Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 8, 2025

TRUSTWORTHY AI FOR IOT USING BLOCKCHAIN

Mehak Riaz khan[*1], Nazia Abrar[2], Muhammad Rizwan Tahir[3], Yusra Mansoor[4], Engr. Sidra Rehman[5]

[*1,5]*Senior Lecturer, Department Computer Science, Iqra University North Campus, Karachi, Pakistan.*
[2]*Lecturer, College of computing and information sciences, Karachi Institute of Economics & Technology, Karachi, Pakistan*
[3]*Affiliation: Machine Learning Engineer, Rootblock Labs, Lahore, Pakistan*
[4]*Senior Lecturer, Department Computer Science, DHA Suffa University, Karachi, Pakistan*

[*1]mehak.riaz@iqra.edu.pk,[2]nazia.abrar@kiet.edu.pk,[3]therizwantahir@gmail.com, [4]yusra.mansoor@dsu.edu.pk,[5]sidra.rehman_n@iqra.edu.pk

DOI: https://doi.org/10.5281/zenodo.16958819

Keywords

Article History
Received on 25 May 2025
Accepted on 07 August 2025
Published on 27 August 2025

Corresponding Author: *
Mehak Riaz khan

Abstract

The convergence of Artificial Intelligence (AI) and the Internet of Things (IoT) has created new opportunities for automation, data-driven optimization, and intelligent decision-making. Despite these advances, critical challenges such as security, privacy, and transparency remain unresolved due to the distributed and heterogeneous nature of IoT environments [1]. Blockchain technology, with its decentralized and immutable characteristics, has emerged as a promising enabler of trust in AI-driven IoT systems [3], [4]. By integrating blockchain with AI, it becomes possible to ensure accountability, resilience, and verifiability of IoT data and AI models. This paper explores the role of blockchain in building trustworthy AI for IoT by reviewing current approaches, identifying key trust dimensions, and proposing a reference architecture that supports decentralized learning, privacy preservation, and auditable decision-making. Potential applications across healthcare, smart cities, and autonomous systems are also highlighted, emphasizing the transformative potential of blockchain-enabled AI for IoT ecosystems.

INTRODUCTION

The propagation of IoT devices has transformed modern computing, enabling billions of connected sensors, actuators, and smart systems to interact with the physical and digital worlds. IoT has penetrated diverse domains, including healthcare, transportation, manufacturing, and smart cities, generating massive volumes of heterogeneous data [1]. Extracting meaningful insights from such data streams relies heavily on AI techniques, especially machine learning (ML) and deep learning (DL). However, deploying AI in IoT raises fundamental trust-related questions, including data reliability, model integrity, transparency of decision-making, and protection against adversarial threats [2].

Blockchain, introduced as the foundation of Bitcoin [3], has evolved into a versatile technology for establishing reorganized trust across untrusted environments. Its characteristics, such as immutability, transparency, and consensus-driven validation, can mitigate many of the challenges inherent in AI-enabled IoT ecosystems [4]. By combining AI, IoT, and blockchain, researchers aim to build systems that are not only intelligent but also secure, auditable, and resilient.

This paper investigates the role of blockchain in achieving trustworthy AI for IoT. The main contributions are:

1. A comprehensive review of current approaches integrating blockchain with AI for IoT.

2. Identification of trustworthiness dimensions including security, privacy, transparency, fairness, and robustness.
3. A proposed reference architecture leveraging blockchain-enabled federated learning for trustworthy AI in IoT.
4. A discussion of potential applications and open research challenges.

Background

Internet of Things and Trust Issues

The Internet of Things (IoT) is a rapidly growing paradigm that connects billions of heterogeneous devices, ranging from sensors and actuators to smartphones and industrial controllers. These devices continuously collect, process, and transmit data, creating opportunities for smarter services and autonomous decision-making across domains such as healthcare, transportation, and manufacturing [5]. However, IoT devices typically operate under resource-constrained conditions, with limited processing power, memory, and battery life. This makes them vulnerable to cyberattacks, unauthorized access, and data manipulation. Moreover, the large-scale deployment of IoT systems introduces complexity and heterogeneity, which further complicates security management.

One of the most critical concerns in IoT networks is the lack of trust. Data collected from sensors can be intercepted or altered, leading to inaccurate or even malicious outcomes. Centralized servers, which are often used to aggregate and analyze IoT data, introduce single points of failure and become prime targets for attackers [6]. Trust issues extend beyond security, encompassing aspects such as the reliability of devices, transparency of system operations, and the accountability of decision-making processes. As IoT applications often involve sensitive or mission-critical environments, such as healthcare monitoring and industrial automation, addressing these trust-related challenges is of paramount importance.

Artificial Intelligence in IoT

Artificial Intelligence (AI), particularly machine learning (ML) and deep learning (DL), has become an indispensable tool for IoT data analysis. AI enables predictive analytics, anomaly detection, intelligent control, and personalized services by learning patterns from massive streams of sensor data [7]. In smart cities, for example, AI systems optimize energy distribution and traffic flow, while in healthcare, they provide diagnostic support based on wearable sensor data.

Despite its advantages, integrating AI into IoT introduces several challenges related to trustworthiness. Machine learning models are highly vulnerable to adversarial inputs, where carefully crafted malicious data can mislead models into making incorrect decisions [8]. Similarly, data poisoning attacks compromise training datasets, leading to biased or faulty models. In addition, many deep learning models function as "black boxes," offering little explanation about their internal reasoning processes. This opacity raises concerns about accountability, fairness, and transparency, particularly in domains where decisions can have life-critical consequences, such as medical diagnosis or autonomous driving.

Blockchain as a Trust Enabler

Blockchain technology has occurred as a talented enabler of trust in distributed and resource-constrained environments. First introduced by Nakamoto as the foundation of Bitcoin [3], blockchain provides a decentralized and tamper-evident ledger where transactions are recorded in immutable blocks validated by consensus mechanisms [9]. This design eliminates the reliance on centralized authorities, ensuring that no single entity can manipulate records or compromise system integrity.

For IoT systems, blockchain offers several advantages. Its decentralized nature reduces the risk of single points of failure, while its cryptographic mechanisms ensure data authenticity and immutability. Smart contracts, programmable rules stored on the blockchain, enable automated and transparent enforcement of policies for data sharing, resource allocation, and service agreements [10]. This not only improves trust but also reduces operational overhead in large-scale IoT environments.

When integrated with AI, blockchain adds an additional layer of trustworthiness. For example, blockchain can record data provenance to ensure that training datasets are reliable and untampered, thereby protecting AI models from data poisoning. It

can also track model updates in collaborative training scenarios such as federated learning, providing accountability and ensuring fairness [11]. Additionally, blockchain's immutable records support auditing of AI decision-making processes, which enhances explainability and accountability in sensitive domains. Together, these features position blockchain as a critical foundation for building trustworthy AI-enabled IoT ecosystems.

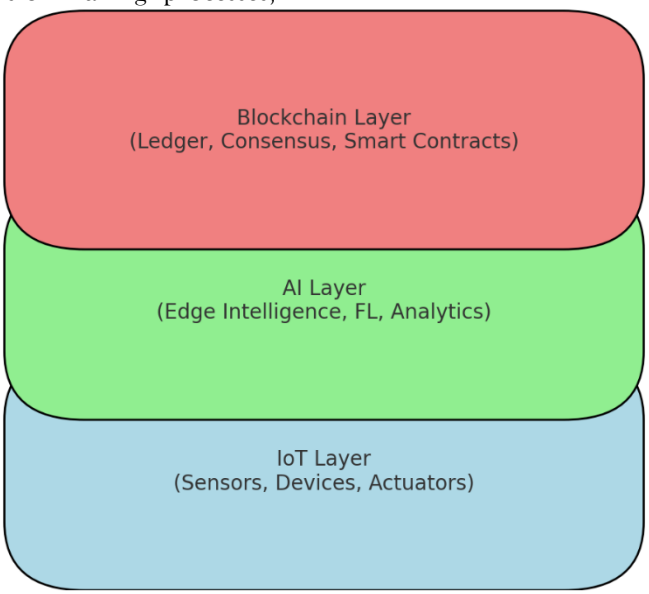

Figure 1: Layered IoT-AI-Blockchain Integration Model

DIMENSIONS AND CHALLENGES

The integration of Artificial Intelligence (AI) into Internet of Things (IoT) ecosystems has created opportunities for automation, smart decision making, and real time analytics. However, the adoption of AI in sensitive and large-scale IoT deployments is limited by concerns over trustworthiness. Trustworthy AI refers to systems that are secure, privacy-preserving, transparent, fair, and reliable, ensuring that users and stakeholders can depend on their outputs in critical environments [12]. These dimensions are crucial, as IoT devices often collect sensitive data and function in safety critical fields such as healthcare, transportation, and industry.

Security and Robustness

AI-enabled IoT systems face significant security challenges. Malicious actors can exploit vulnerabilities in data pipelines or model training processes to launch adversarial attacks, poisoning the learning process or forcing incorrect predictions [12]. For example, adversarial inputs—subtly modified data designed to mislead models—can cause AI systems in autonomous vehicles or medical applications to produce dangerous outcomes [13]. Poisoning attacks on training data can bias models, leading to systematic failures or manipulated outputs. Ensuring robustness requires methods such as adversarial training, anomaly detection, and secure logging of model updates. Blockchain can enhance security by providing immutable records of training contributions and model provenance, thereby protecting against tampering and unauthorized modifications.

Privacy Preservation

Privacy is another critical requirement for trustworthy AI in IoT. Devices such as wearables, smart home systems, and healthcare monitors gather sensitive data, including personal health metrics, behavioral patterns, and location information. Centralized AI training approaches, which rely on aggregating raw data in cloud servers, expose users to privacy risks. Emerging methods such as disparity privacy [14] and amalgamated learning [15] reduce these risks by ensuring that raw data never leaves

local devices. Instead, only model updates are shared. Blockchain further strengthens privacy preservation by securely recording these updates on a distributed ledger, preventing manipulation and reducing reliance on centralized entities.

Transparency and Explainability

The opaque, "black-box" nature of deep learning models is a barrier to trust, especially in critical IoT applications. Users and stakeholders often demand clear explanations for AI-driven decisions, particularly in sensitive areas like healthcare or finance [16]. Explainable AI (XAI) techniques attempt to make model predictions interpretable and understandable to humans. Blockchain complements these efforts by creating immutable logs of data flows, model updates, and decision-making processes, which can later be audited. This combination improves accountability and ensures that IoT applications provide transparent reasoning behind AI outputs [17].

Fairness and Accountability

Bias in AI systems is an increasingly recognized challenge, particularly when models are trained on unrepresentative or skewed datasets. In IoT applications, this could lead to unfair outcomes, such as biased health diagnostics or unequal distribution of smart city resources [18]. Accountability is equally important, as stakeholders must know who is responsible when AI-enabled systems fail or cause harm [19]. Blockchain offers mechanisms to ensure fairness and accountability by recording data provenance, tracking contributions to collaborative learning, and maintaining transparent records of decisions. These features help ensure that outcomes are equitable and that responsibility can be assigned when necessary.

Reliability and Safety

IoT systems often operate in dynamic and uncertain environments, where reliability and safety are non-negotiable. AI models must provide consistent and accurate outputs, even under fluctuating conditions or unforeseen challenges [20]. For instance, industrial IoT systems controlling critical infrastructure or autonomous vehicles navigating urban environments require AI that is fault-tolerant and resilient. Failures in reliability could lead to safety hazards, financial losses, or loss of human life. Trustworthy AI frameworks must therefore incorporate monitoring mechanisms, resilience strategies, and fallback options to ensure safe operations [21].

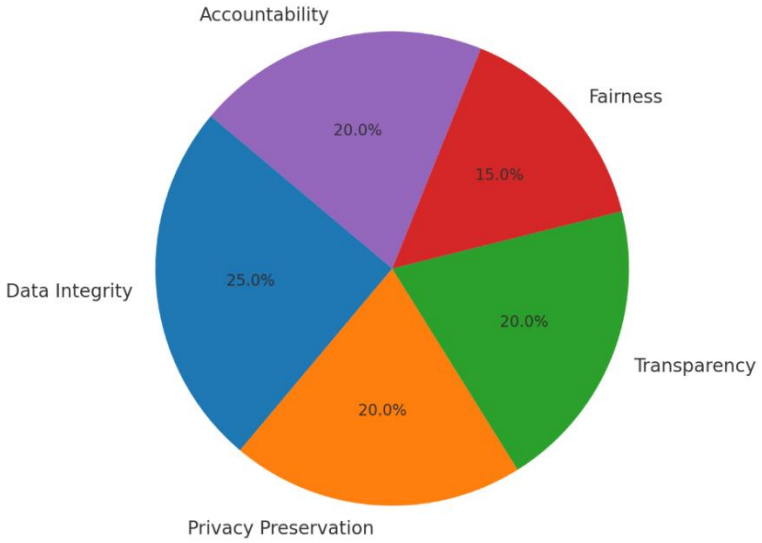

Figure 2: Blockchain Contributions to Dimentions of Trustworthy A

Table 1: Trustworthiness Dimensions and Blockchain Support

Dimension	Challenge in IoT AI	Blockchain Contribution
Security	Vulnerable to adversarial/data	Immutable ledger, consensus

	attacks	validation
Privacy	Raw data exposure in central training	Federated learning, on-chain validation
Transparency	Opaque "black-box" decisions	Immutable audit trails for accountability
Fairness	Bias from skewed datasets	Decentralized participation, on-chain incentives
Reliability	Single-point failures, weak fault tolerance	Decentralized consensus, redundancy

Summary of Challenges

In conclusion, trustworthy AI in IoT must address security, privacy, transparency, fairness, and reliability simultaneously. These challenges are deeply interconnected and cannot be resolved by traditional centralized AI pipelines, which depend heavily on untrusted intermediaries and lack accountability. Blockchain's decentralized, immutable, and auditable nature offers a promising foundation for addressing these gaps. By combining blockchain with privacy-preserving AI techniques such as federated learning, it is possible to create IoT ecosystems where intelligent decision-making is not only efficient but also trustworthy [22].

Blockchain For Trustworthy Ai In Iot

Blockchain has emerged as a critical technology to enhance trust, transparency, and security in AI-enabled IoT systems. It's decentralized, immutable, and auditable properties make it particularly suitable for addressing vulnerabilities and limitations of traditional centralized approaches. In this section, we highlight how blockchain contributes to trustworthy AI in IoT through secure data sharing, federated learning support, decentralized AI marketplaces, and improved explainability.

Secure Data Sharing

IoT systems generate enormous amounts of heterogeneous data distributed across multiple devices, organizations, and geographic regions. Sharing this data securely and reliably is essential for developing accurate AI models and intelligent services. However, conventional centralized data-sharing architectures face risks of tampering, unauthorized access, and single points of failure [17]. Blockchain addresses these concerns by offering a reorganized ledger where data sharing trades are recorded immutably and validated through consensus. This ensures that data cannot be altered once committed, protecting integrity and authenticity. Furthermore, blockchain-based smart contracts allow organizations to define and enforce transparent data-sharing policies automatically. For example, healthcare IoT networks can use smart contracts to regulate access to patient data, ensuring compliance with privacy rules while maintaining transparency [18]. By combining cryptographic guarantees with decentralized governance, blockchain creates a trusted environment for secure IoT data exchange.

Federated Learning with Blockchain

Federated learning (FL) has gained attention as a privacy-preserving approach to AI model training. Instead of sharing raw data, IoT devices collaboratively train models by exchanging local model updates [19]. While FL reduces privacy risks, it remains vulnerable to model poisoning attacks, dishonest participants, and lack of accountability.

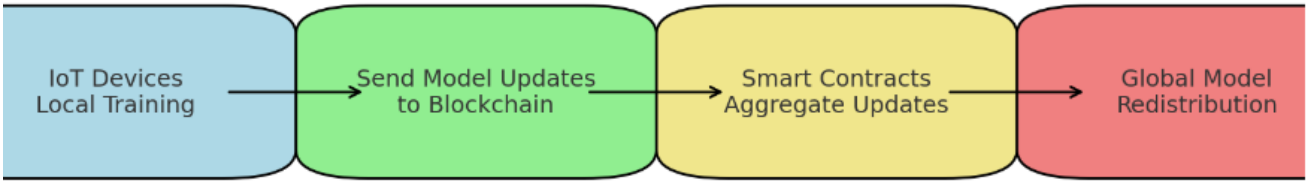

Figure 3: Workflow of Blockchain-Enabled Federated Learning

Blockchain enhances FL by recording model updates immutably on the distributed ledger. This prevents malicious participants from injecting false updates and ensures that the training process remains auditable [20]. Additionally, smart contracts can implement incentive mechanisms to reward devices that contribute high-quality model updates and penalize those providing harmful or low-quality contributions. Token-based reward systems managed on-chain encourage honest participation, making FL more robust and scalable in IoT contexts. Thus, blockchain-enabled federated learning provides both privacy preservation and decentralized accountability, addressing core challenges of trustworthy AI.

Decentralized AI Marketplaces

AI-driven IoT applications often require trained models to be shared across organizations and service providers. Traditional AI marketplaces are usually centralized, raising concerns about bias, monopolistic control, and lack of transparency. Blockchain supports the creation of decentralized AI marketplaces, where stakeholders can securely share, exchange, and monetize AI models without depending on intermediaries [21].

In such marketplaces, blockchain ensures the provenance and authenticity of shared models, while smart contracts govern licensing agreements, payments, and usage rights. Reputation systems built on blockchain further enhance trust by tracking the reliability of model providers and consumers. For IoT ecosystems, decentralized AI marketplaces open opportunities for collaboration while reducing the risks of data misuse or unfair practices.

Explainability and Auditability

Explainability and accountability are crucial dimensions of trustworthy AI, particularly in IoT applications such as healthcare, smart cities, and autonomous transportation. Blockchain enhances these dimensions by providing immutable audit trails of data and decision-making processes [22]. Every data entry, model update, and inference can be recorded on-chain, enabling stakeholders to trace back the origins of AI-driven outcomes.

This auditability is essential for sensitive applications where accountability matters. For example, in an autonomous vehicle system, blockchain can log every sensor input and AI-driven decision, enabling investigators to determine responsibility in the case of accidents [23]. Similarly, in healthcare IoT systems, blockchain can ensure that clinical decisions supported by AI are based on authentic, untampered data. By integrating blockchain with explainable AI methods, systems can achieve both interpretability and accountability, thereby strengthening user trust in IoT-based intelligence.

Table 2: Comparison of Traditional AI and Blockchain-Enabled Trustworthy AI in IoT

Dimension	Traditional AI in IoT	Blockchain-Enabled AI in IoT
Security	Vulnerable to attacks, centralized failures	Immutable ledger, consensus-based validation
Privacy	Requires data centralization	Supports federated learning and differential privacy
Transparency	Opaque 'black-box' models	Immutable audit trails for explainability
Fairness	Risk of biased training datasets	On-chain accountability and

		reputation systems
Reliability	Single point of failure risk	Decentralized, fault-tolerant

Summary

In summary, blockchain serves as a powerful enabler for trustworthy AI in IoT by securing data exchange, enhancing federated learning, enabling decentralized AI marketplaces, and providing immutable audit trails. These capabilities directly address the challenges of privacy, transparency, and accountability that conventional AI and IoT frameworks struggle to solve. The synergy between blockchain, AI, and IoT thus establishes a foundation for building intelligent systems that are not only powerful but also trustworthy [24].

Proposed Architecture

Building trustworthy AI for IoT requires an architecture that ensures decentralization, security, privacy, and auditability while supporting large-scale data processing. To address these requirements, we propose a blockchain-enabled architecture that integrates federated learning, edge computing, and smart contracts. The architecture is designed to mitigate trust issues in AI-enabled IoT environments by decentralizing model training, preserving privacy, and providing immutable audit trails for decision-making [19], [23].

Key Components

1. IoT Devices

IoT devices serve as the data sources in the proposed framework. These devices include sensors, actuators, wearables, and smart home appliances that continuously generate raw data. Given their resource constraints, IoT devices perform lightweight preprocessing such as data cleaning, normalization, or feature extraction before transmitting updates to edge servers [5], [6].

Edge Servers

Edge servers act as intermediaries between IoT devices and the blockchain network. They receive local model updates from IoT devices and perform partial training or aggregation of data. By offloading computation from resource-limited devices, edge servers reduce latency and improve scalability. Moreover, they interact with blockchain nodes to record model updates and ensure data provenance [24].

Blockchain Network

The blockchain network functions as the trust backbone of the architecture. It stores immutable records of model updates, training contributions, and access control rules. Consensus mechanisms, such as Proof of Stake (PoS) or Practical Byzantine Fault Tolerance (PBFT), ensure that all updates are validated and agreed upon in a decentralized manner [9], [26]. This prevents malicious participants from tampering with training results or injecting poisoned data.

Federated Learning Coordinator

The federated learning (FL) coordinator aggregates local model updates from IoT devices and edge servers to produce a global model. Unlike traditional FL frameworks that rely on centralized servers, our architecture records all updates and aggregation steps on the blockchain. This ensures accountability, prevents model poisoning, and provides transparency regarding the training process [18], [19].

Smart Contracts

Smart contracts are integral to the proposed architecture, as they define the rules governing participation, resource sharing, and incentives. For example, participants who contribute accurate model updates can be rewarded with tokens, while malicious or low-quality contributions can be penalized [10], [20]. Smart contracts also enforce data access permissions, ensuring compliance with privacy regulations such as GDPR [14].

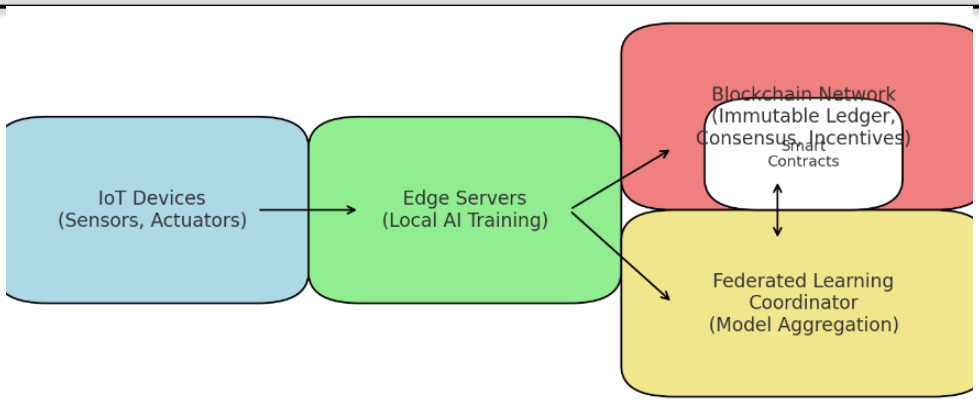

Figure 4: Proposed Blockchain-Enabled Trustworthy AI Architecture for IoT

B. Workflow of the Architecture

The workflow of the proposed system proceeds as follows:

1. IoT devices locally train models using their private datasets and send encrypted updates to edge servers.
2. Edge servers validate updates and forward them to the blockchain network for consensus.
3. The blockchain immutably records model updates and triggers smart contracts to manage incentives and enforce policies.
4. The federated learning coordinator aggregates validated updates to generate a global model.
5. The updated global model is redistributed to IoT devices for further training cycles.

This iterative process enables collaborative AI training across distributed IoT devices without exposing sensitive raw data. Blockchain ensures that each step of the process is transparent, auditable, and resistant to tampering [21], [25].

Advantages of the Proposed Architecture

- **Privacy Preservation:** Sensitive user data remains local to IoT devices, while only encrypted model updates are shared [15].
- **Decentralized Trust:** The blockchain eliminates the need for a centralized authority, ensuring fair participation and resilience against single points of failure [3], [9].
- **Transparency and Accountability:** Immutable logs of all training contributions enable auditing of decisions and model evolution [11], [17].
- **Scalability:** By leveraging edge servers and lightweight consensus mechanisms, the architecture supports large-scale IoT deployments [26], [27].
- **Incentive Alignment:** Smart contracts ensure that honest participants are rewarded, encouraging sustainable collaboration [20].

Limitations and Considerations

While the architecture provides strong guarantees of trustworthiness, certain limitations remain. Blockchain consensus mechanisms may introduce latency, which can hinder real-time IoT applications [28]. Additionally, resource-constrained IoT devices may still struggle with the computational overhead of encryption or local training. Future work must focus on developing lightweight blockchain protocols and adaptive consensus mechanisms to better suit IoT environments [27], [29].

APPLICATIONS

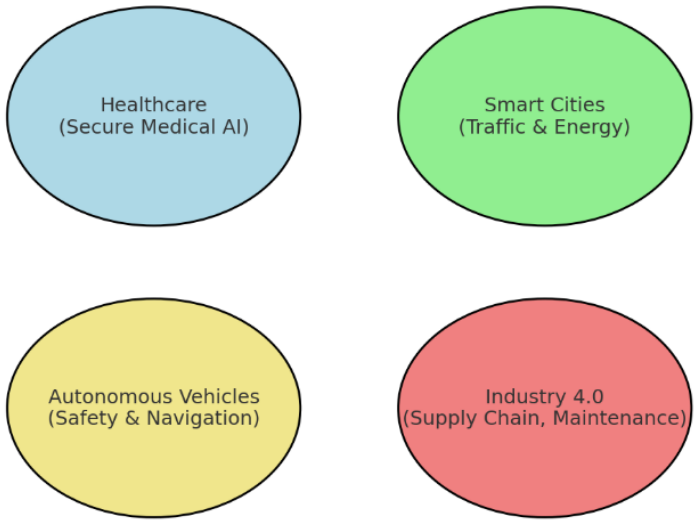

Figure 5: Applications of Blockchain-Enabled Trustworthy AI in IoT

Smart Healthcare

Healthcare is one of the most sensitive domains where IoT devices collect real-time physiological data from patients through wearable sensors, smart implants, and monitoring systems. AI plays a crucial role in analyzing this data for diagnostic support, early disease detection, and personalized treatment recommendations [22]. However, privacy and trust issues are paramount since healthcare data is highly sensitive. Blockchain can ensure secure sharing of patient data among hospitals, insurers, and research organizations, while smart contracts define clear policies for access and usage. Combined with AI, this enables trustworthy clinical decision support systems that are auditable, privacy-preserving, and resistant to manipulation. For example, immutable blockchain records allow healthcare professionals to verify the origin of data used in AI-driven diagnoses, enhancing accountability [23].

Smart Cities

Smart city infrastructures depend heavily on IoT devices for monitoring and managing services such as traffic systems, waste management, and energy grids. AI techniques process vast amounts of sensor data to optimize resource allocation, improve urban mobility, and enhance sustainability [24]. Nevertheless, the accuracy and trustworthiness of such AI-driven optimizations are critical, as any manipulation of underlying IoT data could disrupt essential services. Blockchain ensures the authenticity and integrity of IoT-collected data, preventing malicious tampering. Furthermore, blockchain-enabled AI systems provide transparency in governance, enabling city authorities and citizens to audit decisions. For example, blockchain-based traffic management platforms can make AI-driven routing decisions both accountable and verifiable, which is crucial for public trust in urban automation [25].

Autonomous Systems

Autonomous systems, particularly self-driving cars and unmanned aerial vehicles, rely extensively on IoT data and AI algorithms for perception, navigation, and control. Ensuring trust in these AI-driven systems is vital for safety and user acceptance [24]. AI models can be vulnerable to adversarial attacks, such as manipulated road signs or spoofed GPS signals, which could lead to accidents. Blockchain can mitigate these risks by recording the provenance of IoT sensor data, thereby allowing vehicles to verify the authenticity of the information used for decision-making. Additionally, blockchain's

decentralized consensus mechanisms ensure that no single malicious entity can compromise decision processes across fleets of autonomous systems. In this way, combining blockchain with AI enhances both safety and accountability in autonomous environments [26].

Supply Chain and Industry 4.0

Supply chains increasingly leverage IoT devices for real-time tracking of goods, monitoring of storage conditions, and predictive analytics for logistics. AI provides predictive insights into demand fluctuations, quality assurance, and anomaly detection across global supply networks [25]. However, the integrity and trustworthiness of IoT data in supply chains are often questioned, especially when multiple stakeholders with conflicting interests are involved. Blockchain addresses this challenge by providing a shared, tamper-proof ledger of all transactions and data exchanges. Smart contracts automate verification processes, ensuring transparency and reducing disputes among parties. When combined with AI, blockchain-enabled supply chains can not only predict disruptions but also provide trustworthy and auditable justifications for decision-making. This approach is particularly valuable in Industry 4.0 settings, where trust and automation are fundamental for efficiency and resilience [27].

Table 3: Applications of Blockchain-Enabled Trustworthy AI in IoT

Domain	Role of AI	Role of Blockchain	Example Benefit
Healthcare	Predictive diagnostics	Secure data sharing	Reliable medical AI systems
Smart Cities	Traffic optimization	Auditability of sensor data	Reduced congestion & fraud prevention
Autonomous Vehicles	Navigation & decision-making	Provenance of sensor inputs	Safety and accountability
Industry 4.0	Predictive maintenance	Transparent supply chain	Improved resilience

CHALLENGES AND OPEN RESEARCH ISSUES

Although blockchain has demonstrated significant potential in enabling trustworthy AI for IoT, several challenges remain unresolved. These challenges arise from the inherent limitations of blockchain technology, the unique constraints of IoT devices, and the evolving nature of AI systems. Addressing these issues is essential to make blockchain-enabled AI for IoT scalable, practical, and widely deployable.

Table 4: Open Research Issues in Blockchain-Enabled Trustworthy AI

Challenge	Description	Future Research Direction
Scalability	Consensus overhead limits throughput	Lightweight blockchain protocols
Resource Limits	IoT nodes lack computation/storage	Offloading + edge computing
Interoperability	Heterogeneous IoT platforms	Standardization, cross-chain frameworks
Legal Issues	Unclear liability for AI errors	Governance & regulatory frameworks
Ethics	Bias, fairness, explainability gaps	Explainable and ethical AI + blockchain

Scalability

One of the foremost challenges in integrating blockchain with IoT is scalability. Consensus mechanisms such as Proof of Work (PoW) or Proof of Stake (PoS) often require significant computational resources and time to validate transactions [26]. While these methods are effective in small-scale blockchain networks, they become bottlenecks when applied to IoT environments with billions of devices generating continuous data streams. The transaction throughput and latency of conventional blockchains are insufficient for real-time IoT applications, such as autonomous driving or industrial control systems, where delays of even a few milliseconds can be critical. Future research must explore lightweight and scalable consensus algorithms, such as Proof of Authority or Directed Acyclic Graph (DAG)-based ledgers that are better suited for high-volume IoT scenarios [27].

Resource Constraints

IoT devices typically operate under limited energy, memory, and computational capacity. Running blockchain protocols, which involve cryptographic hashing, digital signatures, and participation in consensus, imposes significant overhead that may not be feasible for low-power devices [27]. For example, lightweight sensors in a smart agriculture system may lack the capacity to participate directly in blockchain operations. This creates the need for edge or fog computing architectures that can offload blockchain tasks from resource-constrained devices, while still ensuring data integrity and trust [28].

Interoperability

Another key challenge lies in interoperability between heterogeneous IoT platforms and blockchain frameworks. IoT ecosystems often consist of devices from multiple vendors operating under different communication standards, data formats, and security protocols [28]. Similarly, blockchain systems are fragmented, with varying consensus mechanisms, privacy features, and governance models. This lack of interoperability creates barriers to seamless integration, hindering the realization of unified, cross-domain trustworthy AI frameworks. Research into standardized protocols, cross-chain communication, and interoperability frameworks will be essential to overcome these limitations [29].

Legal and Ethical Considerations

The deployment of blockchain-enabled AI in IoT introduces complex legal and ethical questions. In decentralized systems, determining liability and accountability when AI-driven decisions cause harm is not straightforward [29]. For example, in the case of an autonomous vehicle accident, identifying whether responsibility lies with the AI model developer, the blockchain network, or the IoT device manufacturer becomes difficult. Moreover, ethical concerns related to bias, fairness, and surveillance arise when AI models are trained on sensitive IoT data. These issues call for the development of governance frameworks, regulatory standards, and ethical guidelines tailored to decentralize AI-driven IoT systems [30].

Standardization and Adoption

The absence of universally accepted standards for blockchain-enabled AI in IoT is a major obstacle to large-scale adoption. Existing solutions are often developed in isolation, leading to fragmented ecosystems and poor interoperability. Standardization across security protocols, consensus mechanisms, and data-sharing models will be necessary to ensure compatibility and foster adoption across industries [30]. In addition, organizational resistance, cost of deployment, and lack of technical expertise also slow down adoption. Collaborative efforts involving academia, industry, and governments will play a critical role in promoting standardized and scalable solutions.

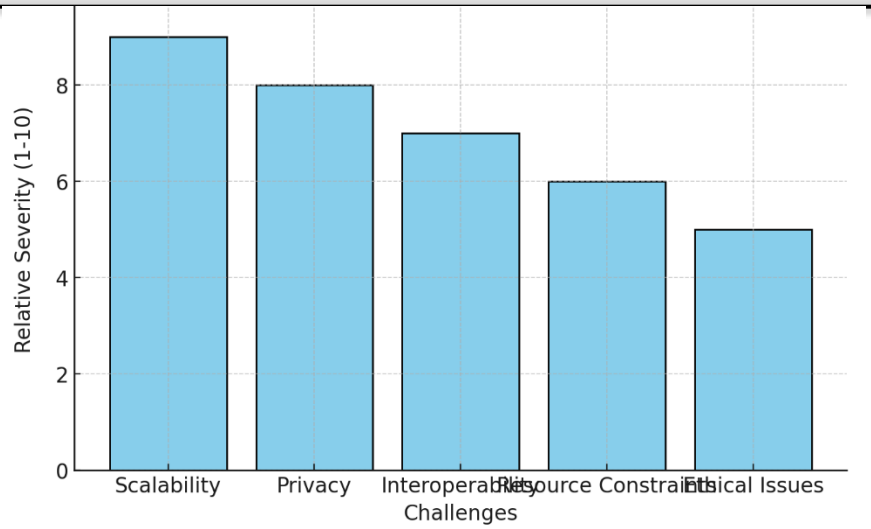

Figure 6: Key Challenges in Achieving Trustworthy AI in IoT

F. Summary of Research Directions

In summary, the key open research issues in blockchain-enabled trustworthy AI for IoT include scalability, resource constraints, interoperability, legal and ethical challenges, and standardization. Overcoming these challenges requires multidisciplinary approaches that combine advances in lightweight cryptography, edge computing, cross-chain interoperability, and regulatory frameworks. Addressing these gaps will enable the design of IoT ecosystems that are not only intelligent but also secure, transparent, and trustworthy [26]–[30].

CONCLUSION

The integration of Artificial Intelligence (AI) into the Internet of Things (IoT) has unlocked unprecedented opportunities for automation, predictive analytics, and intelligent decision-making. However, for such systems to gain widespread adoption in critical domains like healthcare, smart cities, and autonomous transportation, ensuring trustworthiness remains a fundamental challenge. Issues of security, privacy, transparency, fairness, and reliability must be systematically addressed before IoT ecosystems can fully benefit from AI-driven intelligence [12], [16], and [18].

Blockchain technology has emerged as a promising foundation for achieving trustworthy AI in IoT environments. Its decentralized architecture eliminates reliance on central authorities, reducing risks of single points of failure. The immutability and transparency of blockchain records provide strong assurances of data integrity and accountability, while smart contracts enable automated enforcement of policies for secure data sharing and collaborative model training [10], [17]. When combined with privacy-preserving approaches such as federated learning, blockchain ensures that sensitive user data remains secure, while still supporting the collaborative development of robust AI models [15], [19].

The proposed integration of blockchain-enabled federated learning with IoT demonstrates a pathway toward decentralized and verifiable AI ecosystems. This architecture supports auditability, incentivizes honest participation, and ensures that decision-making processes remain transparent and accountable. Such systems are particularly valuable in mission-critical applications, where trust and safety are paramount [20], [21].

Despite these advancements, several open research issues remain. Current blockchain consensus mechanisms face scalability challenges when extended to billions of IoT devices, and lightweight alternatives must be developed to suit resource-constrained environments [26], [27]. Interoperability between heterogeneous IoT platforms and regulatory frameworks also needs to be addressed to support cross-domain integration [28], [29]. Furthermore, ethical considerations such as responsibility for AI

errors and bias mitigation must be incorporated into system design to build public trust [18], [30].

In conclusion, trustworthy AI for IoT using blockchain represents a significant step toward building intelligent, secure, and ethical digital ecosystems. By ensuring privacy, accountability, and resilience, blockchain-enabled AI systems can transform industries ranging from healthcare to transportation. Future research should focus on developing scalable blockchain protocols, establishing interoperability standards, and designing ethical governance frameworks to realize the full potential of trustworthy AI in IoT [22], [29].

REFERENCES

[1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," Computer Networks, vol. 54, no. 15, pp. 2787–2805, 2010.
[2] F. Samie, L. Bauer, and J. Henkel, "IoT technologies for embedded computing: A survey," ACM Trans. Embedded Comput. Syst., vol. 16, no. 3, 2017.
[3] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.
[4] M. Crosby, et al., "Blockchain technology: Beyond bitcoin," Applied Innovation Review, no. 2, 2016.
[5] H. Lin and N. Bergmann, "IoT privacy and security challenges for smart home environments," Information, vol. 7, no. 3, 2016.
[6] A. Al-Fuqaha, et al., "Internet of things: A survey on enabling technologies, protocols, and applications," IEEE Commun. Surveys Tuts., vol. 17, no. 4, 2015.
[7] Y. Sun, H. Song, and A. J. Jara, "Internet of Things and Big Data Analytics for Smart and Connected Communities," IEEE Access, vol. 4, 2016.
[8] N. Papernot, et al., "The limitations of deep learning in adversarial settings," IEEE European Symposium on Security and Privacy, 2016.
[9] X. Xu, et al., "A taxonomy of blockchain-based systems," IEEE Blockchain, 2017.
[10] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the IoT," IEEE Access, vol. 4, 2016.
[11] Z. Zheng, et al., "Blockchain challenges and opportunities: A survey," Int. J. Web and Grid Services, 2018.
[12] I. Goodfellow, et al., "Explaining and harnessing adversarial examples," ICLR, 2015.
[13] C. Dwork, "Differential privacy," ICALP, 2006.
[14] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," arXiv:1702.08608, 2017.
[15] S. Barocas, M. Hardt, and A. Narayanan, Fairness and Machine Learning, 2019.
[16] J. Pearl and D. Mackenzie, The Book of Why, Basic Books, 2018.
[17] S. Underwood, "Blockchain beyond bitcoin," Commun. ACM, vol. 59, no. 11, 2016.
[18] B. McMahan, et al., "Communication-efficient learning of deep networks from decentralized data," AISTATS, 2017.
[19] T. Kim, et al., "Blockchain-based federated learning for trustworthy AI," IEEE Access, vol. 8, 2020.
[20] M. Chen, et al., "AI and blockchain for decentralized marketplaces," IEEE Internet of Things Journal, 2020.
[21] L. Yang, et al., "Blockchain for secure and transparent medical data sharing," IEEE Trans. Engineering Management, 2019.
[22] R. Radanovic and I. Faltings, "Trusted data sharing and trading using blockchain," Proceedings of the 26th IJCAI, 2017.
[23] A. Zanella, et al., "Internet of things for smart cities," IEEE Internet of Things Journal, vol. 1, no. 1, 2014.
[24] Y. Lu, "Blockchain and AI for autonomous vehicle security," IEEE Trans. Industrial Informatics, 2019.
[25] T. M. Fernández-Caramés and P. Fraga-Lamas, "Blockchain in Industry 4.0," Sensors, vol. 18, no. 8, 2018.
[26] E. Androulaki, et al., "Hyperledger Fabric: A distributed operating system for permissioned blockchains," EuroSys, 2018.
[27] S. Dorri, et al., "Blockchain for IoT security and privacy: The case study of a smart home," IEEE PerCom Workshops, 2017.

[28] M. Conoscenti, et al., "Blockchain for the Internet of Things: A systematic review," IEEE International Conference on Smart Computing, 2016.
[29] J. K. Lee, "Legal perspectives on blockchain and AI governance," Computer Law & Security Review, 2020.
[30] H. Treiblmaier, "The impact of blockchain on e-commerce: A framework for trust," Electronic Commerce Research and Applications, vol. 29, 2018.
[31] R. Bano, M. A. Baig, M. A. Hayat, S. H. Channar, and O. Ali, "The role of HR in managing robotic process automation (RPA) displacement anxiety among employees," The Critical Review of Social Sciences Studies, vol. 3, no. 3, pp. 1090–1109, Aug. 2025, doi: 10.59075/f4y5dc30.
[32] F. Irfan, R. Zaka, S. Rehman, B. Sattar, S. A. Haider, and M. A. Hayat, "An IoT-Driven Smart Agriculture Framework for Precision Farming, Resource Optimization, and Crop Health Monitoring," ACADEMIA International Journal for Social Sciences, vol. 4, no. 3, pp. 3329–3342, 2025, doi: 10.63056/ACAD.004.03.0615.
[33] M. A. Hayat, S. Ahmed, M. R. Khan, M. Zaka, F. Irfan, and R. Zaka, "Blockchain-Secured IoT Framework for Smart Waste Management in Urban Environments," The Critical Review of Social Sciences Studies, vol. 3, no. 3, pp. 1462–1467, 2025, doi: 10.59075/mcze1x98.
[34] L. Saeed, R. Khan, S. A. Durrani, C. Y. Mehmood, and M. A. Hayat, "HR Beyond the Office: Leveraging AI to Lead Distributed Teams and Cultivate Organizational Culture in the Age of Remote and Hybrid Work," ACADEMIA International Journal for Social Sciences, vol. 4, no. 3, pp. 291–310, 2025, doi: 10.63056/ACAD.004.03.0361.