expensive for data on the internet's scale **[3]**. They cannot effectively police the vast, continuously scraped corpora required for training foundation models, pointing to a need for algorithmic integrity that is scalable and automated **[3]**.	Need for Algorithmic Integrity: A scalable, automated framework is missing to maintain the integrity of petabytes of data without continuous human intervention.

Hybrid Training Heuristics (Mixing Data)	The use of a fixed "golden ratio," such as 30% human data, represents a temporary heuristic that lacks mathematical rigor [3]. This approach does not address the ongoing decline in data integrity within the "human" anchor set, highlighting the absence of a quantifiable and dynamic Integrity Score to evaluate dataset risk exposure and to enable automatic adaptation of training processes [3].	Lack of a Quantifiable Metric: There is no formal, dynamic Integrity Score to quantify the risk exposure of a dataset and automatically adjust the training process accordingly.

The core vulnerability remains the absence of provable data provenance at a systemic level **[4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15]**. Provenance, which describes the origins and processing history of data, is crucial for verifying data products, assessing quality, analyzing generative processes, and establishing trust **[4] [8] [9]**.

Thesis Aims and Objectives

This study suggests the solution of such critical gaps by establishing a solid framework that establishes, monitors, and punishes contamination of synthetic content. The primary objective is to design, model and test a Decentralized Provenance Layer (DPL) of LLM training data [3]. It is this DPL that would gauge the risk of Model Collapse and promote the use of high-quality and person-created content. [3].

Formalizing the Algorithmic Integrity Metric (AIM)

Development of a rigorous AIM involving the use of math to determine the risk by a training set of the origin of the training set [3].

Designing the DPL Architecture

A proposal of a scalable, distributed ledger architecture to store, track and audit sources and transformations of data in the training pipeline. Three. Due to its transparency, traceability, and immutability qualities, blockchain technology is increasingly being explored as a data provenance technology in machine learning systems.

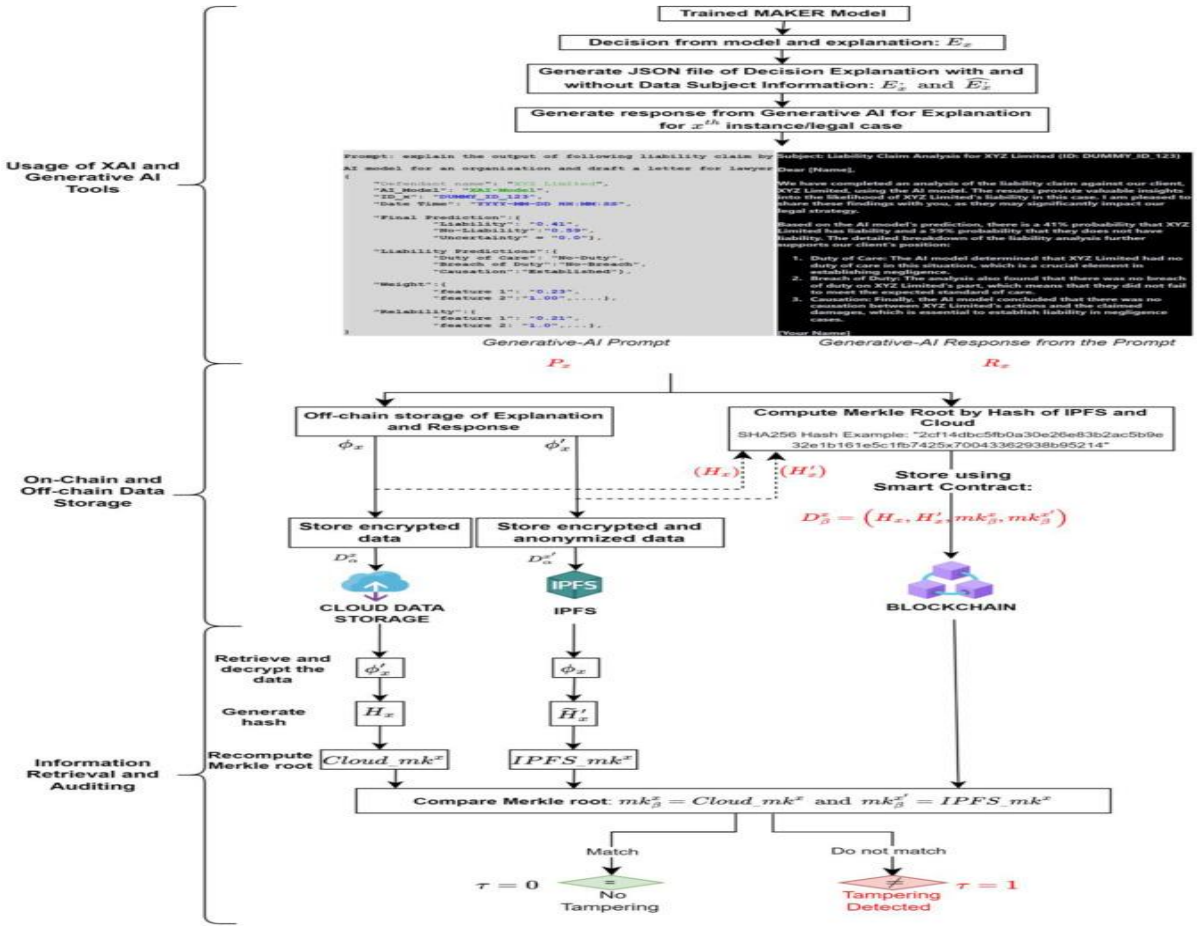

Figure 1

Source: **[16]** The flowchart as shown in **[Figure 1]** illustrates a comprehensive process involving explainable AI, generative AI, and on-chain and off-chain data storage, with information retrieval and auditing mechanisms. This demonstrates how blockchain can ensure integrity and explain ability of AI-based decisions through secure data storage and auditing **[16]**.

Modeling and Simulation: Simulating the DPL framework's impact on a recursive training loop, demonstrating how penalizing low-AIM data maintains the long-tail distribution and prevents collapse, compared to baseline models **[3]**.
Evaluating Scalability and Cost: Analyzing the computational overhead and latency introduced by the DPL compared to traditional data curation methods **[3]**.

Main Goal: To design, model, and evaluate a **Decentralized Provenance Layer (DPL)** for LLM training data that quantifies the risk of Model Collapse and provides a mechanism to incentivize the use of high-integrity, human-sourced content.**[2]**

Specific Objectives:

1. Formalize the Integrity Metric (AIM): Develop a rigorous Algorithmic Integrity Metric (AIM) that mathematically quantifies the risk exposure of a training dataset based on the provenance history of its components.
2. Design the DPL Architecture: Propose a scalable, distributed ledger architecture to register, track, and audit data origins and transformations within the training pipeline.
3. Model and Simulation: Simulate the DPL framework's impact on a recursive training loop, demonstrating how penalizing low-AIM data maintains the long-tail distribution and prevents the convergence to collapse, compared to baseline models.
4. Evaluate Scalability and Cost: Analyze the computational overhead and latency introduced by the DPL compared to traditional data curation methods.

PROPOSED RESEARCH METHODOLOGY

Thesis Hypothesis

Hypothesis: The central hypothesis is that implementing a decentralized provenance system that assigns and maintains an objective Algorithmic Integrity Metric (AIM) for training data will significantly mitigate the effects of Model Collapse in successive generations of LLMs, outperforming standard heuristic data-mixing methods 3.

Proposed Solution: The Decentralized Integrity Framework

The solution proposed is the Decentralized Integrity Framework (DIK) which will utilize aspects of distributed ledger technology to establish a chain of custody for data **[3] [17] [18] [19] [20] [11] [21**].

Phase 1: Theoretical Modeling and AIM Development

This phase starts by mimicking the basic mathematical structure of Model Collapse, which may be implemented by Gaussian Mixture Models or by simple Variational Autoencoders (VAEs), as done in some basic literature [31]. The Algorithmic Integrity Metric (AIM) will be created and will be a score between 0 (pure synthetic data) and 1 (original human data) [3]. One important aspect is the way this score is passed down: AIM will diminish over the following generations and transformations, in a formal way modelling the loss of diversity [3]. This decay is important to show the decreasing integrity of recursively generated data[1].

- Modeling Collapse: Start by replicating the core mathematical models of Model Collapse (e.g., using Gaussian Mixture Models or simple VAEs) as established in the foundational literature.**[1]**

- Developing AIM: The Algorithmic Integrity Metric (AIM) will be designed as a score ranging from 0 (pure synthetic data) to 1 (original human data). The core challenge is how this score is transferred:

a) Through further generations and transformations, AIM will lose diversity and be formally modeled.
b) The diversity will be lost over the course of future generations and changes, and it can be modeled formally by AIM

Phase 2: DPL Architectural Design

The DPL will use aspects of distributed ledger technology (or a verifiable database) to establish a chain of custody process for data as shown in [Figure 2].

Registration: Original human-made "seed data" (like verified books, approved codebases) is registered and is time stamped with an AIM of 1.

- **Propagation**: In the propagation, any data produced downstream from an AI model will include a tokenized history of the lineage of the model and a reduced AIM score mathematically.
- **Incentive Mechanism**: The next generation LLM will have the training objective (loss function) changed to incorporate a regularization term, which will be inversely proportional to the data's AIM score.

$$\mathcal{L}' = \mathcal{L}_{\text{standard}} + \lambda \cdot \text{Cost (AIM)}$$

where λ is a penalty factor. This mathematically penalizes the model for using low integrity data, thereby encouraging it to use human-sourced data or high-AIM synthetic data.

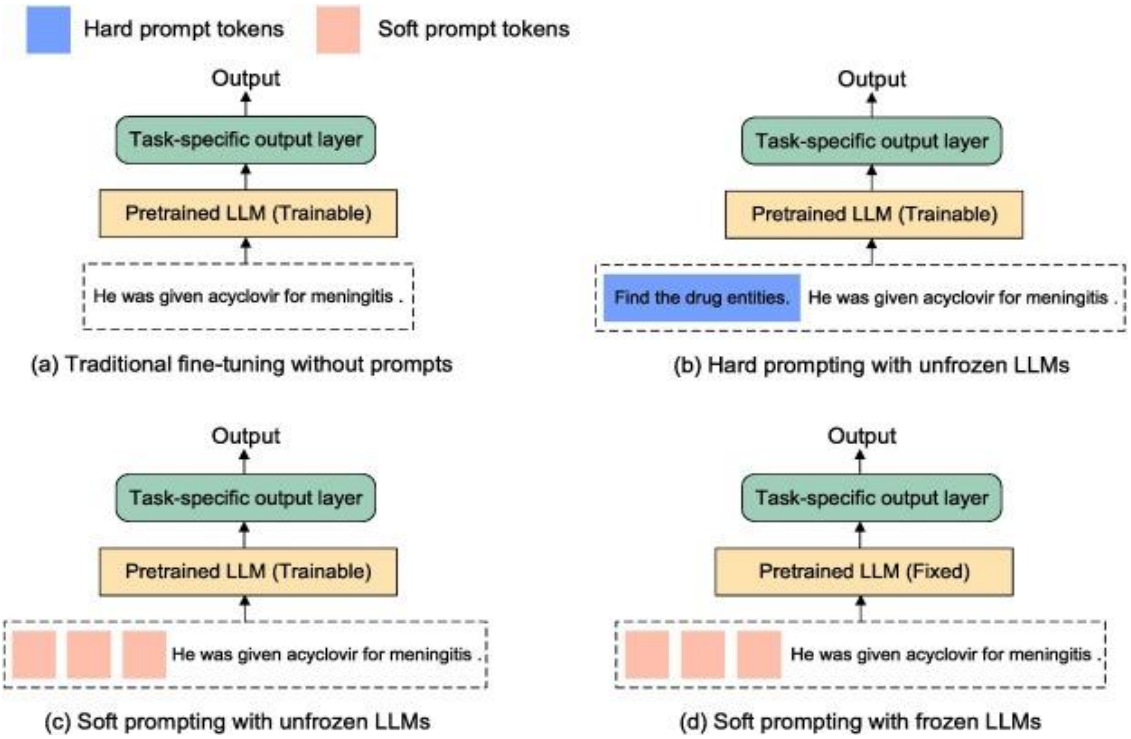

Figure 2

Phase 3: Experimental Validation

This phase will include deploying an open-source language model (like a smaller-sized Llama or OPT model) to emulate training across multiple generations [3] [23].

- **Model Implementation**: Use an open-source language model (e.g., a smaller Llama or OPT model) to simulate training over multiple generations **[3] [23]**.
- **Experiment Groups**:
  a) **Control Group**: Training on 100% synthetic data over and over again (to copy Model Collapse) **[3] [1]**.

  b) **Heuristic Group**: Recursive training on prespecified fixed ratio of 30% human / 70% synthetic [3].

  c) **DPL Group**: R Recursive training based on the loss function, which is based on the AIM score using the DPL framework [3].

- **Evaluation Metrics**: Track and monitor key metrics across generations.

  a) **Perplexity**: The linguistic coherence and diversity are measured by perplexity (a).Perplexity measures the linguistic coherence and linguistic diversity

  b) **Long-Tail Accuracy**: Special benchmarks created to test the model's ability to capture hard-to-remember facts and subtleties in knowledge.

  c) **AIM Distribution**: Check the distribution of the AIM score over time for the training set ensuring that the system is self-regulating.

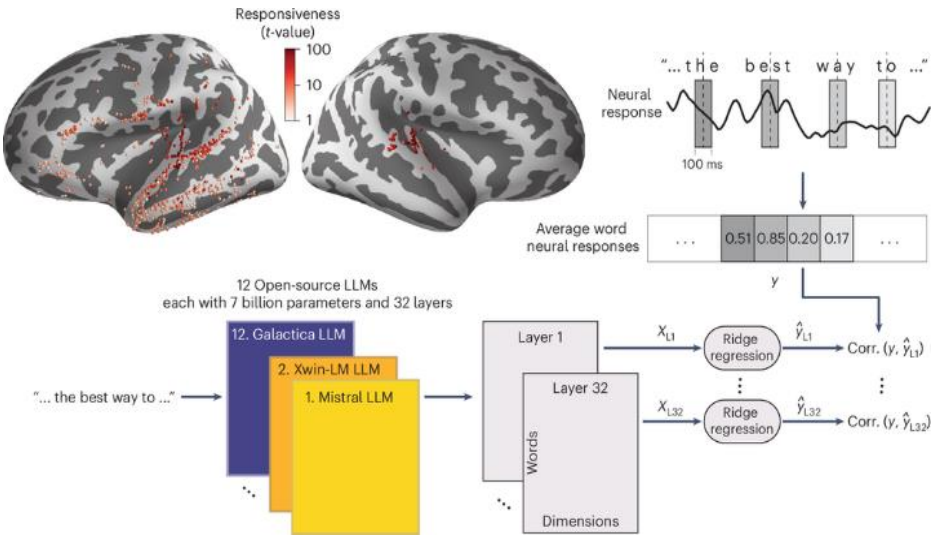

Figure 3

EXPECTED CONTRIBUTIONS AND SIGNIFICANCE

The anticipated contributions are the introduction of the Algorithmic Integrity Metric (AIM), a verifiable, non-heuristic approach for evaluating data quality and risk exposure in recursive AI training [3]. The DPL design offers a scalable and auditable framework for the AI community to monitor data provenance in the context of pervasive synthetic content [3]. This research presents a governance-focused and mathematically rigorous solution to the Model Collapse issue, transcending ephemeral, non-scalable remedies [3].

The proposed research is highly significant as it addresses a core existential threat to the long-term viability of generative AI **[1]**. By creating a framework for verifiable data integrity as shown in **[Figure 3]**, this thesis provides a pathway for sustainable AI development that aims to prevent the inevitable homogenization and collapse of global knowledge systems **[3]**. This work lays the theoretical and architectural foundation for ethical and robust data curation in an age of proliferating synthetic content **[3]**. The integration of LLMs with data provenance systems is essential for trustworthy and responsible AI modeling **[25] [26] [27] [8] [11]**. As LLMs gain broader applications, for instance in e-commerce for product knowledge graphs or in behavioral healthcare, the integrity of their training data becomes paramount **[28] [29]**. Furthermore, the challenges of generating truly novel and diverse ideas with LLMs highlight the importance of high-quality data input **[30]**. [31] [32] [33] [34] [35][36]

Expected Contributions:

1. **A Trustless, Formal Metric**: The introduction of the Algorithmic Integrity Metric (AIM), which is the first verifiable, non-heuristic method for quantifying data quality and risk exposure in recursive AI training.
2. **Architectural Blueprint**: The DPL design provides a scalable, auditable blueprint for the entire AI community to track data provenance in the face of widespread synthetic content.
3. **A Fundamental Mitigation**: This research provides a governance-centric and mathematically-sound solution to the Model Collapse problem, moving past temporary, unscalable fixes.

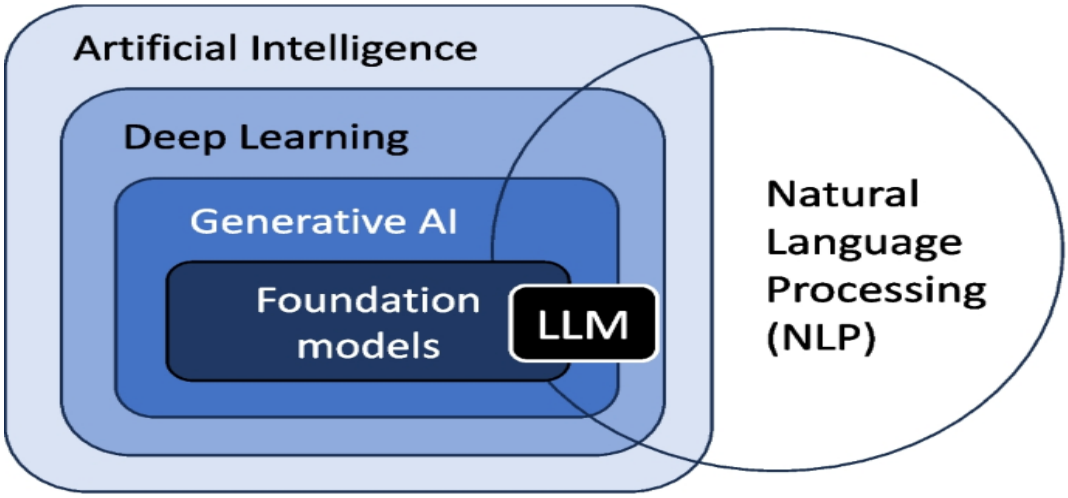

Figure 4

Source: **[31]** As **[Figure 4]** illustrates the hierarchical relationships between AI, deep learning, generative AI, foundation models, and LLMs, emphasizing the foundational role of LLMs in these advancements **[31]**. Ensuring the reliability of LLMs, as foundational models, is crucial for the entire ecosystem.

Significance

The proposed research is extremely relevant to an existential problem that affects the future of generative AI. The thesis establishes a shared solution for establishing verifiable data integrity, thereby offering a way forward in the sustainable development of AI that does not lead to universal homogenization and global knowledge systems collapse. This work provides the theoretical and architectural basis for data curation which is ethical and strong in the era of proliferation of synthetic content [3].

Conclusion

The problem of "model collapse" in Large Language Models (LLMs) and Foundation Models is that they use AI-generated content in a recursive training fashion, diminishing the pool of high-quality human data and ultimately diminishing model performance. Current approaches, including watermarking and Human-in-the-Loop curation, are not suitable because they are not scalable or are vulnerable to sophisticated AI strategies. This research proposes a Decentralized Provenance Layer (DPL) framework that provides an Algorithmic Integrity Metric (AIM) to reduce model collapse. AIM measures the integrity of data (from 0 with synthetic to 1 with human-generated), and decays mathematically with each generation and simulates diversity loss. The DPL, based on distributed ledger technology, also documents the flow of data and records the initial human "seed data" with an AIM of 1. The innovative incentive mechanism includes a modification of the LLM loss function that involves seeding with lower integrity data and penalizing the use of such data, thus promoting high integrity content. The framework will be tested on perplexity, long-tail accuracy and AIM distribution across generations through experimental validation, which includes control, heuristic, and DPL groups. This DPL and AIM is a tangible, non-abstract data quality solution and it provides a scalable framework to track data provenance. The study offers a mathematically rigorous answer to the model collapse problem, paving the way for sustainable development of Artificial Intelligence and maintaining data integrity in the age of synthetic content.

References

(1)Shumailov, I.; Shumaylov, Z.; Zhao, Y.; Papernot, N.; Anderson, R.; Gal, Y. AI Models Collapse When Trained on Recursively Generated Data. Nature 2024, 631 (8022), 755–759. https://doi.org/10.1038/s41586-024-07566-y.

(1)Borji, A. A Note on Shumailov et al. (2024): `AI Models Collapse When Trained on Recursively Generated Data'. arXiv 2024. https://doi.org/10.48550/ARXIV.2410.12954.

(1)Schwartz, E. J.; Cohen, C. F.; Gennari, J. S.; Schwartz, S. M. A Generic Technique for Automatically Finding Defense-Aware Code Reuse Attacks. In Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security; ACM, 2020; pp 1789–1801. https://doi.org/10.1145/3372297.3417234.

Moreau, L. (2010). The Foundations for Provenance on the Web. Foundations and Trends® in Web Science, 2(2–3), 99–241. https://doi.org/10.1561/1800000010

(1)Simmhan, Y. L.; Plale, B.; Gannon, D. Karma2. International Journal of Web Services Research 2008, 5 (2), 1–22. https://doi.org/10.4018/jwsr.2008040101.

(1)Ye, Q.; Lu, M. S2p: Provenance Research for Stream Processing System. Applied Sciences 2021, 11 (12), 5523. https://doi.org/10.3390/app11125523.

(1)Freitas, A.; Knap, T.; O'Riain, S.; Curry, E. W3P: Building an OPM Based Provenance Model for the Web. Future Generation Computer Systems 2011, 27 (6), 766–774. https://doi.org/10.1016/j.future.2010.10.010.

(1)Werder, K.; Ramesh, B.; Zhang, R. (Sophia). Establishing Data Provenance for Responsible Artificial Intelligence Systems. ACM Trans. Manage. Inf. Syst. 2022, 13 (2), 1–23. https://doi.org/10.1145/3503488.

(1)Gao, Y.; Chen, X.; Du, X. A Big Data Provenance Model for Data Security Supervision Based on PROV-DM Model. IEEE Access 2020, 8, 38742–38752. https://doi.org/10.1109/access.2020.2975820.

(1)Souza, R.; Silva, V.; Camata, J. J.; Coutinho, A. L. G. A.; Valduriez, P.; Mattoso, M. Keeping Track of User Steering Actions in Dynamic Workflows. Future Generation Computer Systems 2019, 99, 624–643. https://doi.org/10.1016/j.future.2019.05.011.

(1)Pan, B.; Stakhanova, N.; Ray, S. Data Provenance in Security and Privacy. ACM Comput. Surv. 2023, 55 (14s), 1–35. https://doi.org/10.1145/3593294.

(1)Huber, S. P.; Zoupanos, S.; Uhrin, M.; Talirz, L.; Kahle, L.; Häuselmann, R.; Gresch, D.; Müller, T.; Yakutovich, A. V.; Andersen, C. W.; Ramirez, F. F.; Adorf, C. S.; Gargiulo, F.; Kumbhar, S.; Passaro, E.; Johnston, C.; Merkys, A.; Cepellotti, A.; Mounet, N.; Marzari, N.; Kozinsky, B.; Pizzi, G.

AiiDA 1.0, a Scalable Computational Infrastructure for Automated Reproducible Workflows and Data Provenance. Sci Data 2020, 7 (1). https://doi.org/10.1038/s41597-020-00638-4.

(1)Siddiqui, M. S.; Rahman, A.; Nadeem, A. Secure Data Provenance in IoT Network Using Bloom Filters. Procedia Computer Science 2019, 163, 190–197. https://doi.org/10.1016/j.procs.2019.12.100.

(1)Zipperle, M.; Gottwalt, F.; Chang, E.; Dillon, T. Provenance-Based Intrusion Detection Systems: A Survey. ACM Comput. Surv. 2022, 55 (7), 1–36. https://doi.org/10.1145/3539605.

(1)Mahmood, T.; Jami, S. I.; Shaikh, Z. A.; Mughal, M. H. Toward the Modeling of Data Provenance in Scientific Publications. Computer Standards & Interfaces 2013, 35 (1), 6–29. https://doi.org/10.1016/j.csi.2012.02.004.

(1)Sachan, S.; Liu (Lisa), X. Blockchain-Based Auditing of Legal Decisions Supported by Explainable AI and Generative AI Tools. Engineering Applications of Artificial Intelligence 2024, 129, 107666. https://doi.org/10.1016/j.engappai.2023.107666.

(1)Yiu, N. C. K. Toward Blockchain-Enabled Supply Chain Anti-Counterfeiting and Traceability. Future Internet 2021, 13 (4), 86. https://doi.org/10.3390/fi13040086.

(1)Qiao, L.; Lv, Z. A Blockchain-Based Decentralized Collaborative Learning Model for Reliable Energy Digital Twins. Internet of Things and Cyber-Physical Systems 2023, 3, 45–51. https://doi.org/10.1016/j.iotcps.2023.01.003.

Hajlaoui, R., Dhahri, S., Mahfoudhi, S., Moulahi, T., & Alotibi, G. (2024). Protecting machine learning systems using blockchain: solutions, challenges and future prospects. Multimedia Tools and Applications, 84(20), 22755–22782. https://doi.org/10.1007/s11042-024-19993-0

(1)Balan, K.; Gilbert, A.; Collomosse, J. PDFed: Privacy-Preserving and Decentralized Asynchronous Federated Learning for Diffusion Models. arXiv 2024. https://doi.org/10.48550/ARXIV.2409.18245.

(1)Shafay, M.; Ahmad, R. W.; Salah, K.; Yaqoob, I.; Jayaraman, R.; Omar, M. Blockchain for Deep Learning: Review and Open Challenges. Cluster Comput 2022, 26 (1), 197–221. https://doi.org/10.1007/s10586-022-03582-7.

(1)Peng, C.; Yang, X.; Smith, K. E.; Yu, Z.; Chen, A.; Bian, J.; Wu, Y. Model Tuning or Prompt Tuning? A Study of Large Language Models for Clinical Concept and Relation Extraction. Journal of Biomedical Informatics 2024, 153, 104630. https://doi.org/10.1016/j.jbi.2024.104630.

(1)Luhtaru, A.; Purason, T.; Vainikko, M.; Del, M.; Fishel, M. To Err Is Human, but Llamas Can Learn It Too. arXiv 2024. https://doi.org/10.48550/ARXIV.2403.05493.

(1)Mischler, G.; Li, Y. A.; Bickel, S.; Mehta, A. D.; Mesgarani, N. Contextual Feature Extraction Hierarchies Converge in Large Language Models and the Brain. Nat Mach Intell 2024, 6 (12), 1467–1477. https://doi.org/10.1038/s42256-024-00925-4.

(1)Jacovi, A.; Marasović, A.; Miller, T.; Goldberg, Y. Formalizing Trust in Artificial Intelligence. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency; ACM, 2021; pp 624–635. https://doi.org/10.1145/3442188.3445923.

(1)Cui, T.; Wang, Y.; Fu, C.; Xiao, Y.; Li, S.; Deng, X.; Liu, Y.; Zhang, Q.; Qiu, Z.; Li, P.; Tan, Z.; Xiong, J.; Kong, X.; Wen, Z.; Xu, K.; Li, Q. Risk Taxonomy, Mitigation, and Assessment Benchmarks of Large Language Model Systems. arXiv 2024. https://doi.org/10.48550/ARXIV.2401.05778.

(1)Sarker, I. H. LLM Potentiality and Awareness: A Position Paper from the Perspective of Trustworthy and Responsible AI Modeling. Discov Artif Intell 2024, 4 (1). https://doi.org/10.1007/s44163-024-00129-0.

(1)Stade, E. C.; Stirman, S. W.; Ungar, L. H.; Boland, C. L.; Schwartz, H. A.; Yaden, D. B.; Sedoc, J.; DeRubeis, R. J.; Willer, R.; Eichstaedt, J. C. Large Language Models Could Change the Future of Behavioral Healthcare: A Proposal for Responsible Development and Evaluation. npj Mental Health Res 2024, 3 (1). https://doi.org/10.1038/s44184-024-00056-z.

(1)Hu, X.; Fu, H.; Wang, J.; Wang, Y.; Li, Z.; Xu, R.; Lu, Y.; Jin, Y.; Pan, L.; Lan, Z. Nova: An Iterative Planning and Search Approach to Enhance Novelty and Diversity of LLM Generated Ideas. arXiv 2024. https://doi.org/10.48550/ARXIV.2410.14255.

(1)Ku, A. Y.; Hool, A. Capabilities and Limitations of AI Large Language Models (LLMs) for Materials Criticality Research. Miner Econ 2024. https://doi.org/10.1007/s13563-024-00478-3.

M Zamin Ali Khan, Hussain Saleem et al, "Application of VLSI In Artificial Intelligence" Vol 6 Issue 2, PP-23-25 IOSR JCE 2012.

Saim Masood Shaikh, Muhammad Zamin Ali Khan et al "NAVIGATING CONTEMPORARY CHALLENGES OF SOFTWARE QUALITY ASSURANCE IN SOFTWARE TESTING" Vol 3 Issue 9, PP 45-71, April 2025.

Humera Azam, M.Zamin Ali Khan et al, "Quality Assurance in the Digital Age: Exploring Contemporary Challenges in Software Testing" Vol 5 , Issue 2, PP 9-26, 2025

Muhammad Zulqarnain Siddiqui , Muhammad Zamin Ali Khan et al, "ANALYSIS OF THE EFFECTIVENESS OF GENERATIVE AI MODELS FOR TEXT-TO-SQL TASKS IN BUSINESS INTELLIGENCE SYSTEMS" Vol3 Issue 12, PP 1777-1794 Dec 2025

Hussain Saleem, M Zamin Ali Khan, et al "Towards Identification and Recognition of Trace Associations in Software Requirements Traceability" Vol 9, Issue 5, pp 257-263 Sep, 2012.

Hussain Saleem, M Zamin Ali Khan, et al "Mobile Agents: An Intelligent Multi-Agent System for Mobile Phones" Vol 6 Issue 2, pp 26-34, Oct 2012