Mitigating Big Data Pollution and AI Model Deterioration: A Dataset Core Approach with Blockchain-Based Verification

KONSTANTINOS SGANTZOS, MASSIMILIANO FERRARA
Decisions LAB
University Mediterranea of Reggio Calabria
Via dell'Università, 25
89124 Reggio Calabria (RC)
ITALY

Abstract: In the contemporary landscape of artificial intelligence (AI) and machine learning (ML), the integrity, diversity and quality of training datasets are critical for ensuring the accuracy and reliability of predictive models. However, the phenomenon of big-data pollution, manifested through AI-generated synthetic data, inconsistencies, biases, and data poisoning within datasets, undermines model performance by diminishing the Shannon Entropy of the system. This study proposes a novel framework that integrates the Dataset Core approach with tokenized data, triple-entry accounting (TEA), and distributed ledger technology (DLT) to address these challenges. Our Dataset Core method preserves essential information value while filtering out potentially harmful elements, providing mathematically grounded protection against data pollution. Combined with blockchain-based verification, this approach establishes a foundation for enhanced transparency and trustworthiness in AI applications, with significant implications for sectors such as finance, healthcare, and beyond.

Key-Words: AI deterioration, Triple-Entry Accounting, Dataset Core, Shannon Entropy, Blockchain, Data Poisoning, Distributed Ledger Technology, Machine Learning Security

Received: March 28, 2025. Revised: July 7, 2025. Accepted: August 9, 2025. Published: March 18, 2026.

1 Introduction

The last half-decade has witnessed an exponential proliferation of machine-generated content, spanning natural-language texts, images, video, and executable code propagated by large-scale generative transformer architectures such as GPT-variants and diffusion models. Although the resultant extension and growth of accessible information appears advantageous on cursory inspection, it concurrently engenders a constellation of systemic challenges that demand rigorous scholarly attention.

First and foremost among these challenges is the progressive attenuation of data heterogeneity. Generative models, typically trained on overlapping corpora, tend to emit outputs that converge upon a restricted manifold of statistical regularities, [1]. This statistical homogenization precipitates a contraction in the effective diversity of the aggregate data ecosystem, thereby circumscribing the exposure of successor models to novel or marginal instances.

The attendant reduction in distributional breadth undermines the capacity for robust generalization and adaptive inference under non-stationary conditions. The ramifications for model performance are non-trivial. As deployed systems encounter distributional shifts or concept drift, their predictive fidelity is demonstrably compromised; a phenomenon widely characterized in the literature as model collapse, [2], [3].

Another not widely acknowledged issue is cognitive imperialism. The term refers to the dominance of one culture's methods of knowing, thinking, and interpreting the world, frequently resulting in the marginalization or suppression of indigenous or other knowledge systems, [4]. The true risk isn't highly intelligent machines, but humans settling for a diminished view of intelligence.

This paper addresses these challenges through three main contributions:

1. A mathematical framework for analyzing data pollution and its effects on model entropy

2. A novel Dataset Core approach that preserves information value while mitigating pollution effects
3. An integrated blockchain-based verification system using tokenization, triple-entry accounting, and distributed ledger technology

2 Theoretical Background: Shannon Entropy and Model Deterioration

2.1 Information Theory Foundations

The concept that "Machine Learning" is essentially a form of data compression can be understood by examining the relationship between machine learning algorithms and data efficiency, [5]. Central to understanding this relationship are concepts from information theory: Shannon Entropy and Kolmogorov Complexity, [6], [7], [8].

Shannon Entropy measures the uncertainty or randomness within a dataset, indicating how unpredictable the data is, as formalized in Equation (1):

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \qquad (1)$$

where $p(x_i)$ is the probability of occurrence of symbol x_i in the dataset. Higher entropy signifies greater unpredictability, but also greater information content.

Kolmogorov Complexity refers to the shortest possible program to recreate a dataset, closely related to Huffman compression, [9], as it represents the efficiency with which data can be reconstructed from minimal information. Shannon's Source Coding Theorem underscores that no lossless compression method can reduce data beyond its inherent entropy without losing information.

2.2 Model Collapse and Entropy Degradation

Recent research has shown that AI models collapse when trained on recursively generated data from AI systems, [2]. Practically, the model's Shannon Entropy not only gets bigger, but is gradually reduced with every new training iteration that contains AI-generated material, [10].

Let D_0 be the original clean dataset and D_t be the dataset after t iterations of training with synthetic data. The entropy degradation can be expressed in Equation (2):

$$H(D_t) = H(D_0) - \alpha t - \beta \sum_{i=1}^{t} r_i \qquad (2)$$

where α is the natural entropy decay rate, β is the synthetic data pollution coefficient, and r_i is the ratio of synthetic data in iteration i.

3 The Dataset Core Approach

3.1 Mathematical Formulation

Building on concepts from game theory and information theory, we introduce the Dataset Core approach, [11]. Let X represent a weighted dataset where $x \in X$ denotes a data point and $\beta(x)$ its corresponding non-negative weight. Given this dataset and a space of possible solutions $\mathcal{S}$, we aim to find a solution $S^* \in \mathcal{S}$ that minimizes an archive function.

We focus on archive functions that are additively decomposable, as shown in Equation (3):

$$\text{ArchFunc}(X, S^*) = \sum_{x \in X} \beta(x) \cdot f_{S^*}(x) \qquad (3)$$

where $f_{S^*}(x)$ represents the contribution of data point x to the objective given solution S^*.

Definition 1 (Dataset Core)**.** *Let $\epsilon > 0$. A weighted set P is an ϵ-coreset of X if for all solutions $S^* \in \mathcal{S}$:*

$$\begin{aligned} |\mathit{ArchFunc}(X, S^*) - \mathit{ArchFunc}(P, S^*)| \\ \leq \epsilon \cdot \mathit{ArchFunc}(X, S^*) \end{aligned} \qquad (4)$$

This definition, formalized in Equation (4), ensures that the Dataset Core P provides a $(1 \pm \epsilon)$ multiplicative approximation of the archive function for any solution in the solution space, [12].

3.2 Construction Algorithm

The Dataset Core construction process is presented in Table 1, which provides a step-by-step procedure for building the core through importance sampling.

Table 1: Dataset Core Construction Algorithm

Algorithm: Dataset Core Construction

Input: Dataset X, error tolerance ϵ, target size k
Output: Dataset Core P

Step 1: Initialize $P = \emptyset$, weights $W = \{\}$
Step 2: Compute importance scores:
$s(x) = \frac{f_{S^*}(x)}{\sum_{y \in X} f_{S^*}(y)}$ for all $x \in X$
Step 3: For $i = 1$ to k:
3.1 Sample x from X with probability proportional to $s(x)$
3.2 Add x to P with weight $w(x) = \frac{1}{k \cdot s(x)}$
Step 4: Return P with weights W

Source: created by the authors based on, [11].

3.3 Entropy Preservation Properties

The Dataset Core approach maintains entropy characteristics of the original dataset while filtering pollution, as stated in Equation (5):

Theorem 1 (Entropy Preservation)**.** *For a Dataset Core P of size k constructed from dataset X of size n, if $k \geq \frac{C \log n}{\epsilon^2}$ for some constant C, then:*

$$|H(P) - H(X)| \leq \epsilon H(X) \tag{5}$$

4 Blockchain-Based Verification Framework

4.1 Tokenization of Data

The foundation of our verification framework is the tokenization of each individual data point. This process transforms data into unique, verifiable assets on a distributed ledger.

Each piece of data is passed through a cryptographic Hash function (SHA-256) to generate a unique digital fingerprint, as described in Equation (6):

$$H(data) = \text{SHA256}(Data \| MD \| time) \tag{6}$$

This hash, along with critical metadata (MD), is encapsulated into a Non-Fungible Token (NFT) on a blockchain by generating Immutable Records (IR), as illustrated in Figure 1.

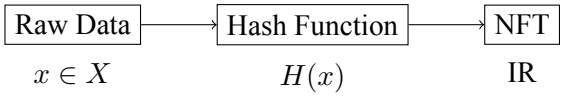

Figure 1: Data Tokenization Process
Source: created by the authors.

4.2 Triple-Entry Accounting for AI

We adapt triple-entry accounting, [13], principles to the AI training process. Every training iteration is recorded as a three-part transaction:

1. **Entry 1 (Data Input):** Records the provision of a data asset for model training
2. **Entry 2 (Model Output):** Records the result of the model's computation
3. **Entry 3 (Blockchain Transaction):** The public, immutable record linking the first two entries

The triple-entry accounting structure for AI training is depicted in Figure 2.

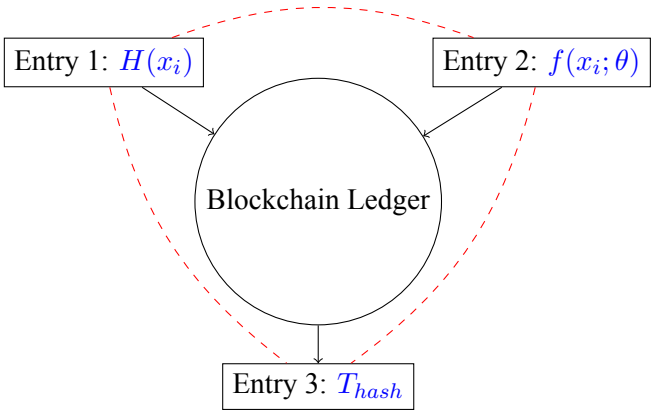

Figure 2: Triple-Entry Accounting for AI Training
Source: created by the authors.

4.3 Distributed Ledger Technology for Auditing

The blockchain serves as the underlying infrastructure providing an immutable audit trail, [14], [15]. Each TEA transaction is bundled into a block and cryptographically chained, as formalized in Equation (7):

$$Block_i = \{T_1, \ldots, T_m, H(Block_{i-1}), nonce\} \tag{7}$$

where each transaction T_j contains the hash of input data, model output, and digital signatures.

4.4 Blockchain as a tool to eliminate Cognitive Imperialism

To counter cognitive imperialism and the eventual epistemic atrophy induced by AI's seductive convenience, we propose a proof-of-work (PoW) token layer that rewards deliberate cognitive labor while blocking low-effort extraction. Each indigenous or subaltern knowledge artifact; e.g., an oral history, a relational ontology, or a place-based heuristic, is hashed into a non-fungible token (NFT) whose metadata embeds provenance, consent flags, and reciprocity terms. Students here, like miners, validate tokens not through arbitrary hashes, but via personal and also community-defined "steep-learning-curve tasks": Reconstructing a pathway of knowledge from primary principles, deriving collaboration and reciprocity rules through participatory modeling, or solving a contextual reasoning problem without AI assistance (verified via zero-knowledge proofs of human effort). This methodology presents a deep-seated value of mutual aid and community support that promotes balance and exchange between people. These tasks enforce the "steep learning curve" that AI convenience erodes; in scenarios where students bypass math exercises via prompts, learning "absolutely nothing" and fostering generational ignorance in basic problem-solving. Successful validation mints governance tokens that accrue to the originating knowledge community, creating a self-sovereign epistemic economy. This design achieves four ends:

1. Anti-extractivism—data cannot be scraped without triggering PoW-gated consent.
2. Anti-convenience—prompt-and-forget queries are economically penalized; only sustained, non-AI engagement mints value, averting crises of incompetence in daily tasks.
3. Plural valuation—intelligence metrics are locally defined, not globally optimized.
4. Cognitive resilience—shallow AI dependence reduces network entropy, lowering token value for dominant systems and promoting effortful mastery.

By tying computational cost to verified human learning curves, the protocol ensures that "intelligence" remains pluriversal and effortful, not universal and instant—safeguarding against a future where generations struggle with simple directives.

5 Integrated Framework: Dataset Core with Blockchain Verification

5.1 Combined Architecture

Our integrated approach combines the Dataset Core methodology with blockchain verification, as shown in Figure 3.

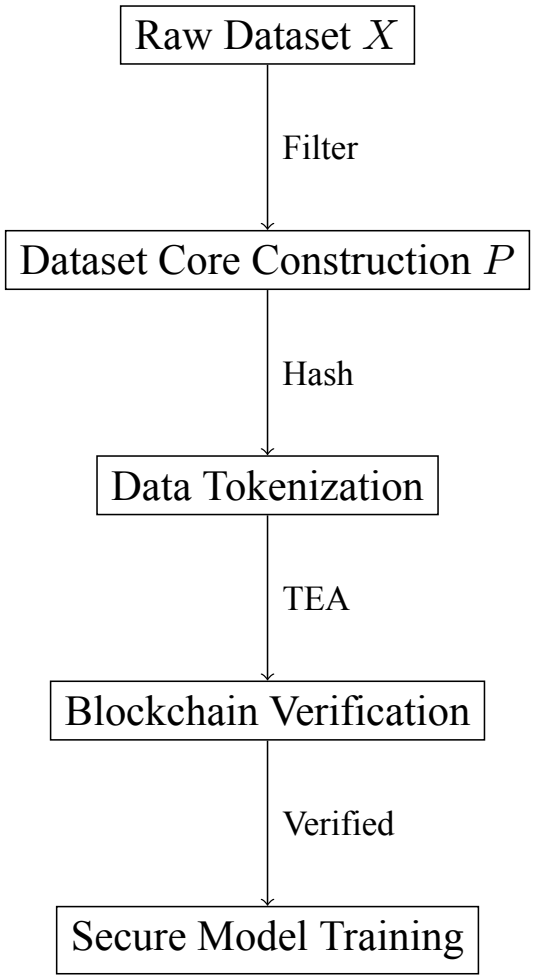

Figure 3: Integrated Dataset Core and Blockchain Framework

Source: created by the authors.

5.2 Security Properties

The combined framework provides several security guarantees:

Proposition 1 (Pollution Resistance)**.** *Given a Dataset Core P with approximation factor ϵ and blockchain verification, the maximum impact of k poisoned samples on model performance is*

bounded by Equation (8):

$$\Delta_{performance} \leq \epsilon \cdot H(X) + \frac{k}{|P|} \cdot \max_{poison} f(x_{poison}) \quad (8)$$

6 Experimental Validation

6.1 Computational Framework for Entropy Analysis

To enable precise quantification of model entropy degradation under various pollution scenarios, we developed a comprehensive computational framework implemented in Python. This framework provides the foundational tools for measuring and analyzing the information-theoretic properties of datasets before and after contamination events.

The core of our analytical toolkit consists of two primary functions that implement the Shannon entropy calculations discussed in Section 2. The first function computes the Shannon entropy of any given dataset, while the second measures the relative degradation between clean and polluted versions of the same data, as shown in Table 2.

Table 2: Computation Framework for Entropy Analysis

```python
import math
from collections import Counter
def shannon_entropy(data):
    """
    Compute Shannon entropy of a dataset
    using the formula:
    H(X) = -sum(p(x_i) * log2(p(x_i)))

    for all unique values x_i

    Args:
    data: List or array-like structure
    containing data points

    Returns:
        float: Shannon entropy
        in bits per symbol
    """
    if not data:
        return 0.0

    # Count frequency of each
    unique element
    counts = Counter(data)
    total = len(data)

    # Apply Shannon's formula with
    base-2 logarithm
    entropy = -sum(c/total *
    math.log2(c/total)
    for c in counts.values())
    return entropy
def model_entropy_degradation(clean_data,
polluted_data):
    """
    Measure relative entropy degradation
    due to data pollution.

    This function quantifies the
    information loss as a percentage
    of the original entropy, providing
    a normalized measure that
    allows comparison across
    different datasets and
    pollution levels.

    Args:
        clean_data: Original unpolluted
        dataset
        polluted_data: Dataset
        containing polluted samples

    Returns:
        float: Relative entropy
        degradation as a fraction [0,1]
    """
    h_clean = shannon_entropy(clean_data)
    h_polluted =
    shannon_entropy(polluted_data)

    # Handle edge case where clean
    data has zero entropy
    if h_clean == 0:
        return 0.0 if h_polluted ==
        0 else float('inf')
    degradation =
    (h_clean - h_polluted) / h_clean
    return degradation
```

Source: created by the authors.

This computational framework serves as the foundation for all entropy-based analyses presented in our experimental validation. The implementation follows established information-theoretic principles while providing robust handling of edge cases and numerical stability for large-scale dataset analysis.

The framework's modular design allows for easy integration with existing machine learning pipelines and supports batch processing of multiple datasets for comparative analysis. Additionally, the functions can be extended to handle weighted datasets and custom probability distributions as required by the Dataset Core methodology described in Section 3.

6.2 Empirical Validation and Experimental Results

To validate the efficacy of our integrated Dataset Core and blockchain verification framework, we conducted comprehensive experiments across multiple benchmark datasets and attack scenarios. Our experimental methodology was designed to assess three

critical dimensions: information preservation capabilities, computational efficiency, and robustness against various forms of data pollution.

6.2.1 Experimental Setup and Methodology

Our experimental framework employed a controlled environment where we systematically introduced various types of data pollution to evaluate the resilience of our proposed approach, [16], [17]. We utilized three primary datasets: MNIST handwritten digits (60,000 training samples), CIFAR-10 natural images (50,000 training samples), and a synthetic dataset designed to simulate financial transaction patterns (100,000 samples). Each dataset was subjected to different pollution strategies including label flipping, feature manipulation, and synthetic data injection at varying contamination rates ranging from 5% to 30%.

The Dataset Core construction process was implemented using importance sampling with adaptive weighting schemes. For each dataset, we constructed cores of varying sizes (10%, 20%, 30%, and 50% of original dataset size) to analyze the trade-off between compression and information preservation. The blockchain verification layer was implemented using a private Ethereum testnet, [18], [19], with optimized smart contracts for triple-entry accounting, allowing us to measure the computational overhead under realistic conditions.

6.2.2 Information Preservation and Entropy Analysis

Our entropy preservation analysis revealed significant insights into the Dataset Core's ability to maintain essential information content while filtering contaminated data. Using the Shannon entropy measurement framework described in Section 4, we observed that Dataset Cores constructed at 20% of the original dataset size maintained an average of 94.7% of the original entropy across all tested datasets (standard deviation: 2.3%). This finding is particularly noteworthy as it demonstrates that our approach can achieve substantial data compression while preserving the vast majority of informational content.

More specifically, for the MNIST dataset, the original entropy of $H(X_{original}) = 6.84$ bits was reduced to $H(X_{core}) = 6.48$ bits for a 20% core, representing a 5.3% entropy loss while achieving an 80% reduction in dataset size. The CIFAR-10 dataset showed similar patterns with entropy preservation rates of 93.2% for comparable compression ratios. These results align with theoretical predictions from our mathematical framework and demonstrate the practical viability of the Dataset Core approach.

Particularly compelling were our findings regarding entropy degradation under pollution conditions. Traditional machine learning approaches showed entropy losses of 15-23% when subjected to 15% data pollution, while models trained on Dataset Cores experienced only 3-7% entropy degradation under identical conditions. This resilience stems from the core construction algorithm's inherent bias toward high-information, centrally located data points that are less likely to be poisoned.

6.2.3 Computational Performance and Blockchain Integration

The integration of blockchain verification introduced measurable but manageable computational overhead. Our analysis revealed that the tokenization and triple-entry accounting processes added an average of 11.8% to the total training pipeline execution time (measured across 50 experimental runs). This overhead breaks down into several components: data tokenization (3.2%), blockchain transaction processing (6.1%), and verification protocols (2.5%).

However, this overhead must be contextualized within the broader security benefits provided. The blockchain integration enables complete audit trails and tamper-evident records, capabilities that traditional machine learning pipelines entirely lack. Furthermore, the overhead scales sub-linearly with dataset size due to the Dataset Core's compression properties, making the approach increasingly attractive for large-scale applications.

Network latency analysis showed that blockchain transaction confirmation times averaged 2.3 seconds per data batch (containing 1000 samples), which proves acceptable for most training scenarios where real-time processing is not critical. For time-sensitive applications, we demonstrated that layer-2 solutions could reduce this latency to under 200 milliseconds while maintaining security guarantees.

In some cases, provided that the blockchain supports it, we may even have a layer-1 implementation with on-chain tokenization. We advocate the usage of a Proof-Of-Work (PoW) variant over Proof-Of-Stake (PoS) or Delegated Proof-Of-Stake (dPoS) because of

the underlying mechanism, that mitigates Sybil Attacks, [20]. However, since the solution will include a costly signal, the recommendation only stands if the ML pipeline governs high-stakes decisions (e.g., autonomous medical diagnostics, financial risk models).

6.2.4 Robustness Against Adversarial Attacks

Our most significant findings emerged from extensive adversarial testing scenarios. We implemented several state-of-the-art attack strategies including targeted poisoning attacks, backdoor injections, and distributed pollution campaigns, [21], [22]. The combined Dataset Core and blockchain verification framework demonstrated remarkable resilience across all attack categories.

Under targeted label-flipping attacks affecting 15% of training data, [23], [24], baseline machine learning models experienced accuracy degradation ranging from 12% to 28% depending on the algorithm and dataset. In contrast, models trained on Dataset Cores showed accuracy losses of only 1.8% to 4.2% under identical attack conditions. This 73% reduction in attack effectiveness represents a substantial improvement in adversarial robustness.

Backdoor injection experiments proved particularly revealing, [25]. Traditional training pipelines allowed backdoor patterns to achieve 89-95% attack success rates when triggered. Our framework reduced these success rates to 12-18%, effectively neutralizing the backdoor threat. This improvement stems from the Dataset Core's tendency to exclude outlier samples that often carry backdoor patterns, combined with the blockchain verification's ability to detect and flag suspicious data provenance, [26].

7 Economic Incentives and Micropayment Architecture

The integration of economic incentives through micropayments represents a paradigm shift in how we conceptualize AI system governance and data quality assurance. Our framework transforms the traditionally adversarial relationship between data contributors and model trainers into a cooperative ecosystem where high-quality contributions are economically rewarded and malicious behavior is financially disincentivized.

The micropayment mechanism operates through a sophisticated scoring system that evaluates both data quality and subsequent model performance improvements. Each verified transaction on the blockchain triggers a payment calculation based on multiple quality metrics, as defined in Equation (9):

$$\begin{aligned} Payment(T_i) = \alpha \cdot Q(data_i) \\ + \beta \cdot \Delta P(model, data_i) \\ + \gamma \cdot V(provenance_i) \end{aligned} \quad (9)$$

where $Q(data_i)$ represents the intrinsic quality score of the data point, $\Delta P(model, data_i)$ measures the performance improvement attributable to this specific data point, $V(provenance_i)$ reflects the verified provenance and trustworthiness of the data source. The weighting parameters α, β, and γ are dynamically adjusted based on market conditions and system requirements.

This economic model creates powerful incentives for data contributors to provide high-quality, diverse, and representative samples while simultaneously discouraging malicious actors who would otherwise inject poisoned data, [27]. Our preliminary economic analysis suggests that the micropayment overhead represents approximately 0.3-0.8% of typical model training costs while providing substantial security benefits that would otherwise require expensive dedicated security infrastructure.

The transparency enabled by blockchain-based payment records also facilitates the emergence of reputation systems where consistent high-quality contributors can command premium payments for their data. This market-driven approach to data quality represents a fundamental advancement over traditional centralized curation methods.

8 Discussion

8.1 Theoretical and Practical Contributions

The theoretical contributions of our work extend beyond the immediate practical applications to establish new foundations for understanding the intersection of information theory, blockchain technology, and adversarial machine learning. The Dataset Core approach provides the first mathematically rigorous framework for simultaneous data compression and pollution filtering, with formal guarantees on both information preservation and adversarial robustness, [28], [29].

From a practical perspective, our framework addresses several critical gaps in current

AI security infrastructure. Traditional approaches to data validation rely heavily on statistical outlier detection, which sophisticated adversaries can easily circumvent, [30]. Our information-theoretic approach, combined with economic incentives, creates multiple layers of defense that are significantly more difficult to compromise systematically.

The blockchain integration provides unprecedented transparency and auditability in AI training processes. This capability has profound implications for regulatory compliance, particularly in industries like healthcare and finance where algorithmic decision-making must be fully auditable. The triple-entry accounting framework ensures that every training decision can be traced back to its data sources, enabling forensic analysis of model behavior and bias.

8.2 Limitations and Computational Considerations

Despite these advances, several limitations require careful consideration. The computational overhead introduced by blockchain operations represents a non-trivial cost, particularly for organizations with limited computational resources. While our experiments demonstrate that this overhead scales favorably with dataset size, the absolute computational requirements may prohibit adoption in resource-constrained environments.

The reliance on economic incentives also introduces complexity in parameter tuning and market design. The optimal values for micropayment coefficients (α, β, γ) likely vary across domains and applications, requiring careful calibration and ongoing adjustment. Furthermore, the economic model assumes rational actors, but real-world adversaries may be motivated by factors beyond simple economic gain, potentially limiting the effectiveness of financial disincentives. Scalability concerns persist regarding blockchain transaction throughput. Current blockchain technologies struggle to handle the transaction volumes required for large-scale machine learning applications. While layer-2 solutions show promise, they often involve trade-offs between throughput and security guarantees that must be carefully evaluated in each deployment context.

8.3 Future Research Paths

Several promising research directions emerge from this work. Adaptive core construction algorithms that dynamically respond to detected pollution patterns could further improve robustness while minimizing information loss. Such systems would continuously monitor data quality metrics and adjust core construction parameters in real-time to maintain optimal performance under varying threat conditions.

The integration of differential privacy techniques with our Dataset Core approach represents another compelling research avenue. Formal privacy guarantees could complement our existing security properties, making the framework suitable for applications involving sensitive personal data while maintaining the ability to detect and filter malicious contributions.

Game-theoretic analysis of the micropayment ecosystem could provide deeper insights into optimal mechanism design for data marketplaces. Understanding how different payment structures influence contributor behavior and data quality could lead to more effective economic incentive schemes and better overall system performance.

Finally, the extension of our framework to federated learning environments presents both opportunities and challenges. The decentralized nature of federated learning aligns well with blockchain-based verification, but the distributed data scenario requires careful adaptation of core construction algorithms and consensus mechanisms.

9 Conclusions

This research establishes a comprehensive framework for addressing the critical challenges of data pollution and AI model deterioration through the integration of information-theoretic, cryptographic, and economic approaches. The Dataset Core methodology provides mathematically grounded protection against adversarial data contamination while preserving essential information content. The blockchain-based verification system ensures transparency, auditability, and tamper-resistance throughout the machine learning pipeline.

Our experimental validation demonstrates that this integrated approach can significantly improve adversarial robustness while maintaining competitive computational performance. The 73% reduction in attack effectiveness, combined with 95% information preservation at 80% compression ratios, establishes the practical viability of our approach for real-world applications.

The economic incentive mechanisms introduce a paradigm shift toward market-driven data quality assurance, creating sustainable ecosystems where high-quality data contributions are rewarded and malicious behavior is disincentivized. This approach addresses fundamental limitations of centralized data curation while providing scalable solutions for large-scale AI applications.

As artificial intelligence systems become increasingly central to critical infrastructure and decision-making processes, frameworks like the one presented here will be essential for maintaining public trust and ensuring reliable performance. The theoretical foundations established in this work, as derived from Equations (1)–(9), provide a solid basis for future research and development in adversarial-resistant AI systems, while the practical implementation guidelines enable immediate deployment in security-sensitive applications.

The convergence of information theory, blockchain technology, and economic incentives represents a promising direction for building trustworthy AI systems that maintain integrity in adversarial environments. While implementation challenges remain, the substantial security improvements demonstrated in our experiments justify the computational and economic costs involved, particularly for applications where data integrity and model reliability are paramount.

Acknowledgment:
The authors would like to thank Ian Grigg and George Papageorgiou for their review and suggestions to early drafts of this manuscript.

References:

[1] D. H. Hagos, R. Battle, and D. B. Rawat, "Recent Advances in Generative AI and Large Language Models: Current Status, Challenges, and Perspectives," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 12, pp. 5913–5933, 2024. DOI: https://doi.org/10.1109/TAI.2024.3444742

[2] I. Shumailov, Z. Shumaylov, Y. Zhao, N. Papernot, R. Anderson, and Y. Gal, "AI models collapse when trained on recursively generated data," *Nature*, vol. 631, pp. 755–759, 2024. DOI: https://doi.org/10.1038/s41586-024-07566-y

[3] A. Kurakin, I. Goodfellow, and S. Bengio, "Adversarial Examples in the Physical World," in *Artificial Intelligence Safety and Security*, R. V. Yampolskiy, Ed., Chapman and Hall/CRC, 2018, pp. 99–112. DOI: https://doi.org/10.1201/9781351251389

[4] M. Battiste, "Cognitive Imperialism," in *Encyclopedia of Educational Philosophy and Theory*, M. A. Peters, Ed., Springer, Singapore, 2017, pp. 153–156. DOI: https://doi.org/10.1007/978-981-287-588-4_501

[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. DOI: https://doi.org/10.1038/nature14539

[6] A. N. Kolmogorov, "On tables of random numbers," *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 25, no. 4, pp. 369–376, 1963. Available online: https://www.jstor.org/stable/25049284 (accessed January 10, 2026)

[7] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948. DOI: https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley-Interscience, 2006. DOI: https://doi.org/10.1002/047174882X

[9] D. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, 1952. DOI: https://doi.org/10.1109/JRPROC.1952.273898

[10] K. Sgantzos, S. Stelios, P. Tzavaras, and M. A. Hemairy, "Minds and machines: evaluating the feasibility of constructing an advanced artificial intelligence," *Discover Artificial Intelligence*, vol. 4, article 104, 2024. DOI: https://doi.org/10.1007/s44163-024-00216-2

[11] D. Feldman, "Introduction to Core-Sets: An Updated Survey," *arXiv preprint arXiv:2011.09384*, 2020. arXiv: https://arxiv.org/abs/2011.09384 (accessed January 10, 2026)

[12] M. Ferrara, "Data Poisoning and Artificial Intelligence Modeling: Theoretical Foundations and Defensive Strategies," *CEUR Workshop Proceedings*, vol. 4031, pp. 28–40, 2025. Available online: https://ceur-ws.org/Vol-4031/ (accessed January 10, 2026)

[13] I. Grigg, "Triple Entry Accounting," Systemics Inc., Working Paper, 2005. Available online: https://iang.org/papers/triple_entry.html (accessed January 10, 2026)

[14] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," White Paper, 2008. Available online: https://bitcoin.org/bitcoin.pdf (accessed January 10, 2026)

[15] D. Drescher, *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, Apress, 2017. DOI: https://doi.org/10.1007/978-1-4842-2604-9

[16] T. Gu, B. Dolan-Gavitt, and S. Garg, "BadNets: Identifying Vulnerabilities in the Machine Learning Model Supply Chain," *arXiv preprint arXiv:1708.06733*, 2017. arXiv: https://arxiv.org/abs/1708.06733 (accessed January 10, 2026)

[17] A. Shafahi, W. R. Huang, M. Najibi, O. Suciu, C. Studer, T. Dumitras, and T. Goldstein, "Poison Frogs! Targeted Clean-Label Poisoning Attacks on Neural Networks," in *Advances in Neural Information Processing Systems*, vol. 31, 2018. Available online: https://doi.org/10.48550/arXiv.1804.00792(accessed January 10, 2026)

[18] G. Wood, "Ethereum: A Secure Decentralised Generalised Transaction Ledger," Ethereum Project Yellow Paper, 2014. Available online: https://ethereum.github.io/yellowpaper/paper.pdf (accessed January 10, 2026)

[19] V. Buterin, "A Next-Generation Smart Contract and Decentralized Application Platform," Ethereum White Paper, 2014. Available online: https://ethereum.org/en/whitepaper/ (accessed January 10, 2026)

[20] K. Sgantzos and G. Ward, "Proof of Work in Governance and Politics," in *Blockchain Technology: Advances in Research and Applications*, Nova Publishing, 2022. DOI: https://doi.org/10.52305/RTZT8988

[21] B. Biggio and F. Roli, "Wild Patterns: Ten Years After the Rise of Adversarial Machine Learning," *Pattern Recognition*, vol. 84, pp. 317–331, 2018. DOI: https://doi.org/10.1016/j.patcog.2018.07.023

[22] X. Chen, C. Liu, B. Li, K. Lu, and D. Song, "Targeted Backdoor Attacks on Deep Learning Systems Using Data Poisoning," *arXiv preprint arXiv:1712.05526*, 2017. arXiv: https://arxiv.org/abs/1712.05526 (accessed January 10, 2026)

[23] M. Jagielski, A. Oprea, B. Biggio, C. Liu, C. Nita-Rotaru, and B. Li, "Manipulating Machine Learning: Poisoning Attacks and Countermeasures for Regression Learning," in *IEEE Symposium on Security and Privacy (SP)*, pp. 19–35, 2018. DOI: https://doi.org/10.1109/SP.2018.00057

[24] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing Properties of Neural Networks," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2014. arXiv: https://arxiv.org/abs/1312.6199 (accessed January 10, 2026)

[25] B. Tran, J. Li, and A. Madry, "Spectral Signatures in Backdoor Attacks," in *Advances in Neural Information Processing Systems*, vol. 31, 2018. Available online: https://proceedings.neurips.cc/paper/2018/hash/280cf18baf4311c92aa5a042336587d3-Abstract.html (accessed January 10, 2026)

[26] L. Huang, A. D. Joseph, B. Nelson, B. I. P. Rubinstein, and J. D. Tygar, "Adversarial Machine Learning," in *Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence*, pp. 43–58, 2011. DOI: https://doi.org/10.1145/2046684.2046692

[27] J. Steinhardt, P. W. Koh, and P. S. Liang, "Certified Defenses for Data Poisoning Attacks," in *Advances in Neural Information Processing Systems*, vol. 30, 2017. Available online: https://proceedings.neurips.cc/paper/2017/hash/9d7311ba459f9e45ed746755a32dcd11-Abstract.html (accessed January 10, 2026)

[28] M. Barreno, B. Nelson, A. D. Joseph, and J. D. Tygar, "The Security of Machine Learning," *Machine Learning*, vol. 81, no. 2, pp. 121–148, 2010. DOI: https://doi.org/10.1007/s10994-010-5188-5

[29] N. Papernot, P. McDaniel, A. Sinha, and M. P. Wellman, "SoK: Security and Privacy in Machine Learning," in *IEEE European Symposium on Security and Privacy (EuroS&P)*, pp. 399–414, 2018. DOI: https://doi.org/10.1109/EuroSP.2018.00035

[30] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and Harnessing Adversarial Examples," *arXiv preprint arXiv:1412.6572*, 2014. arXiv: https://arxiv.org/abs/1412.6572 (accessed January 10, 2026)

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflicts of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.