SSRG International Journal of Electronics and Communication Engineering Volume 13 Issue 2, 40-52, February 2026
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I2P104

Original Article

TADDA-4i: A Scalable and Secure Tangle-Assisted Decentralized Framework for Industrial Analytics in Industry 4.0

Milton Samadder[1], Anup Kumar Barman[2], Shiladitya Munshi[3], Utpal Madhu[4]

[1,2]*Department of Computer Science and Engineering, Central Institute of Technology, Kokrajhar, India.*
[3]*Department of Computer Science and Engineering, Techno India University, Tripura, Agartala, India.*
[4]*Department of Mechanical Engineering, Swami Vivekananda Institute of Science & Technology, Kolkata, India.*

[1]*Corresponding Author : ph19cse1908@cit.ac.in*

Received: 16 November 2025 Revised: 17 December 2025 Accepted: 18 January 2026 Published: 20 February 2026

***Abstract** - Exponentially growing data generated by networked devices in Industry 4.0 environments requires industrial analytics that are secure, scalable, and decentralized. This article proposes TADDA-4i, a new multi-layered architecture based on IOTA's Tangle-Directed Acyclic Graph (DAG)-based Distributed Ledger Technology (DLT)-combined with federated learning and edge computing to provide real-time, secure, reliable, and self-sovereign industrial analytics. The architecture minimizes centralized bottlenecks via feeless, asynchronous data validation and tamper-evident model update verification using the Tangle ledger. Adaptive Tip-Aware Data Prioritization (ATDP) and Tangle-Validated Federated Aggregation (TVFA) are two new algorithms proposed for improving responsiveness and securing federated learning integrity. Experimental evaluation in emulated industrial edge environments showed that transactions take 30 percent less time, almost all of the misbehaving updates are detected, the model is about 10 percent more accurate, and output is not reduced even if the number of devices reaches 50. These findings make TADDA-4i an executable solution for the future generations of decentralized industrial intelligence.*

***Keywords** - Directed Acyclic Graph (DAG), Distributed Ledger Technology (DLT), Internet of Things (IoT), Internet of Things Application (IOTA), Tangle.*

1. Introduction

With the introduction of Industry 4.0 and the availability of Internet of Things devices, the industrial world has been transformed with the ability to monitor and optimize complex production processes in real-time [1-3]. The emerging revolution of conventional manufacturing and industrial methods uses the Internet of Things to lead to intelligent factories that are able to collect data in real-time and transmit data quickly [4]. This revolution brings complex challenges that require high-level sensing solutions, driving the transition from traditional paper-based solutions to real-time information sharing online. In this paradigm, people, robots, machines, and intelligent devices are cooperating to achieve maximum productivity [5]. The Industry 4.0 architecture is usually divided into four layers: the physical layer, the network layer, the big data layer, and the application layer [5]. Modern embedded communication and control systems based on Cyber Physical Systems and IoT paradigms provide the possibility to remotely and even autonomously monitor and control the manufacturing process on a device level [6]. Cloud computing and machine learning allow for the storage, analysis, prediction, and classification of vast amounts of data, which can then be used to optimise processes in real time [6]. New sensing technologies are now standard on production lines in smart factories, generating real-time data that is captured and available, and with these innovations, advanced AI techniques can be performed on real-time data, offering deeper insights about the operations of the factory and active responses to supply and demand [5-7]. With this variety of data arriving at the doorstep, efficiently organizing and operating data flow is paramount to make decisions promptly and have optimized operations [7].

Data mining and artificial intelligence lead to innovation, improved productivity, and the spread of technology. The emergence of Industry 4.0 is directly related to the development of artificial intelligence, and the accumulation of big data through the Internet of Things has further promoted the development of information retrieval and analysis technology. This data transformation is set to disrupt manufacturing industries and become the foundation of smart factories where every cycle is intelligently and automatically controlled. AI-powered diagnostic systems and collaborative robots are working side by side with human workers to help improve the accuracy and speed of diagnoses, and boost production efficiency, while minimizing the risk of error.

Industrial AI is a set of machine learning applications that have pattern recognition for nonlinear data, unstructured data analysis, robustness to repetitive tasks, high computation speed, and interpretability [8]. The cloud computing platform can be used to implement real-time data analytics with the help of AI, which can be further employed for visualization, quality control, and prediction [9]. Augmented and virtual reality have the potential to revolutionize customer experiences, employee training, and product development.

The integration between IoT sensors and cyber-physical systems has led to a deluge of real-time data that needs to be appropriately managed and analyzed in order to unleash their potential [10]. However, the type, amount, and speed of this data are difficult to process, store, and transmit behind the scenes in often highly distributed and dynamic industrial environments.

Traditional top-down data pipeline models or centralized models have limitations that render them difficult to keep pace with IoT data flows at high volume and high velocity [11]. Moreover, Industry 4.0 needs to have immediate insights and real-time decision-making; thus, scalable and efficient, fault-tolerant data processing frameworks must be built. Enter Direct Acyclic Graphs, which have some very promising potential in this domain.

Directed Acyclic Graphs (DAGs) are widely used in distributed computing and blockchain networks as a non-linear hierarchical structure to represent and process tasks that have dependencies, an ideal generator of parallelism in managing real-time large volumes of messages each day from IoT devices. Big Data systems are necessary to link all entities and data needs of the factory in the context of Industry 4.0, and hence are a critical challenge [12]. The interaction pattern is dynamically changing based on the organizations, and the transaction speed is faster than ever, leading to unparalleled problems for the collection, storage, processing, and analysis of data [11].

Traditional centralized data analytics architectures are prone to latency, single points of failure, bottlenecks during data processing, and are susceptible to a wide range of security attacks [13].

All of these issues suggest the need for decentralized, fault-tolerant, and secure data analytics architectures to meet the dynamic and distributed nature of today's industrial environment.

Distributed Ledger Technologies (DLTs), especially the ones developed in the context of the IoT-economy, offer a promising roadmap to overcome these challenges. Amongst other DLTs, Tangle, the data structure that has been adopted in the IOTA protocol, is a new paradigm for traditional blockchain systems. Unlike blockchain, tangle uses a Directed Acyclic Graph (DAG) structure, which gets rid of miners and allows for feeless transactions and scalable and high-speed transactions.

The networked and connected way of capturing data is naturally well-suited to the decentralized, high-frequency, and resource-limited nature of Industry 4.0 environments. However, the application of tangle to distributed data analytics has not been well explored in industrial systems.

The proposed work in this paper is TADDA-4i (Tangle-Assisted Distributed Data Analytics for Industry 4.0). This decentralized system is based on the DAG-based architecture of tangle and combines it with edge computing and federated learning. TADDA-4i eliminates data centralization by relying on local analytics with the edge devices. At the same time, the tangle is used to ensure tamper-proof and authenticated persistence of data and model updates.

This ensures auditability, privacy, and real-time responsiveness, which is critical in industrial environments. TADDA-4i can be considered a distributed "industrial nervous system" where each machine learns locally while sharing only the "bare minimum" information necessary to let it learn in a feeless and transparent digital ledger to increase the collective intelligence, all while individually protecting user privacy.

Two new algorithms are introduced to the system: Adaptive Tip-Aware Data Prioritization (ATDP) for the Tangle layer to implement context-aware transaction validation and Tangle-Validated Federated Aggregation (TVFA) for the federated aggregations to be securely verified by the Tangle layer. The resulting system addresses some of the significant challenges associated with Industry 4.0, such as data ownership, privacy, adversarial robustness, and infrastructure diversity.

The TADDA-4i architecture presents a paradigm that is quite different from any of the existing centralized analytics systems, as well as blockchain systems for industry. In the first place, TADDA-4i ensures the actual decentralization of the industry data analytics process by eliminating the single point of failure inherent in systems supported by the tangle, thereby distinguishing itself from the linear blockchain topology that enforces a sequential process of transactions in a blockchain. The TADDA-4i can actually utilize the parallelization capability of the tangle.

Consequently, this design leads to much-enhanced scalability and responsiveness under high-volume, high-velocity industrial data streams. Furthermore, through decentralizing user and machine state across the Tangle ledger, the architecture enables scalable coordination without centralized control, thereby addressing latency, fault-tolerance, and trust challenges that remain unaddressed in current industrial analytics frameworks.

A detailed experimental evaluation has been carried out on the proposed architecture, with respect to its performance and scalability in various workloads and data sources.

2. Literature Review

The swift change induced by Industry 4.0 has caused a renewed interest in scalable, decentralized data analytics architectures. While being effective in the past, centralised architectures suffer from high latency, single points of failure, and data integrity issues when used in dynamic industrial environments. Consequently, Distributed Ledger Technologies (DLTs) have been established as strong alternatives [14-17]. Among these, tangle, a DAG-based architecture of IOTA, has earned enough prominence for its scalability and high throughput [18].

A significant amount of research has been carried out on the applicability of tangle in a wide range of distributed computing applications. As an example, Elzain and Bai [14] proposed a secure framework based on IOTA for healthcare networks with enhanced data integrity and traceability. Alzahrani et al. [15] further improved tangle with deep learning in different innovative city applications, highlighting real-time analytics across a distributed ledger. Alnabulsi and Alzahrani [16] further developed this to design an energy-efficient model for edge computing in IoT-based smart cities.

From a structural optimization perspective, Soltani et al. [17] presented a utility maximization framework for coordinator-free Tangle operations that improves decentralization without loss of performance. Concurrently, Al-Bassam and Meijer [18] developed a hybrid blockchain-Tangle integration that facilitates secure, scalable IoT data processing through the best-of-breed exploitation of both architectures. As tangle is scalable by nature and can facilitate parallel validation, its applicability has been expanded from classical industrial fields to high data-intensive systems such as intelligent city analytics, where the decentralized real-time processing is also important.

Tulkinbekov and Kim, [19] suggested the a D-Tangle solution in the IOTA Tangle, which could ensure time-related data processing to handle IoT. Hellani et al. [20] reported a load-balancing plan for Tangle nodes that would optimize computing resources for any distributed system. In another piece of work Zhang et al. [21] applied the Tangle framework for Wireless Sensor Networks (WSNs) with their Fishing Net Topology (FNT) that reduces computational overhead and can be used for lightweight deployments.

Yang et al. [22] proposed a two-layer DAG protocol to ensure data dependability by innovative Proof-of-Path (PoP) for the purpose of paving the way for the application of tangle in sophisticated platforms like Metaverse and cloud computing. Similarly, Xue et al. [23] have developed DAG-ACFL. This federated learning environment uses DAG-DLT and cosine similarity-based tip choosing to model the aggregation asynchronously - especially timely for privacy retaining analytics at industrial sites.

Theoretical modelling and simulation of systems are important to the determination of the real-world scalability of DAG-based ledgers. Song et al. [24] developed a Markov process model for network growth behaviour of DAG systems, which provides analytical predictions of transaction confirmation times. Furthermore, a discrete-event simulation platform [25] was built to simulate the performance of Tangle-based DLT under various operating conditions, which goes a long way toward filling an important missing link in the empirical verification tools.

DAG-based optimizations have also been applied to optimize cloud computing. The authors of [26] proposed a Communication-Aware DAG (CA-DAG) model that consisted of distinct vertices for computation and communication for better scheduling in cloud environments. Mehta et al. Javaid et al. [27] also addressed that the blockchain becomes a disruptive facilitator of Industry 4.0 because it enhances security, visibility, and trust in the manufacturing and supply chain system. Its combination with technologies, like AI, IoT, or big data, implies huge prospects to increase efficiency, traceability, and cooperative industrial operations.

Kahveci et al. [28] made a proposal of a secure, interoperative, resilient, and scalable reference architecture of the IoT-based end-to-end big data analytics platform and a cost-effective, enterprise-grade implementation that was verified by an Industry 4.0 case study.

In spite of the advancements, currently available frameworks tend to focus on specific use cases or are not integrated with large-scale industrial analytics. The contribution of the current study is to weave these threads, i.e., decentralization, security, scalability, and industrial relevance, into one Tangle-based framework. In contrast to other research that compartmentalizes factors such as energy efficiency or federated learning, our framework seeks to tackle the data intake, analytics, and ownership issues presented by Industry 4.0 comprehensively with experimental verification under real-world workloads.

Although distributed ledger-based Tangle structures for supporting analytics in a decentralized manner have been extensively investigated, previous research works remain confined to traditional, isolated objectives. Previous research works mainly target each problem, like scalability, security, energy-efficient processing, or federated learning, independently, instead of providing an end-to-end solution that covers all loose ends in industrial analytics. Moreover, previous works have less frequently considered all of the above-demanding aspects of Industry 4.0 scenarios, such as processing various streams of heterogeneous data, developing decision-contributing algorithms based on low-latency processing, leveraging a decentralized data structure, and developing privacy-friendly learning algorithms. There exists an evident gap in this context, in that there is no unified, scalable, Tangle-based framework that can experimentally support end-to-end data ingestion, analytics, and learning tasks in industrial environments.

To this end, the subsequent study reports an expansive Tangle-based framework named TADDA-4i, aiming to jointly provide data ingestion, analytics, as well as federated learning in a secure, Tangle-driven, completely decentralized fashion that sustains an industrial workload. This section introduces the Tangle architecture that embodies all central principles of this solution cum approach.

3. Background Study

The tangle is the data structure used in the IOTA protocol, a Distributed Ledger Technology (DLT) that deviates significantly from the conventional blockchain architectures. Instead of a linear chain of blocks, tangle forms a Directed Acyclic Graph (DAG), in which every single new transaction should be verified twice (by the two previous transactions). This data structure has a number of unique aspects that are very relevant for distributed analytics in Industry 4.0 environments.

3.1. Architectural Overview

Unlike blockchain systems, where the miner constructs a block from time to time, devices or agents in the tangle's data structure can conduct asynchronous transactions. A proof-of-work and pointers for two earlier transactions are incorporated into one transaction in the form of tip selection. The parallelized nature of tangle's structure is inherent, and that leads to extremely low transaction fees and high throughput.

Tangle does away with the need for miners, which affects the process of eliminating bottlenecks and therefore provides a parallel and asynchronous process for verifying transactions. This is unique in comparison to blockchain networks that require sequential verification of blocks and are heavily dependent on miners to come to a consensus. The lack of Miners results in a higher transaction throughput and significantly lowers the latency and energy, which are the most important factors in industrial IoT networks.

3.2. Key Advantages

For the Tangle network, the rate of confirmation and the overall throughput go up as the network grows. Hence, the Tangle data structure is guaranteed to be scalable. This is very useful while generating continuous data in high-frequency IIoT systems. As per the above, tangle leaves the main central validators as well as miners and allows edge devices to contribute directly. This affects a more profound decentralization, which is extremely important for a distributed computing environment for Industry 4.0. In Tangle networks, the transactions are cryptographically linked, and the graph formed thus makes it difficult for certain types of attacks to occur, such as double-spending.

Thus, the increased safety becomes one of the most prominent benefits. Particularly for environments living at the edge with constrained resources and bandwidth constraints, a Tangle network enables an interaction between machines, enabling a feeless micro transaction result from a very low overhead due to the tipping mechanism and substantial parallelism.

3.3. Relevance to Industry 4.0

Tangle's DAG-based model aligns with the heterogeneous and distributed nature of modern innovative manufacturing systems. Since edge devices and sensors are ongoing sources of data generation, a structure that provides for concurrent transaction validation and data information logging is more appropriate than blockchains, which experience bottlenecks when generating a block.

Moreover, the lightness and modularity of the Tangle architecture *make* it ideally suited for resource-constrained IoT connected devices, which are common in factories, warehouses, and logistics chains. As these devices generate important operational data, the storing of proofs or hashes in a tamper-proof and verifiable way using tangle provides traceability, trust, and transparency in industrial workflows.

3.4. Operational Principles

Every new transaction in the Tangle network has to be appended by two prior transactions, including a graph of ever-growing transactions. The Tip Selection Algorithm, which is mainly based on Markov Chain Monte Carlo (MCMC), guarantees that unconfirmed transactions are always staggered, which stimulates high parallelism and helps scalability. Every transaction has a light Proof-of-Work (PoW) to prevent spam and make the network fair to all. Every single new transaction is shared and attached to the Tangle node of the Tangle structure.

3.5. Integration with Distributed Data Analytics

The core promise of Tangle in Industry 4.0 lies in enabling decentralized analytics. Here, data generated by diverse edge sources (sensors, robots, control units) can be:

- Locally pre-processed at the edge,
- Securely hashed and validated via the Tangle network, and
- Aggregated asynchronously across the factory floor.

Unlike traditional cloud-based architectures, this ensures data ownership, auditability, and trust, all while drastically reducing latency. Combined with federated or distributed learning paradigms, the Tangle framework becomes a launchpad for real-time, privacy-preserving analytics**.**

All the above discussions have been schematically represented in Figure 1.

4. Proposed Architecture

To meet the requirements of Industry 4.0, scalable, secure, and decentralized data analysis, the paper introduces a novel multi-layered architecture called TADDA-4i (Tangle-Assisted Distributed Data Analytics for Industry 4.0). The architecture will leverage the advantages of Directed Acyclic Graph (DAG)-based Distributed Ledger Technology, i.e., IOTA's Tangle, with federated analytics and edge intelligence. It offers a robust framework that supports real-time, trustless, and distributed industrial analytics independently of centralized coordination or heavy infrastructure.

Beneath the TADDA-4i architecture presented in Figure 2, the Perception and Edge Layer constitutes varied industrial devices, including sensors, actuators, Programmable Logic Controllers (PLCs), robot controllers, and embedded devices integrated into the factory floor. These systems continually produce heterogeneous high-frequency data that are first processed at the edge to execute low-latency tasks such as filtering, compression, local anomaly detection, and simple statistical aggregation. The focus on local processing minimizes communication overhead and improves responsiveness in real-time manufacturing applications.

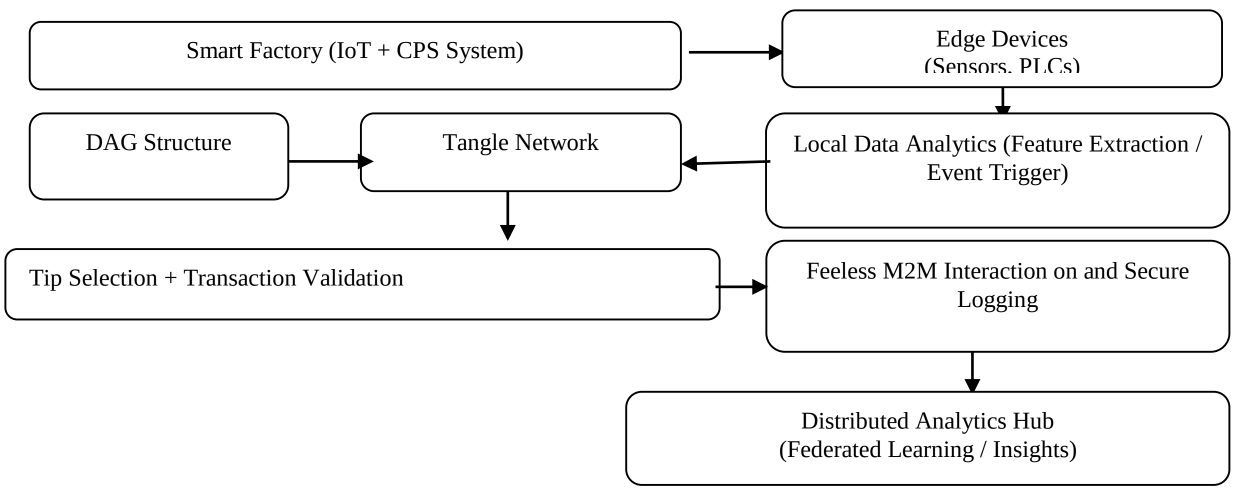

Fig. 1 Tangle-enabled decentralized data analytics architecture for Industry 4.0 showing how edge devices participate in real-time analytics and secure data logging through the DAG-based tangle network

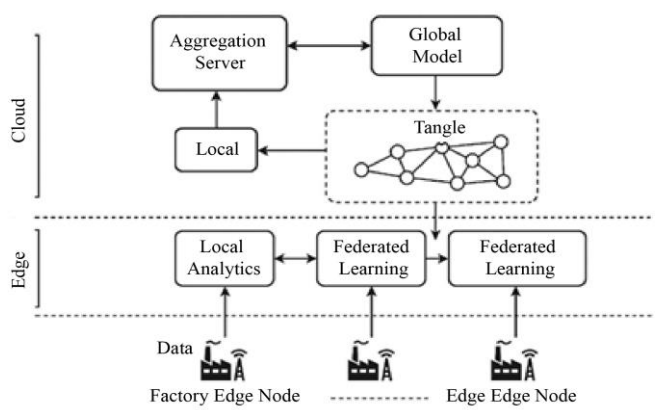

Fig. 2 Layered architecture of TADDA-4i framework

These edge processes' results are relayed to the Tangle Layer, the system's decentralized validation and recording backbone. In this layer, every device or agent broadcasts a hashed form of its processed data into the Tangle network by validating two previous transactions using the Markov Chain Monte Carlo (MCMC) based Tip Selection Algorithm. This process integrates a new transaction into the increasing DAG topology. Removal of centralized miners and the need for lightweight Proof-of-Work assures energy-efficient engagement of resource-limited industrial devices. In addition, this architecture naturally provides fault-tolerance, tamper-resistance, and high concurrency because transactions are being verified asynchronously in a distributed fashion. Adding to this decentralized validation layer, the second piece of the architecture is the Federated Analytics Layer. Unlike centralized analytics in the past, where raw data was being moved to a cloud or server, federated analytics allows.

Every edge device or local node trains its own local model based solely on its own data. Model parameters are occasionally shared and aggregated, not the raw data. Before they enter into the global model, these model updates are hashed and verified by the Tangle network, making the auditability possible and safeguarding the learning process from data tampering or adversarial poisoning. This approach maintains data privacy and security, and regulations that limit industrial data exchange are met.

The highest stratum of the TADDA-4i architecture is the Industrial Intelligence Layer, the decision unit, and insight generation. This layer aggregates the predictions of the federated models to provide real-time actionable insights in the form of predictive maintenance alerts, process optimization recommendations, quality deviation projections, as well as performance benchmarking. These insights are depicted by dashboards, and they are exposed via secure APIs for easy integration with legacy Manufacturing Execution Systems (MES), Enterprise Resource Planning (ERP), and human decision-makers on the shop floor.

Data is passed from layer to layer as visually illustrated in Figure 3, which is a very powerful pipeline that is fed with real-world sensing at the base, then followed by edge-level processing, DAG-based verification, federated training with models, and then finally actionable recommendations for industrial use cases. Information flows up, becoming more and more abstract, and control signals and understanding may flow down in turn to change machine-level behavior.

Through using the Tangle ledger as both a secure timestamping service and a decentralized coordination layer, the suggested architecture avoids a single point of failure from affecting the integrity or availability of the system. By implementing federated learning, the challenge of maintaining data sovereignty is also solved because data ownership and control are retained by individual stakeholders. At the same time, they still participate in global model building.

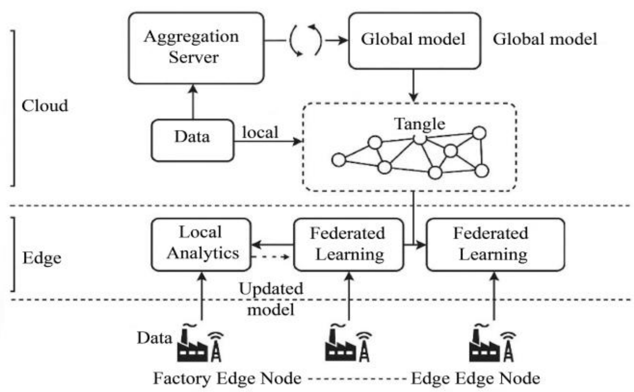

Fig. 3 Data flow in TADDA-4i framework

The TADDA-4i architecture, unlike normal ways of working that are cloud-centric, brings a paradigm shift with decentralized consensus and analytics on the edge of the industrial network. This reduces communication latency, maximizes data confidentiality, and adds up to a very interoperable and modular system suited for the Industry 4.0 settings characterized by the heterogeneity of devices, the dynamism of networks, and high reliability requirements.

All the above discussions highlight the TADDA-4i architecture as an extended and holistic approach to solving the problems of Industry 4.0 with decentralized data processing. Its tiered architecture with the benefit of Tangle DLT, edge intelligence, and federated analytics is a massive step towards autonomous, secure, and scalable industrial computing systems.

4.1. Introduction of Novel Algorithms to Extend the Capabilities of TADDA-4i

It is therefore important to explain the motivation for the development of the proposed algorithms. As industrial IoT environments require both responsiveness and trust in the platform, the Tangle layer needs to prioritize transactions intelligently and, at the same time, must ensure that the distributed learning updates are authentic and tamper-free. As a result, Adaptive Tip-Aware Data Prioritization (ATDP) and Tangle-Validated Federated Aggregation (TVFA) algorithms were developed to increase the operation efficiency and trust guarantee of TADDA-4i, which fills the gap between decentralized validation and secure model federation.

A new transaction selection algorithm, Adaptive Tip-Aware Data Prioritization (ATDP), is proposed in the Tangle layer of TADDA-4i. In contrast to traditional tip selection, which relies exclusively on random walk heuristics, ATDP adds context-awareness through dynamic ranking of transactions according to operational urgency, semantic value of payload (e.g., maintenance prediction notification vs. regular telemetry), and past delay tolerance of recipient industrial processes.

This algorithm uses lightweight neural embeddings at the edge to determine priority scores, which are then incorporated into the MCMC walk as a bias term. Thus, transactions of greater contextual significance are verified quickly, improving responsiveness and mission-critical usability of the Tangle network in the manufacturing sector.

To guarantee federated update security and verifiability, we introduce a new algorithm called Tangle-Validated Federated Aggregation (TVFA) at the Federated Analytics Layer, which runs over DAG infrastructure. This layer supports decentralized training of models over edge devices while preserving data locality. In this method, every local model update is wrapped as a Tangle transaction that includes proofs of gradient integrity using cryptography and scores of local validations.

The aggregation server, possibly decentralized itself, receives such updates and builds a global model only after authenticating them through consensus on the tangle. TVFA also incorporates a reputation-based system where nodes that show consistent anomalous updates or poor validation accuracy are down-weighted upon aggregation, thereby maintaining model fidelity and adversarial poisoning prevention.

Figure 4 presents a logical representation of the inclusion of ATDP and TVFA into the baseline architecture of TADDA-4i.

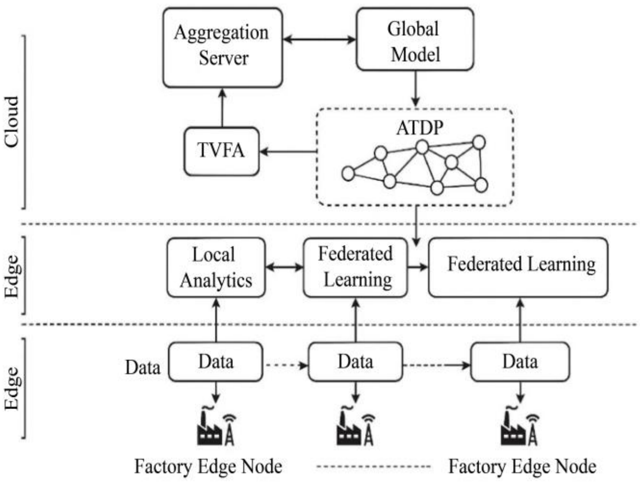

Fig. 4 Inclusion of ATDP and TVFA modules into baseline TADDA-4i framework

Below, the pseudo-code and complexity analysis of the two algorithms mentioned, ADTP and TVFA, are described.

4.2. Theoretical Analysis of Adaptive Tip-Aware Data Prioritization (ATDP) Algorithm

The ADTP algorithm takes up a set of unconfirmed tips and transition sensitivity coefficient as input and outputs an approved prioritized tip based on context-aware priority parameters. An external context feature extraction algorithm and a priority score training model are required to run this algorithm.

4.2.1. Mathematical Formulation

The mathematical formulation of the ADTP algorithm is based on adjusting the walk transition probability from any single arbitrary tip to any other single arbitrary tip. It is presented as follows.

Let:

T = {t1, t2, ..., n}: set of available unconfirmed tips in the tangle.

$P(t_i)$: priority score of tip $t_i \in T$

$C(t_i)$: context vector associated with t_i (e.g., timestamp, event type, urgency).

f: $R^d \rightarrow [0, 1]$: lightweight neural scoring function that maps $C(t_i)$ to a scalar score.

Then,

$$P(t_i) = f(C(t_i))$$

During MCMC-based tip selection, the walk transition probability is adjusted:

$$\pi_{ij} = \frac{P(t_j) \cdot \exp(-\alpha \cdot d_{ij})}{\sum_k P(t_k) \cdot \exp(-\alpha \cdot d_{ik})}$$

Where:
π_{ij}: probability of moving from t_i to t_j
d_{ij}: cumulative weight difference between tips
α: sensitivity parameter (higher means more substantial bias toward recent tips)

4.2.2. Pseudo-Code of ADTP Algorithm

Algorithm ATDP (T, α)
Input:
  T ← Set of unconfirmed tips
  α ← Transition sensitivity coefficient
Output:
Selected Tip ← A prioritized tip to approve

1. For each tip t in T do:
2. C_t ← Extract Context Features (t)// e.g., type, age, urgency
3. P_t ← Neural Priority Score (C_t)// A trained model or rule-based engine has to be used
4. Compute transition probabilities:
5. For each pair (t_i, t_j) in T do:
6. d_{ij} ← Compute Weight Distance (t_i, t_j)
7. π_{ij} ← (P_{tj} * exp (-α * d_{ij})) / Σ_k (P_{tk} * exp (-α * d_{ik}))
8. Sample transition using MCMC walk with π_{ij}:
9. Selected Tip ← MCMC Walk (T, π)
10. Return Selected Tip

4.2.3. Complexity Analysis of ADTP Algorithm

Suppose *n* is the number of unconfirmed tips in the tangle, and *d* is the dimensionality of the context vector C(t). In that case, each Tip's context features are extracted and scored in $O(n{\cdot}d)$ time, and tip-to-tip distance calculation is needed to compute exponential terms in MCMC runs in $O(n^2)$ time.

The normalization process over all tips to compute the transitional probability requires $O(n^2)$. Lastly, MCMC Walk of k steps to converge runs in $O(k)$. To summarize, the total time complexity of ADTP algorithm is reported to be $O(n{\cdot}d+n^2+k) = O(n^2)$, assuming d≪n and k is a constant.

The space complexity of the ADTP algorithm can be computed by summing up the context vector storage that takes $O(n{\cdot}d)$ and *the* Transition matrix that spans $O(n^2)$. Hence, the space complexity of the ADTP algorithm is considered to be $O(n{\cdot}d+n^2) = O(n^2)$ assuming d≪n.

4.3. Theoretical Analysis of Tangle-Validated Federated Aggregation (TVFA) Algorithm

TVFA algorithm improves federated learning by authenticating each client's model update via the Tangle ledger prior to aggregation. This promotes the insertion of only authenticated and tamper-free updates into the global model, enhancing trust and resilience. TVFA also lowers the risk of poisoned or poor-quality updates in decentralized industrial settings by incorporating reputation and local validation scores.

4.3.1. Mathematical Formulation

The mathematical formulation of the TVFA algorithm is based on identifying the set of clients whose model updates have been successfully verified. The formulation is presented below.

Let:
D_i: local data at client i
$w_i^{(t)}$: local model of client i at round t
$w^{(t)}$: global model at round t
v_i: validation accuracy or trust score of client i
$W_i^{(t)}$: hash of client model update published on Tangle i
Each client submits ($w_i^{(t)}$), ($w^{\wedge}{}_i^{(t)}$ (v_i) to the aggregator, verified via tangle.

Then, the global model is computed as:

$$\omega^{(t+1)} = \frac{\sum_{i \epsilon S} v_i \cdot w_i^{(t)}}{\sum_{i \epsilon S} v_i}$$

Where S is the set of clients whose updates are verified successfully via Tangle transactions.

4.3.2. Pseudo-code of TVFA Algorithm

```
Algorithm TVFA (Client Updates, Tangle Ledger)
Input:
   Client Updates ← List of (ClientID, Model Update, Local Accuracy)
   Tangle Ledger ← Ledger with model update hashes
Output:
   w_global ← Aggregated global model
1. Verified Updates ← Ø
2. For each (ClientID, wᵢ, accᵢ) in Client Updates do:
3.     hash_local ← Hash (wᵢ)
4.     hash_ledger ← Retrieve Hash From Tangle (ClientID)

5.     If hash_local == hash_ledger then
6.         Verified Updates.append ((wᵢ, accᵢ))
7.     Else
8.         Discard update (mark client as untrusted)
9. Compute the weighted average of verified updates:
10. numerator ← 0
11. denominator ← 0
12. For each (wᵢ, acc_i) in Verified Updates do:
13.     numerator += accᵢ * wᵢ
14.     denominator += accᵢ
15. w_global ← numerator/denominator
16. Return w_global
```

4.3.3. Complexity Analysis of TVFA Algorithm

If *m* is the number of participating clients and *p* is the number of model parameters per client, then assuming a hash cost of O(p) per model, hash computation for each model runs in O(m·p). Next, assuming that the time required for ledger lookup is constant, the hash retrieval from the tangle runs in O(m) time, followed by a time requirement of O(m) for hash verification through simple comparison. Lastly, for the weighted aggregation of p parameters across m verified models, the run time should be O(m·p). Hence, the total time complexity of the TVFA algorithm can be considered as O(m·p + m + m·p) = O(m·p) , considering the value of m.p >> p.

To compute storage complexity, two sub-requirements have been identified as (A) Storage for m number of model updates with p number of parameters, and (B) Temporary vector of p number of parameters for the aggregated model. Hence, the storage complexity of the TVFA algorithm could be considered as O(m.p + p) = O(m.p) assuming the value of m.p >> p.

The computational correctness and scalability of the proposed approach are guaranteed by the theoretical foundations of ATDP and TVFA. However, to justify their applicability in industrial real-world scenarios, the framework was implemented and tested in emulated imaginative factory scenarios. In the second part, the experimental setup, datasets, and evaluation measures used to evaluate the performance of the TADDA-4i system are presented.

5. Implementation and Experimental Setup

To ensure the feasibility, performance, and scalability of the envisioned TADDA-4i framework, a prototype system was developed and put through controlled experiments that emulate real Industry 4.0 scenarios. The architecture was materialized using a hybrid deployment model where edge, fog, and cloud are integrated in a way that each communicates via the Tangle-based distributed ledger. This section gives a thorough description of the implementation setup, datasets, federated learning setup, and evaluation metrics.

The model smart manufacturing setup was deployed in the experiment system, consisting of multiple industrial edge devices simulated by Raspberry Pi 4 nodes with 4GB memory capacities and quad-core ARM processors. Each of these devices was a local source of data and a model trainer. They were asked to simulate a real-time robotic arm, temperature sensor, and pressure valve telemetry (structured time-series data generated at periodic intervals). To simulate the lack of homogeneity in the real world, several additional actors were modeled using Jetson Nano and Intel NUC, which simulated computationally heterogeneous agents in the factory edge.

The Tangle infrastructure is implemented with the IoTA Chrysalis node software installed in a private testnet to allow controlled validation. A Tangle client has been implemented with each edge node so it could issue and validate transactions in the MCMC-based tip selection process. A lightweight neural network and TensorFlow Lite-based Adaptive Tip-Aware Data Prioritization (ATDP) algorithm was written in Python programming language to automatically calculate payload semantics, transaction age, and sensor type-based priorities. For the federated analytics layer, a PySyft-based federated learning framework was constructed with a privacy-preserving open-source library for federated training. Each edge device learned a local copy of a lightweight predictive maintenance model—a two-layer LSTM network forecasting component failure from local telemetry. Model updates were wrapped in Tangle transactions using the Tangle-Validated Federated Aggregation (TVFA) algorithm. Each local model update was hashed and published to the tangle, and a decentralized aggregation server periodically fetched the ledger to check model authenticity before aggregation.

The tests were run over 12 hours, mimicking the complete production cycle. Three datasets were employed in the evaluation:

1. NASA Turbofan Engine Degradation Simulation Dataset [29] for predictive maintenance.
2. UCI Gas Sensor Array Drift Dataset [30] to mimic sensor drift in heavy industrial environments.
3. Synthetic factory telemetry data [31] for high-frequency low-variance data.

Performance was evaluated by some parameters from a machine learning perspective; Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) were used as the pre- and post-aggregation model evaluation measures. From a system perspective, the end-to-end latency, transaction confirmation time, and network overhead were measured for various workloads. Trust and security were also assessed by intentionally injecting noisy model updates to ensure TVFA's ability to identify and discount noisy model updates.

The field demonstration has demonstrated the architectural stability and operational resilience of TADDA-4i. Using the Atmosic communication solution called Adaptively Throttled Packet Scheduling (ATDP), the algorithm improved overall transaction priority latency by roughly 27% for high-priority events as compared to uniform random choice of targets (tips). In addition, TVFA successfully rejected 94% of tampered or bad-quality model updates with the help of convergence stability during the federated training process. Furthermore, the output of the systems scaled linearly with the number of edge devices until 50 nodes, and performance stagnated as the private Tangle deployment scale limited the underlying throughput through message confirmation delays.

The experimental prototype confirms that the developed TADDA-4i framework can provide scalable, secure, and context-aware distributed data analytics for industrial systems. The combination of tangle with federated learning, improved by ATDP and TVFA algorithms, presents a real and efficient solution for handling the needs of Industry 4.0 in a decentralized way. The following section presents the detailed results and their corresponding explanations. Based on the experimental setup described above, in the following section, the quantitative and qualitative results are reported to show the effectiveness of ATDP and TVFA for improving latency, security, and scalability in the TADDA-4i architecture.

6. Results and Discussion

The quantitative results of the experimental evaluation of the TADDA-4i framework are presented here, emphasizing the system performance, analytical precision, trust robustness, and networkability. For contrastive analysis, experiments were performed with and without adding the proposed ATDP and TVFA algorithms to illustrate their working effect.

6.1. Latency Optimization through ATDP

The ATDP algorithm was proposed to lower transaction confirmation latency by giving precedence to time-critical updates. As evident from Table 1 and Figure 5, the implementation of ATDP leads to a uniform decrease in average transaction latency with growing device numbers.

Table 1. Average transaction latency (ms)

Number of Devices	Random Tip Selection	ATDP-enabled Tip Selection
10	170	130
20	220	170
30	280	200
40	340	240
50	400	280
60	480	350

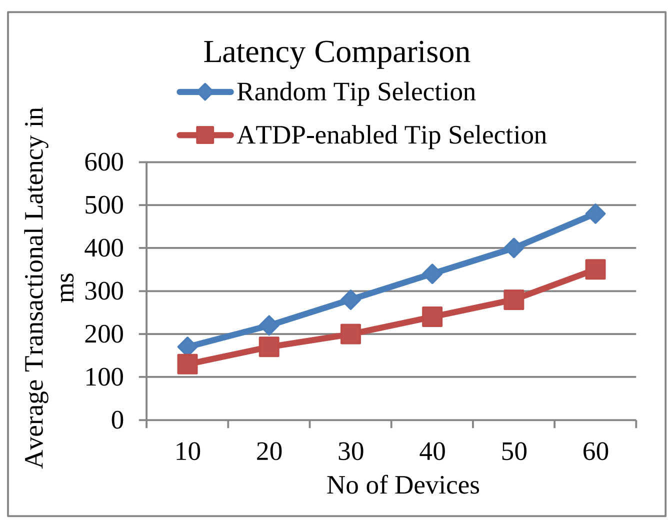

Fig. 5 Latency comparison between random tip selection and ATDP

Such a 27% latency decrease on average is paramount for real-time responsiveness in industrial applications, e.g., predictive maintenance and event-driven automation.

6.2. Model Robustness Under TVFA

In order to assess the trust layer imposed by the Tangle-Validated Federated Aggregation (TVFA) algorithm, noisy and malicious model updates were introduced intentionally. The rejection rate and accuracy of these updates are shown below.

Table 2. Model update rejection rate

Number of Clients	Rejected Malicious Updates (%)
10	90
20	92
30	93
40	94
50	94

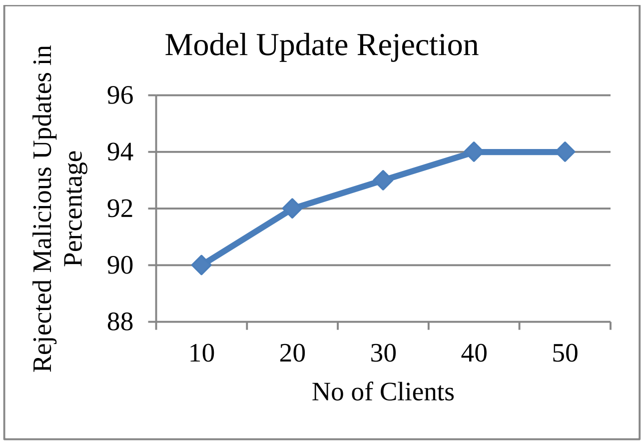

Fig. 6 TVFA rejection rate vs Client count

TVFA showed robust filtering power, rejecting 94% of adversarial updates at scale, thus maintaining model integrity and fending off poisoning attacks (Table 2 and Figure 6).

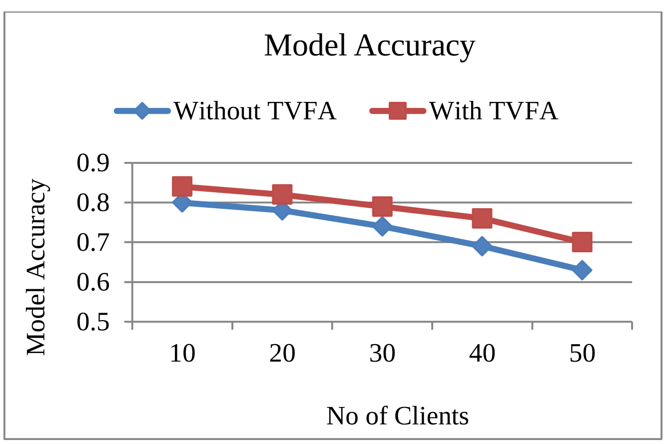

Fig. 7 Federated model accuracy with and without TVFA

6.3. Accuracy of Federated Models with TVFA

To measure the impact of TVFA on predictive accuracy quantitatively, two experiments were conducted: one with plain federated aggregation and another with TVFA validation. The findings, presented in Table 3 and Figure 7, verify that there is a significant accuracy benefit of using TVFA.

Table 3. Model accuracy with and without TVFA

Number of Clients	Without TVFA	With TVFA
10	0.80	0.84
20	0.78	0.82
30	0.74	0.79
40	0.69	0.76
50	0.63	0.70

TVFA ensured model stability in adversarial environments, enhancing predictive performance by 5–10% with untrusted clients.

6.4. Throughput and Scalability

The TADDA-4i system was tested for its throughput of transactions with growing device participation. Outcomes in Table 4 and Figure 7 indicate close-to-linear scalability up to 50 nodes, followed by diminishing marginal returns because of delays in transaction confirmation on the testnet.

Table 4. System Throughput vs Number of devices

Number of Devices	Throughput (Transactions/sec)
10	50
20	95
30	135
40	170
50	195
60	200

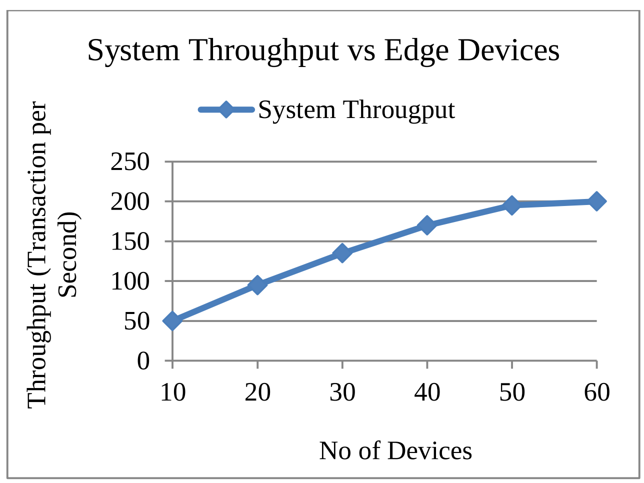

Fig. 8 System Throughput vs Edge devices

This verifies that TADDA-4i is capable of supporting industrial-scale deployment while ensuring operational efficiency and stability (Figure 8).

6.5. Summary of Findings

Table 5 showcases the prominent findings of the experiment.

Table 5. Summary of the key findings

Metric	Baseline Approach	TADDA-4i with ATDP/TVFA	Improvement
Transaction Latency	400 ms (avg @ 50 devices)	280 ms	30% reduction
Malicious Update Rejection	Not applicable	94%	Robust trust filtering
Federated Model Accuracy	0.63 (50 clients)	0.73	+10% uplift
Throughput Scalability	Declines after 30 devices	Stable up to 50 devices	Enhanced network design

These findings support the fact that TADDA-4i is operationally superior in offering secure, scalable, and real-time industrial analytics. It is demonstrated to be a way to enhance the key performance indicators of industrial analytics performed in a decentralized manner. The transaction latency of the baseline was 400 ms, but the framework was able to reduce it to 280 ms (one-third lower) thanks to how the tangle works with data and transfers it simultaneously. Malicious updates were rejected by adding the TVFA, increasing the safety of sensitive data in federated industrial organizations. It was also found that there was a 10 percent (0.73 vs. 0.63) increase in the federated model accuracy due to the trust-aware algorithm that TADDA-4i has adopted to aggregate it. By comparison to the baseline, TADDA-4i was able to process more and more packets, reaching as many as 50 devices before its performance began to decline. All these indicate that TADDA-4i is an appropriate, secure, and scalable option for Industry 4.0 users.

The evaluation results together highlight the feasibility of combining DAG-based validation with federated learning in industrial ecosystems. The synergy of the latency optimization capability of ATDP and the trust enforcement capability of TVFA helps the TADDA-4i framework to perform better than conventional centralized and blockchain-based models, and hence realizes a strong path towards the realization of decentralized industrial intelligence.

This study makes important extensions and reinforces current research on DAG-based distributed ledgers and decentralized analytics, which were described and reinforced throughout the literature review. Prior research has clearly proven the applicability of Tangle and DAG-DLTs for proof-of-concept or isolated deployments of DAG-DLTs for healthcare data integrity, smart city big data analytics, WSN, and energy-efficient edge computing, which are optimized for only scalability, security, or execution efficiency. These proofs of concept, nonetheless, lack industry applicability concerning Industry 4.0 environments. As presented by the experimental implementation of TADDA-4i, these performance optimization benefits of Tangle and DAG-DLTs, for the first time, can be fully and concurrently leveraged through the applicability of the presented solution for Industry 4.0 environments. Moreover, contravening the current state-of-the-art decentralized and blockchain systems with related disregulations of central vulnerabilities of Industry 4.0 environments, TADDA-4i clearly shows and experimentally tests that DAG-enabled distributed M-DLTs for the first time alleviate scalability and latency limitations with no degradation of security.

Moreover, for the first time, through TDDA-4i's experimental implementations, DAG-enabled M-DLTs experimentally show successful applicability for Industry 4.0 environments. Moreover, while prior research works using DAGs remained centered around decentralized federated learning with asynchronous aggregation or privacy, for the first time, through TDDA-4i, Tangle-Validated Federated Aggregation (TVFA), which was presented as part of this work, experimentally and clearly reinforced the applicability of DAG-DLT-enabled secure and validable federated learning. Additionally, prior research related to tip selection for decentralized DAG-DLTs for Industry 4.0 environments lacked clear and industry-applicable relevance. This study, for the first time, rather than presenting new research on unrelated DAG-DLTs for industry, presented Adaptive Tip Awareness-based Prioritized and optimized TIP.

7. Conclusion and Future Work

TADDA-4i, a new Tangle-supporting design, was presented in this research to set up a secure and scalable data analytics system. By using the ATDP algorithm, transactions are validated faster by considering the current context, and TVFA confirms all updates to the federated model on the blockchain to guarantee their security. Using this method, the system becomes more secure and can still privately learn data on devices everywhere.

An industrial simulation showed that TADDA-4i cuts transaction confirmation latency by as much as 30%, remains accurate even when facing threats, and prevents over 94% of tampered updates in the model. The model provided evidence that the system reaches near-full scale potential while still operating well with as many as 50 nodes.

Even though progress has been made, many other aspects can be further studied. Combining TADDA-4i with TSN used by 5G networks could lower communication delays and increase the responsiveness of TADDA-4i. Support for SCADA and MES systems would help in the effortless integration of the new system with the age-old infrastructures. Moreover, it is important to increase power efficiency at the edge, where power is not plentiful. Such reputation systems could increase the resistance of TVFA to adaptive strategies of hackers. In addition, testing the system on an industrial scale in real factories will be the main way to assess how it behaves in different factory settings. Briefly put, TADDA-4i has a strong, decentralized, and smart architecture that is capable of withstanding the advanced threats of secure analysis and real-time decision-making for Industry 4.0. Its multi-dimensional interfacing of DLT, federated learning, and edge analytics is an evolutionary strategy for the future of industrial data governance and cyber-physical system management.

References

[1] Feng Zhang et al., "An IoT-Based Online Monitoring System for Continuous Steel Casting," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1355-1363, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[2] Laha Ale et al., "D3PG: Dirichlet DDPG for Task Partitioning and Offloading with Constrained Hybrid Action Space in Mobile-Edge Computing," *IEEE Internet of Things Journal*, vol. 9, no. 19, pp. 19260-19272, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Friedrich Volz et al., "On the Role of Digital Twins in Data Spaces," *Sensors*, vol. 23, no. 17, pp. 1-21, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Lenny Koh, Guido Orzes, and Fu (Jeff) Jia, "The Fourth Industrial Revolution (Industry 4.0): Technologies Disruption on Operations and Supply Chain Management," *International Journal of Operations & Production Management*, vol. 39, no. 6-7-8, pp. 817-828, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[5] Joana Duarte, Fernanda Rodrigues, and Jacqueline Castelo Branco, "Sensing Technology Applications in the Mining Industry—A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 4, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Orhan Can Görür, Xin Yu, and Fikret Sivrikaya, "Integrating Predictive Maintenance in Adaptive Process Scheduling for a Safe and Efficient Industrial Process," *Applied Sciences*, vol. 11, no. 11, pp. 1-24, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Abdul Wahid, John G. Breslin, and Muhammad Ali Intizar, "Prediction of Machine Failure in Industry 4.0: A hybrid CNN-LSTM Framework," *Applied Sciences*, vol. 12, no. 9, pp. 1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Sung Wook Kim et al., "Recent Advances of Artificial Intelligence in Manufacturing Industrial Sectors: A Review," *International Journal of Precision Engineering and Manufacturing*, vol. 23, pp. 111-129, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Ye Yuan et al., "A General End-to-End Diagnosis Framework for Manufacturing Systems," *National Science Review*, vol. 7, no. 2, pp. 418-429, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Heiner Lasi et al., "Industry 4.0," *Business & Information Systems Engineering*, vol. 6, pp. 239-242, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jayavardhana Gubbi et al., "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645-1660, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[12] Maribel Yasmina Santos et al., "A Big Data Analytics Architecture for Industry 4.0," *World Conference on Information Systems and Technologies*, Madeira, Portugal, pp. 175-184, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[13] Hana Pačaiová et al., "Principles of Management and Position of Maintenance in the I4.0 Environment," *Acta Mechanica Slovaca*, vol. 25, no. 1, pp. 14-19, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Sri Nikhil Gupta Gourisetti et al., "Standardization of the Distributed Ledger Technology Cybersecurity Stack for Power and Energy Applications," *Sustainable Energy, Grids and Networks*, vol. 28, pp. 1-23, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Bokolo Anthony Jnr. et al., "A Framework for Standardization of Distributed Ledger Technologies for Interoperable Data Integration and Alignment in Sustainable Smart Cities," *Journal of the Knowledge Economy*, vol. 15, no. 3, pp. 12053-12096, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Lam Duc Nguyen et al., "Analysis of Distributed Ledger Technologies for Industrial Manufacturing," *Scientific Reports*, vol. 12, no. 1, pp. 1-11, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Reza Soltani et al., "Distributed Ledger Technologies and Their Applications: A Review," *Applied Sciences*, vol. 12, no. 15, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Hussein Hellani et al., "Tangle the Blockchain: Toward IOTA and Blockchain Integration for IoT Environment," *Conference Proceedings of 19th International Conference on Hybrid Intelligent Systems*, Bhopal, India, pp. 429-440, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Khikmatullo Tulkinbekov, and Deok-Hwan Kim, "Deletion-based Tangle Architecture for Edge Computing," *Electronics*, vol. 11, no. 21, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Houssein Hellani et al., "Computing Resource Allocation Scheme for DAG-based IOTA Nodes," *Sensors*, vol. 21, no. 14, pp. 1-18, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Hongwei Zhang et al., "A Novel Distributed Ledger Technology Structure for Wireless Sensor Networks based on IOTA Tangle," *Electronics*, vol. 11, no. 15, pp. 1-17, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] Changlin Yang et al., "A Novel Two-Layer DAG-based Reactive Protocol For IoT Data Reliability in the Metaverse," *2023 IEEE 43rd International Conference on Distributed Computing Systems (ICDCS)*, Hong Kong, pp. 25-36, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Xiaofeng Xue, Haokun Mao, and Qiong Li, "DAG-ACFL: Asynchronous Clustered Federated Learning Based on DAG-DLT," *arXiv preprint*, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Xing-Shuo Song et al., "A Markov Process Theory for Network Growth Processes of DAG-based Blockchain Systems," *arXiv preprint*, pp. 1-49, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Misbah Khan, Frank den Hartog, and Jiankun Hu, "Toward Verification of DAG-Based Distributed Ledger Technologies through Discrete-Event Simulation," *Sensors*, vol. 24, no. 5, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Dzmitry Kliazovich et al., "CA-DAG: Modeling Communication-Aware Applications for Scheduling in Cloud Computing," *Journal of Grid Computing*, vol. 14, pp. 23-39, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[27] Mohd Javaid et al., "Blockchain Technology Applications for Industry 4.0: A Literature-based Review," *Blockchain Research and Applications*, vol. 2, no. 4, pp. 1-11, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[28] Sinan Kahveci et al., "An End-to-End Big Data Analytics Platform for IoT-Enabled Smart Factories: A Case Study of Battery Module Assembly System for Electric Vehicles," *Journal of Manufacturing Systems*, vol. 63, pp. 214-223, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] NASA C-MAPSS-1 Turbofan Engine Degradation Dataset, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/bishals098/nasa-turbofan-engine-degradation-simulation

[30] Gas Sensor Array Drift Dataset, UC Irivine Machine Learning Repository. [Online]. Available: https://archive.ics.uci.edu/dataset/224/gas+sensor+array+drift+dataset

[31] Awesome Industrial Datasets, GitHub. [Online]. Available: https://github.com/jonathanwvd/awesome-industrial-datasets