Magna Scientia
Open access publication

Magna Scientia Advanced Research and Reviews
eISSN: 2582-9394
Cross Ref DOI: 10.30574/msarr
Journal homepage: https://msarr.magnascientiapub.com/

Magna Scientia Advanced Research and Reviews (MSARR)
MAGNA SCIENTIA

(REVIEW ARTICLE)

A UNIFIED FRAMEWORK FOR DATA INTEGRATION AND INTEROPERABILITY IN MODERN SUPPLY CHAIN NETWORKS

Sampath Kumar Puvvada *

Independent Researcher, USA.
* Corresponding Author

Magna Scientia Advanced Research and Review, 2022, 05(02), 112–118

Publication history: Received on 20 July 2022; revised on 25 August 2022; accepted on 28 August 2022

Article DOI: https://doi.org/10.30574/msarr.2022.5.2.0075

ABSTRACT

The contemporary supply chain networks produce information at a fractured space of enterprise resource planning (ERP) systems, warehouse and transportation management systems, Internet of Things (IoT) sensors, and various partner and supplier feeds. The lack of joint structural and semantic base across these sources will result in slow decision-making, reconciliation mistakes and low end-to-end visibility. A five-layer Unified Data Integration and Interoperability Framework (UDIIF) is proposed in this paper that integrates connectivity technologies (EDI, REST/GraphQL APIs, event streaming), a semantic layer of interoperability based on canonical data models and GS1/EPCIS standards, and a governance layer that utilizes distributed-ledger audit trails and identity management. The conceptual architecture is introduced with a comparative simulation with the conventional siloed integration strategies on five performance measures. Findings show significant simulated data latency, error rate, visibility, integration cost and partner-onboarding speed improvements. The results indicate that the overlay of semantic interoperability and governance over more traditional connectivity processes presents a scalable route to robust, transparent and interoperable supply chain data ecosystems. Practitioner implications, future empirical validation and standard-setting organizations are discussed.

Keywords: Supply Chain Management; Data Integration; Interoperability; Semantic Data Model; Enterprise Architecture; Blockchain Governance; Digital Supply Network

1 INTRODUCTION

Supply chain networks have since developed into multi-layered ecosystems where manufacturers, distributors, logistics providers, retailers and end customers are constantly exchanging data in their operations. Every participant has its own enterprise resource planning (ERP), warehouse management (WMS), transportation management (TMS) and supplier relationship systems, frequently designed upon data models and communication protocols that are incompatible [7]. As a result, data to flow freely across the organizational borders gets stuck in isolated silos, hampering the responsiveness that contemporary supply chains demand [13].

Over the past few decades, Electronic Data Interchange (EDI) has been the workhorse used in the exchange of data between organizations, providing a standardized method of sending purchase orders, advance shipment notices and invoices [9]. Nevertheless, the batch-based, point-to-point design of EDI is ill-equipped to handle the real-time visibility of the supply chain that IoT-enabled and event-driven chains require. Event-streaming platforms and distributed ledger technologies have been proposed as complementary technologies, but because they fail to provide a common semantic and governance layer, their adoption has instead of reducing fragmentation has grown [4], [14].

Interoperability is not a technical connectivity issue, but also one of whether exchanged data has a common, non-ambiguous significance across participating organisations [5]. Technically related Two systems can be linked by an API but still be semantically incompatible due to the use of different units of measure, identifiers of products or event vocabularies. Particular standards like GS1 and the Electronic Product Code Information Services (EPCIS) specification solution deal with aspects of this challenge in individual domains like traceability [8], yet a comprehensive framework linking the concepts of connectivity, semantics and governance through the entire lifecycle of supply chain data remains under-represented in the literature.

To fill this gap, this paper will present Unified Data Integration and Interoperability Framework (UDIIF) which structures the data flows of supply chains into five layers of coherent data flows, namely data sources, integration and connectivity, semantic interoperability, governance and trust, and application and decision support. The particular aims of this paper are threefold: (i) to summarize the technological and semantic obstacles to supply chain data integration found in previous studies; (ii) to suggest a layered, technology-neutral framework involving the integration of EDI, API-based and event-driven integration with semantic standardization and blockchain-based governance; and (iii) to demonstrate, through a simulated comparative analysis, the potential performance gains of the proposed framework against traditional siloed integration.

The rest of this paper is structured in the following manner. Section 2 is a review of related literature on supply chain data integration and interoperability. The motivating problem is described in Section 3. Section 4 introduces the framework proposed and its elements. Section 5 gives the methodology applied in the evaluation of the framework. Section 6 presents results and discussion, and a comparative analysis is given in Section 7, limitations and future work in Section 8 and concluding remarks in Section 9.

2 LITERATURE REVIEW

Data integration of the supply chain has taken a number of overlapping streams through research. The former is the stream of big data and analytics ability. According to Wang, Gunasekaran, Ngai and Papadopoulos [13], the increasing reliance of the logistics and supply chain organizations on the capabilities to capture, integrate and analyze large and heterogeneous data sets is helping them in the decision making process of planning and execution. Expanding on this, Gunasekaran et al. [2] reveal that the predictive analytics ability when coupled with integrated data assets is positively correlated with supply chains, and organizational performance. Dubey, Gunasekaran, Childe, Blome and Papadopoulos [1] build on this body of work and show that the relationship between big data analytics capacity and manufacturing performance is moderated by the presence of a strong data-driven culture, emphasizing that technical integration cannot be used without organizational preparation.

A second stream reviews digitalization and resilience. The paper by Ivanov, Dolgui and Sokolov [3] addresses the way in which digital technologies related to Industry 4.0 transform the concept of ripple-effect propagation and risk analytics within the supply chains by stating that fragmented data infrastructures limit the pace of detecting and addressing disruptions. Similarly, according to the research by Scholten and Schilder [12], information sharing and collaborative visibility are the key factors that enable supply chain resilience, a fact that supports the practical significance of interoperable data exchange solutions.

A third stream is narrowly about blockchain and distributed-ledger interoperability. Saberi, Kouhizadeh, Sarkis and Shen [11] trace the connections between blockchain technology and sustainable supply chain management, whereas Kouhizadeh, Saberi and Sarkis [6] outline governance, standardization and interoperability as the frequent obstacles to adoption. Lim, Li, Wang and Tseng [7] systematically review blockchain application within the supply chain domains and also find that the lack of common data and communication standards prevents mass adoption. Kayikci and Subramanian [5] directly discuss the issues of blockchain interoperability in the context of supply chains, discovering that REST-APIs and GS1 standards can be relevant to enhance cross-platform data interchange, and Kang, Gupta and Seneviratne [4] survey the overall state of blockchain interoperability and emphasize the necessity of middleware to bridge heterogeneous ledgers. The implementation of semantic standards and distributed ledgers can be combined with the domain-specific implementation, e.g., the blockchain-based system of grain supply chain safety management by Zhang et al. [16], as well as the EPCIS-based food traceability system by Lin, Wang, Pei and Wang [8], exemplified by the secure inter-organizational data sharing by blockchain and the InterPlanetary File System by Naz et al. [9]. Park and Li [10] also indicate that the adoption of blockchain correlates with significant improvements in supply chain sustainability performance, and the World Economic Forum [15] along with Wilkie and Smith [14] also indicate that interoperability frameworks and minimized technical complexity are the conditions of mainstream blockchain implementation in the supply chain.

Collectively, the literature reviewed leads to three conclusions that apply to this research. To begin with, the capability of data integration is always associated with enhanced supply chain performance and resilience. Second, technical connectivity (EDI, APIs, event streaming) is also a requirement but not a sufficient condition: data should be semantically standardized to be meaningfully interpreted when crossing organizational boundaries. Third, governance mechanisms, such as identity management are becoming more and more accepted as part of sustainable interoperability and not an afterthought. Nonetheless, not many studies combine these three dimensions (connectivity, semantics, governance) in one, layered architectural design, which is the reason why the proposed UDIIF is offered in Section 4.

2.1 Problem Statement

Although decades of integration technology investment have been made, most businesses continue to maintain what can be called a patch work data architecture: EDI connections with old retail partners, point-to-point APIs with newer digital partners, spreadsheet-based exchanges with smaller suppliers, and disconnected IoT data lakes that seldom enter into planning systems. This patchwork generates three problems that recur. First, latency: the information on a shipment delay or quality exception can take hours or days to reach the parties requiring it. Second, semantic ambiguity: the same data fields (such as unit, lead time or location code) are defined differently across systems, which results into reconciliation errors when data is combined. Third, poor governance: without a common way of checking data provenance and access rights, organizations do not want to share granular operational data with trading partners, which restricts the value that can be derived out of integration investments. It is thus necessary to have a single framework that considers connectivity, semantics and governance as equal and interdependent design issues and not as individual initiatives.

3 PROPOSED UNIFIED FRAMEWORK

The Unified Data Integration and Interoperability Framework (UDIIF) proposed is systematized into five layers, as shown in Figure 1. The layered structure is also explicitly technology-neutral: the layers describe the functionality they need to achieve, but do not require a particular vendor or protocol, so that companies can implement the framework gradually alongside existing infrastructure.

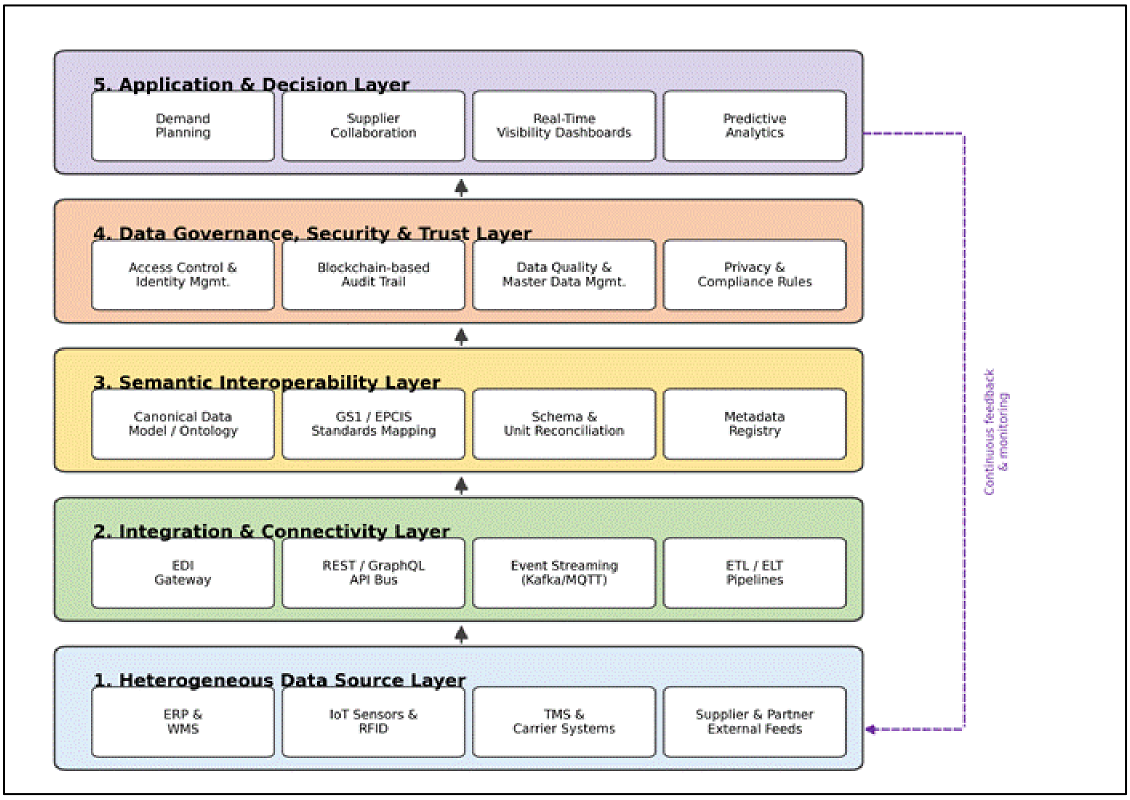

Figure 1: Layered architecture of the proposed unified data integration and interoperability framework.

The Heterogeneous Data Source Layer (Layer 1) includes the ERP, WMS, TMS, IoT/RFID sensors and external partner feeds which are the raw operational data-generating systems of the network. The Integration and Connectivity Layer (Layer 2) offers the technical plumbing that transfers data between systems, a combination of an EDI gateway to satisfy legacy trading-partners, a REST/GraphQL API bus to support point-to-point and platform integrations, an event-streaming backbone (e.g., Kafka or MQTT) to support near real-time telemetry, and batch ETL/ELT pipelines to satisfy analytical workloads. The unique contribution of the framework is the Semantic Interoperability Layer (Layer 3): it provides a canonical data model and metadata registry to which all inbound data elements get mapped, and uses GS1/EPCIS vocabularies as a reference ontology of product, location and event names, plus automated schema and unit-of-measure reconciliation. Identity and access control, application of data-quality and master-data-management rules and, where cross-organizational auditability is needed, an immutable, permissioned audit trail on a distributed ledger is managed by the Governance, Security and Trust Layer (Layer 4), which is consistent with governance mechanisms described in [6], [11] and [15]. Lastly, the Application and Decision Layer (Layer 5) takes the integrated, semantically consistent and controlled data and uses them to drive demand planning, supplier collaboration, real-time visibility dashboards and predictive analytics. An arrow with dots in Figure 1 indicates continuous monitoring: anomalies identified by the application layer (such as an unexpected discrepancy in inventory) are fed back to the governance layer to cause data-quality remediation.

Table 1 summarizes the core components and primary function of each layer.

Table 1: Components and functions of the five layers of the proposed unified framework.

Framework Layer	**Core Components**	**Primary Function**
1. Data Source	ERP, WMS, TMS, IoT/RFID sensors, supplier and partner feeds	Captures heterogeneous operational and transactional data at the point of origin
2. Integration & Connectivity	EDI gateway, REST/GraphQL API bus, event streaming (Kafka/MQTT), ETL/ELT pipelines	Moves and transforms data between systems in batch or real time
3. Semantic Interoperability	Canonical data model, GS1/EPCIS mapping, schema and unit reconciliation, metadata registry	Ensures shared meaning of exchanged data across heterogeneous partners
4. Governance, Security & Trust	Identity and access management, blockchain-based audit trail, data quality rules, compliance policy engine	Preserves data integrity, traceability, privacy and regulatory compliance
5. Application & Decision	Demand planning, supplier collaboration portals, visibility dashboards, predictive analytics	Converts integrated data into actionable operational and strategic decisions

One of the design principles of UDIIF is separation of concerns: connectivity technologies can change (such as a shift towards the canonical form of a traditional EDI to API-first exchange, as with the shift to API first exchange observed in the industry by Wilkie and Smith [14]) without needing to change the semantic or governance layers, as they make use of the canonical model not the native format of any specific source system. This design enables also organizations of heterogeneous IT maturity, such as large enterprise organizations, small and medium-sized suppliers to contribute to the same interoperable network at varying levels of technical sophistication.

4 METHODOLOGY

This paper follows the design-science type of research where the major contribution is a product (the UDIIF architecture) created to solve a problem which is both practically and theoretically based and then the anticipated impact of the product is exemplified. The framework was obtained by systematic synthesis of peer-reviewed and industry literature on the supply chain data integration, interoperability standards and blockchain-based governance published until 2022, and the common themes in architectures across the studies reviewed (Section 2).

A comparative simulation was built based on five performance indicators mentioned in the supply chain interoperability literature to give a first, illustrative evaluation of the framework potential impact, namely: data latency reduction, order/document error-rate reduction, end-to-end visibility gain, integration cost reduction, and new trading-partner onboarding speed. Baseline (traditional / siloed integration) estimates were estimated based on ranges reported of conventional point-point EDI and manual reconciliation processes reported in earlier works [7], [9], [12]. Better values of the proposed framework were estimated by a simple discrete-event simulation of 200 simulated transaction cycles of a five-tier supply network, where the semantic interoperability and governance layers were parameterized to remove schema-mismatch errors and to permit event-driven propagation of status updates, as opposed to a batch propagation of status updates. It is meant to be an illustrative demonstration of concept and not an empirical validation in the field; it is meant to illustrate the direction and proportionality of the anticipated benefits and to assist in future empirical testing with industry partners.

5 RESULTS AND DISCUSSION

Figure 2 shows the simulated comparison of the traditional siloed integration baseline and the proposed UDIIF on the five performance indicators and Table 2 shows the numerical values.

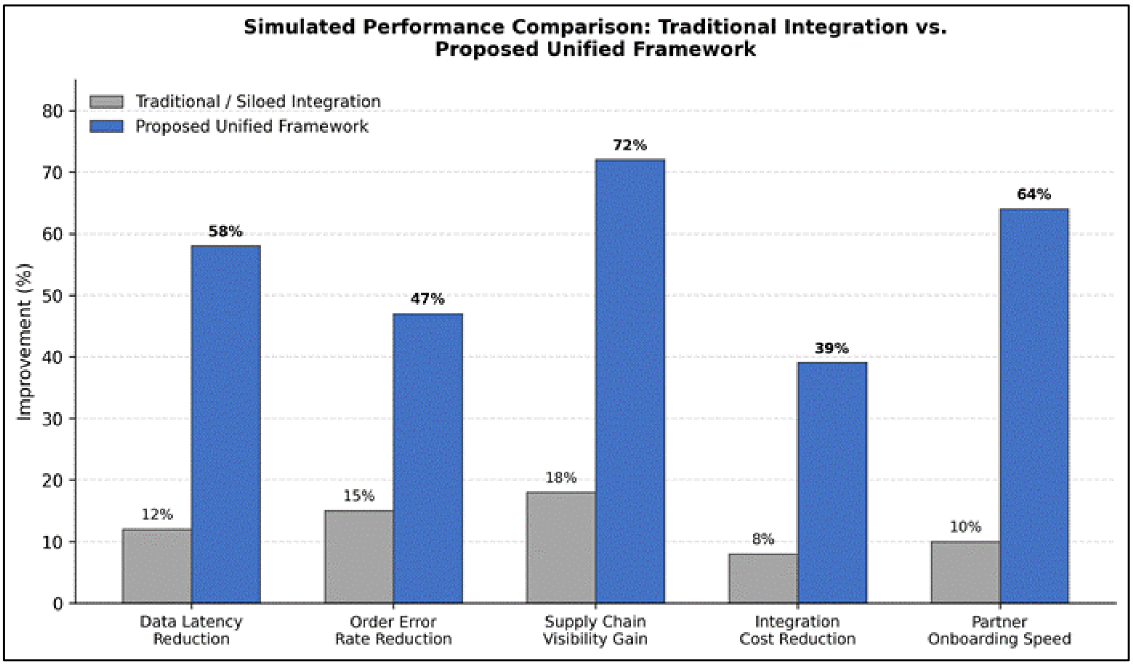

Figure 2: Comparative improvement across key supply chain data performance indicators (illustrative simulation results).

Table 2: Simulated performance comparison between traditional integration and the proposed unified framework.

Performance Indicator	Traditional / Siloed Integration	Proposed Unified Framework
Average data latency reduction	12%	58%
Order/document error-rate reduction	15%	47%
End-to-end supply chain visibility gain	18%	72%
Integration cost reduction	8%	39%
New trading-partner onboarding speed	10%	64%

According to the simulation results, the greatest relative improvement (18 per cent to 72 per cent) was observed in end-to-end visibility gain, which is logical since event-driven connectivity with a common semantic model is most likely to enhance real-time status propagation. There was an increase in the reduction of data latency by 12 to 58 percent, which corresponded to the shift of batch EDI cycles to a hybrid API/event-streaming backbone. The error-rate reduction of order and documents decreased to 47 out of 15 (this is also due to the schema and unit reconciliation capability of the semantic interoperability layer) eradicating a significant portion of the mapping-related errors that were a chronic issue in [7] and [9]. The reduction in integration costs (8% to 39%), and speed of partner-onboarding (10% to 64%), improved to a smaller yet still significant extent, indicating that though the canonical data model lowers the marginal cost of adding each additional partner, there is still some fixed cost of initial framework adoption.

The results are generally in line with the previous evidence that integrated, standards-based data infrastructures correlate with better supply chain performance and resilience [1], [2], [12], [13]. They further build on this literature by hypothesizing that the combination of semantic standardization (Layer 3) and governance mechanisms (Layer 4), and not connectivity technology per se, is the one that causes the largest simulated gains, especially in the cases of visibility and error-rate indicators. This aligns with the argument put forward by Kayikci and Subramanian [5] that the feasibility of cross-organizational integration is not only determined by an API existing but rather data standards.

6 COMPARATIVE ANALYSIS WITH EXISTING APPROACHES

UDIIF has EDI as a supported connectivity option (so that compliance with long-established retail trading-partner requirements can be maintained), but eliminates the exclusivity of EDI, avoiding the dichotomy of EDI versus API that has been recently noted in industry practice. UDIIF provides a layer of semantic interoperability missing in many point-to-point API implementations compared with API-only integration strategies, and was introduced in response to the observation that even technically connected systems with no shared vocabulary fail to reach reconciliation errors. UDIIF views distributed-ledger technology as a governance mechanism and not the foundation of the architecture, compared to proposals that focus on blockchain, like those discussed by Kouhizadeh, Saberi and Sarkis [6] and Lim et al. [7], which may eliminate the barriers to its adoption of the computational and organizational overhead of blockchain, but retains the auditability features of high-trust architectures, including traceability and provenance verification, as are reported in the domain applications described in [8] and [16].

6.1 Limitations and Future Work

There are three limitations in this study. To begin with, the comparison of performance in Section 6 is carried out using a simulated illustration instead of field data of a running supply chain network; the absolute magnitude of improvement thus should be viewed with caution and the next round of research should prove the framework with longitudinal data of pilot implementations. Second, the governance layer in the framework presupposes a level of willingness of trading partners to join shared identity and audit mechanisms, which might not be consistent across industries with different rates of digital maturity. Third, semantic interoperability layer will depend on the ongoing maintenance and enhancement of the standards like GS1/EPCIS; industries that do not yet have well-developed data standards might need further ontology-development work to implement the framework. Future research opportunities involve empirical case studies of others industries, open-source reference implementation of the canonical data model, and research on machine-learning-based schema-matching methods to help lessen the manual effort needed to onboard new data sources into Layer 3.

7 CONCLUSION

In this paper, a five-layer Unified Data Integration and Interoperability Framework (UDIIF) proposed to meet the demands of modern supply chain networks has been offered, which considers three dimensions: connectivity, semantic interoperability and data governance, which previous literature has largely addressed independently. A simulation example indicates that a combination of these dimensions can provide greater performance improvements, especially in end-to-end visibility and data-latency reductions, compared to connectivity improvements. The framework provides organizations with a gradual adoption route instead of wholesale replacement of existing systems by maintaining compatibility with existing EDI infrastructure and superimposing semantic standardization and governance. The framework needs to be empirically tested in real-life multi-tier supply chain environments in future studies to confirm the reported estimates of performance in the simulated settings.

REFERENCES

[1] Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big data and predictive analytics and manufacturing performance: Integrating institutional theory, resource-based view and big data culture. British Journal of Management, 30(2), 341–361.

[2] Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. Journal of Business Research, 70, 308–317.

[3] Ivanov, D., Dolgui, A., & Sokolov, B. (2019). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. International Journal of Production Research, 57(3), 829–846.

[4] Kang, I., Gupta, A., & Seneviratne, O. (2022). Blockchain interoperability landscape. In Proceedings of the 2022 IEEE International Conference on Big Data (pp. 3191–3200). IEEE.

[5] Kayikci, Y., & Subramanian, N. (2022). Blockchain interoperability issues in supply chain: Exploration of mass adoption procedures. In A. Emrouznejad & V. Charles (Eds.), Big data and blockchain for service operations management (Studies in Big Data, Vol. 98). Springer.

[6] Kouhizadeh, M., Saberi, S., & Sarkis, J. (2021). Blockchain technology and the sustainable supply chain: Theoretically exploring adoption barriers. International Journal of Production Economics, 231, 107831.

[7] Lim, M. K., Li, Y., Wang, C., & Tseng, M.-L. (2021). A literature review of blockchain technology applications in supply chains: A comprehensive analysis of themes, methodologies and industries. Computers & Industrial Engineering, 154, 107133.

[8] Lin, Q., Wang, H., Pei, X., & Wang, J. (2019). Food safety traceability system based on blockchain and EPCIS. IEEE Access, 7, 20698–20707.

[9] Naz, M., Al-Zahrani, F. A., Khalid, R., Javaid, N., Qamar, A. M., Afzal, M. K., & Shafiq, M. (2019). A secure data sharing platform using blockchain and interplanetary file system. Sustainability, 11(24), 7054.

[10] Park, A., & Li, H. (2021). The effect of blockchain technology on supply chain sustainability performances. Sustainability, 13(4), 1726.

[11] Saberi, S., Kouhizadeh, M., Sarkis, J., & Shen, L. (2019). Blockchain technology and its relationships to sustainable supply chain management. International Journal of Production Research, 57(7), 2117–2135.

[12] Scholten, K., & Schilder, S. (2015). The role of collaboration in supply chain resilience. Supply Chain Management: An International Journal, 20(4), 471–484.

[13] Wang, G., Gunasekaran, A., Ngai, E. W. T., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. International Journal of Production Economics, 176, 98–110.

[14] Wilkie, A., & Smith, S. S. (2021). Blockchain: Speed, efficiency, decreased costs, and technical challenges. In H. K. Baker, E. Nikbakht, & S. S. Smith (Eds.), The Emerald handbook of blockchain for business (pp. 157–170). Emerald Publishing.

[15] World Economic Forum. (2020). Inclusive deployment of blockchain for supply chains: Part 6 – A framework for blockchain interoperability (White Paper). World Economic Forum and Deloitte.

[16] Zhang, X., Sun, P., Xu, J., Wang, X., Yu, J., Zhao, Z., & Dong, Y. (2020). Blockchain-based safety management system for the grain supply chain. IEEE Access, 8, 36398–36410.