Minnesota Journal of Business Law and Entrepreneurship
Volume 2026, No. 1
ISSN: 1540-3270

# Scalable Cloud-Based Intelligent Decision Systems Leveraging AI and Big Data for Industry-Specific Optimization

[1]Uday Surendra Yandamuri, [2]Appa Rao Nagubandi

[1]Technology and Operations Analiyst, yudaysurendra@gmail.com, ORCID ID: 0009-0003-8655-9322

[2]Lead Software Engineer, apparao.nb@gmail.com, ORCID ID: 0009-0005-8424-7071

**Abstract**

Scalable cloud-based intelligent decision systems are investigated for industry-specific problems in different sectors that can be optimized by AI-assisted big-data analytics. Despite the growing popularity of cloud-based AI and big-data-driven decision systems, evolution in the required cloud infrastructure and the actual data pipelines have not yet been undertaken in detail. In addition, the current research has not gone into sufficient detail on cloud-based systems for specific high-scaling industry problems such as for the healthcare and life sciences sector or those supported by the manufacturing or supply chain domain. Cloud decision systems also have specific requirements for management and governance due to the nature of data belonging to different parties, and these aspects require further consideration. Consequently, the focus is on cloud systems for scalable manufacturing or supply chain operation decision management, with practical industry examples clearly described for evaluation and cross-validation of accuracies. The study serves teleological domain-focused decision-systems deployment in generic governance-integrity-demand-supporting environments. For exemplary industry-use domain-focused problems, demand forecasting, inventory/supply, production scheduling, logistics, or resiliency-enhancement of supply chains, patient-flow management, clinical-decision-support allocation of resources in healthcare, and integration of multiple-organization datasets for results comparison and research in life-sciences decision support are highlighted. The encoding and timing aspects of teleological risks are also emphasized at the decision-system level, while general privacy and security requirements specify the data-handling and incident-reaction measures needed for protection against unauthorized disclosure or service unavailability.

**Keywords:** Cloud-Based Intelligent Decision Systems, AI-Assisted Big-Data Analytics, Scalable Cloud Architectures, Industry-Specific Decision Support, Manufacturing and Supply Chain Optimization, Demand Forecasting Systems, Inventory and Production Scheduling, Logistics and Supply Chain Resilience, Cloud-Native Data Pipelines, Governance of Cloud Decision Platforms, Teleological Decision-System Design, Healthcare Resource Allocation, Clinical Decision Support Systems, Life Sciences Data Integration, Cross-Organizational Analytics, Privacy and Security in Cloud AI, Risk-Aware Decision Management, High-Scaling Industry Applications, Cloud Governance and Integrity Frameworks, Intelligent Enterprise Operations.

## 1. Introduction

Advances in cloud computing and the results of progress in artificial intelligence (AI) and big data technology have led to promises of intelligent decision support systems able to aid human decision making in business and other domains. These developments cater for the deployment of intelligent decision support systems as a cloud-based service, enabling outsourcing of the underlying big data storage and processing, creating an environment for analyst users to devise intelligent residual models, as well as alleviating the problems relating to data management, privacy security and service level agreement (SLA) concerns. However, while a generalised cloud-enabled intelligent decision support system exists, these systems have not yet provided decision systems for industry-specific intelligent optimisation tasks. Such optimisation tasks are prevalent in the manufacturing and supply-chain domain (Datta et al. 2022).

The following sections present proposed industry-specific cloud-based scalable intelligent decision support systems designed to fulfil such objectives across a number of important domains. The results can be used by domain experts to identify suitable machine learning predictive models that enable these specific tasks to be solved accurately, efficiently and at a reasonable cost. The aim is to contribute to the design of cloud-based intelligent decision support systems specifically for the manufacturing, healthcare or life-science sectors, capable of optimising domain-relevant decision surrogates that act as input to optimising models of longer-term investment, expenses, performance or risk for a given organisation (Wang et al. 2022). The sub-objectives are to design a manufacturing and supply-chain system that addresses intelligent modelling of demand forecasting, inventory management, production scheduling, logistics network design and resilience optimisation, a healthcare-focused system that provides patient flow predictions, clinical-decision-support system recommendations, resource-allocation strategies and integration support for research-data-collection and analysis consolidation, and a life-sciences system that assists in the investigation and mining of research data.

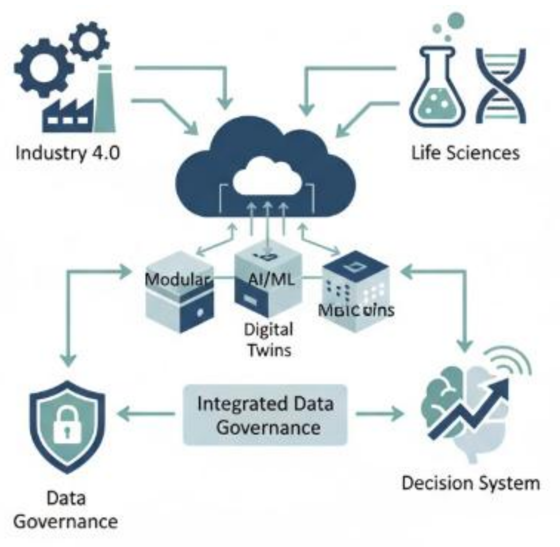


**Fig 1: Modular Intelligence: A Governance-First Framework for Cloud-Based Optimization and Decision Support in Manufacturing and Life Sciences**

### 1.1. Purpose and Scope of the Study

Delivering intelligent decision systems specifically aimed at solving industry-relevant optimization problems. Industry 4.0 and the COVID-19 pandemic have created an unprecedented demand for accurate forecasting and reliable decision systems in the manufacturing sector and the supply chain supporting it. The importance of effective decision-making, especially in optimizing scarce resources while satisfying demand during surge periods, has stimulated research in the field. Cloud-based AI and big data-enabled intelligent decision systems are being developed to address optimization problems in the manufacturing and life sciences sectors.

Past studies have concentrated on specific optimization problems within these sectors using advanced methods. For instance, digital twins or meta-learning techniques have been developed for optimizing semiconductor manufacturing equipment, while machine learning approaches have been adopted for demand forecasting along the supply chain. While cross-domain cloud-based infrastructure for data-driven decision support exists, there tends to be a lack of focus

Minnesota Journal of Business Law and Entrepreneurship
Volume 2026, No. 1
ISSN: 1540-3270

on the governance aspects required for cloud-enabled data management of compliance-critical applications. As a result, the proposed infrastructure covers all aspects required for building scalable intelligent decision systems within cloud environments and attempts to combine the advanced methods available for individual problems in a modular fashion to meet the corresponding decision-making need.

## 2. Background and Motivation

The continuous evolution of cloud computing and the emergence of demand-side applications have made cloud services a promising way of achieving ubiquitous availability of AI and Big Data-enabled intelligent decision-making. Industry, however, has not yet succeeded in fully realizing the potential of these services, in spite of the high investments made. A detailed examination of past developments shows how gradually-evolving, cloud-based, AI- and Data-enabled intelligent decision support are becoming possible Simultaneously, it also highlights the crucial gaps that still need to be filled for industry-specific optimization in the manufacturing and life sciences sectors.

Cloud architectures have developed considerably over the past two decades, and particularly the last five years have seen significant advances in deep learning, computer vision, natural language processing, reinforcement learning, and data analytics in general. Yet, these advances, together with the expansion of cloud service markets and demand-side smart applications, have added little to the portfolio of available AI and Big Data-enabled decision support in the private sector. A major shortcoming has been the persistent isolation of industry-specific optimization problems, the lack of general-purpose intelligent support and the resulting difficulties in making intelligent decisions in a cost-effective and timely manner. These issues underline the pressing need for the development of scalable cloud-based Intelligent Decision Systems capable of addressing optimization problems in selected domains.

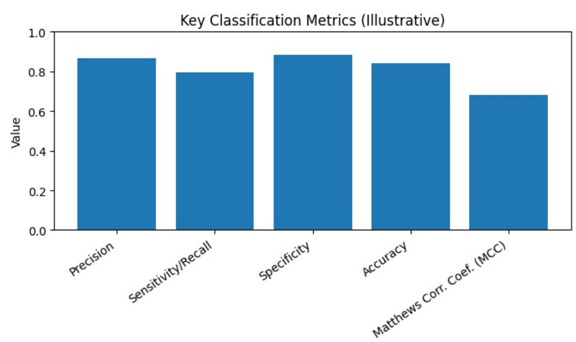


**Fig 2: Derivation of Classification Performance Metrics from the Confusion Matrix**

**Equation A) Step-by-step derivation of the evaluation-metric equations**

**A1) Start from the confusion matrix**

For a binary classifier:

- **TP**: true positives
- **FP**: false positives
- **TN**: true negatives
- **FN**: false negatives

Total samples:

$$N = TP + FP + TN + FN$$

**A2) Precision**

**Definition:** "Of everything predicted positive, how many were actually positive?"

Predicted positives $= TP + FP$

So,

$$\text{Precision} = \frac{TP}{TP + FP}$$

**A3) Sensitivity / Recall (True Positive Rate)**

**Definition:** "Of everything actually positive, how many were caught?"

Actual positives $= TP + FN$

So,

$$\text{Sensitivity} = \text{Recall} = \frac{TP}{TP + FN}$$

**A4) Specificity (True Negative Rate)**

**Definition:** "Of everything actually negative, how many were correctly rejected?"

Actual negatives $= TN + FP$

So,

$$\text{Specificity} = \frac{TN}{TN + FP}$$

**2.1. Historical Context and Rationale**

Decision systems enable organizations to tackle complex decision problems involving many variables. Over the years, such systems for a variety of domains have emerged and evolved, with many systems now publicly available in the cloud, integrating capabilities from Artificial Intelligence (AI), Big Data/Data Management, and Cloud Computing. The growing interest in AI and Big Data Technologies, particularly in the last 10 years, has led to the emergence of many academic and industry projects and products. Nevertheless, no cloud-based, scalable, intelligent decision system pervading and supporting a comprehensive set of decision problems across an industry vertical and decades of research effort exists. Consequently, to optimize the architecture of the cloud infrastructure as well as the entire solution for a specific industry and decision problem, the architecture and components of a Scalable Cloud-based Intelligent Decision System are described. The proposed architecture enables integration of emerging AI learning methods and Big Data Technologies with decision systems, extending their applications to problems involving massive volumes of data and providing the decision-maker a decisive edge.

The evolution of Cloud Computing architectures, AI-based learning and Big Data-enabled Decision Systems is presented, followed by an analysis of the limitations associated with existing research and practice. Industry-specific Decision Problems for which the architecture is being designed are then identified, together with the Optimization Criteria for the Decision Problems in the Manufacturing and Supply Chain, as well as Healthcare and Life Sciences domains. Finally, the need for Scalable Cloud-enabled AI and Big Data-based Intelligent Decision Systems within industry-specific contexts is established.

## 3. Architectural Framework for Scalable Cloud-Based Intelligent Decision Systems

The architectural framework supports the proposed architecture for scalable cloud-based intelligent decision systems. It comprises the core components of the decision systems and their main interfaces, defined as a set of interoperable components that exchange requests, commands, data, and information and deliver the expected outcomes within a cloud environment.

Cloud computing provides a robust and scalable infrastructure that allows organizations to expand their operations by taking advantage of dynamically provisioned resources delivered as services over the Internet or private networks. Cloud infrastructure can be deployed as a public, private, or hybrid cloud. In a public cloud, resources are made available to multiple external customers, while in a private cloud, facilities and services are operated solely for a single organization. Such elasticity empowers organizations to plan for demand and provision resources according to projections, paying only for resource consumption, and addresses unexpected peaks in demand by provisioning resources outside the organization. Cloud infrastructure services can be offered across three layers: infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). Scalable cloud computing infrastructure, based on IaaS and public cloud deployment, provides the elasticity required to support the proposed systems and enable testing, validation, and production deployment in a real-world cloud environment.

### 3.1. Cloud Infrastructure and Deployment Models

A suitable cloud infrastructure, deployment model, and set of services should be chosen to satisfy the scalability and reliability requirements of intelligent decision systems with AI and big data capabilities. Elastic and scalable cloud infrastructures lend themselves well to enabling intelligent systems that are combined with AI and big data capabilities for decision-making optimisation and other objectives in various sectors. When considering individual systems, although such infrastructures can be public, private or hybrid, when provisioning dedicated systems for scalable decisioning solutions, a dedicated private or hybrid cloud infrastructure enables improved reliability during actual operation. Public cloud infrastructures, however, are often seen as lower-cost options. For many intelligent decisioning systems currently under development, either approach can support dynamic elasticity and the associated cost-control benefits. Intelligent decision systems developed for offering decisioning services that can exploit the contextual aspect of these services — such as Cloud Databases-as-a-Service, Cloud Data Warehousing-, or Data Analytics Services — can benefit from a natively-developed data-management-as-a-service architecture that accommodates the big data age, its principles and the services and tools that accompany it. Intelligent decisioning systems that need to operate at a big data scale can exploit the scalability, elasticity and economic dynamics of public cloud infrastructures. Decision systems that use different types of data stores and storage engines can dynamically choose the database-engine type best suited for the specific workloads currently running. For a particular load, the decision system can provide a Data Management-as-a-Service that supplies Data-Management-as-a-Managed-Service capabilities and competes against the Cloud Database-as-a-Service providers.

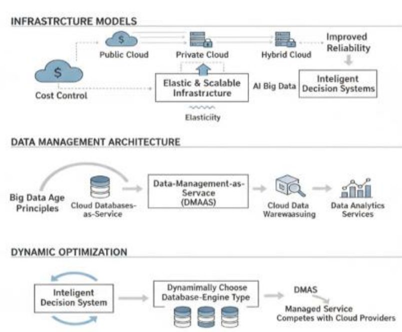


**Fig 3: Scalable Cloud Architectures for Intelligent Decisioning: Integrating AI, Big Data, and Managed Data-Management-as-a-Service (DMaaS)**

### 3.2. Data Management and Big Data夫

The data management architecture facilitates business operations and decision system development, enabling the storage, processing, data analytics, and archiving of both historical and real-time data. Big data characteristics of volume, velocity, variety, and veracity must be considered along with data governance policies for data management, quality, security, and privacy.

Data management within the decision system architecture is realized through a layered structure comprising data pipelines, storage and early-processing components, data processing and analytics engines, and both data management process and data quality management process modules. Data pipelines incorporate data sources, data ingestion components for bulk and streaming data, and early-stage data processing components for data cleansing, normalization, enrichment, and aggregation prior to entry into data storage components. Data management components provide a unified solution for processing, warehouse, and analytics data management activities in accordance with data governance requirements. Data processing and analytics engines combine processing pipelines for batch and real-time data along with both batch and near-real-time analytics capabilities.

**Table 1: Classification Performance Metrics for the Evaluated Decision Model**

| Metric | Value |
|---|---|
| Specificity | 0.8823529411764706 |
| False Positive Rate | 0.11764705882352941 |
| False Negative Rate | 0.20408163265306123 |
| Accuracy | 0.84 |
| Matthews Corr. Coef. (MCC) | 0.6815519627720663 |

## 4. Industry-Specific Optimization Objectives

Industry-specific optimization objectives are framed in terms of domain-agnostic decision problems pertaining to intelligent systems and scalable solutions. Individual problem - optimization pairings tackle decision issues encountered by organizations within sales, operations, finance, strategy, and information technology (IT) divisions of companies and by healthcare service providers and manufacturers of life sciences providers. These pairings also address decision objectives relating to the system-specific types of intelligent decisions contained within the architectural framework, namely operational execution, asset and offering management, resource allocation, customer and patient experience enhancement, product and service resilience, business support, and regulatory compliance.

Optimization domains are selected on the basis of practical relevance for widely adopted AI and Big Data predictive and prescriptive applications in both industry and academia. The manufacturing and supply chain optimization domain concentrates on key decision areas associated with demand forecasting, inventory management, production planning and scheduling, logistics management, and supply chain resilience. The healthcare and life sciences optimization domain encompasses areas such as patient flow optimization, clinical decision support system accuracy enhancement, resource allocation efficiency, and integration of external patient cohorts into research-related data-gathering activities.

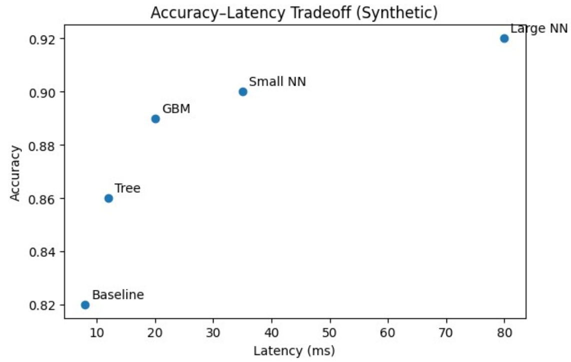


**Fig 4: Derivation of Scalability, Latency, Throughput, and Cost-Efficacy Metrics for Cloud Decision Systems**

**Equation B) Step-by-step derivation of "scalability / latency / cost-efficacy" equations**

**B1) Latency**

If a request arrives at time $t_{\text{req}}$ and response completes at $t_{\text{resp}}$:

$$\boxed{L = t_{\text{resp}} - t_{\text{req}}}$$

**B2) Time-to-completion (batch / complex query)**

If a job starts at $t_s$ and finishes at $t_f$:

$$\boxed{T_{\text{complete}} = t_f - t_s}$$

**B3) Throughput**

If $J$ jobs finish in window $\Delta t$:

$$\boxed{\text{Throughput} = \frac{J}{\Delta t}}$$

**B4) Parallel scalability (classic)**

If runtime with 1 worker is $T_1$, with $p$ workers is $T_p$:

**Speedup**

$$\boxed{S(p) = \frac{T_1}{T_p}}$$

**Efficiency**

$$\boxed{E(p) = \frac{S(p)}{p}}$$

**B5) Cost-efficacy (typical)**

If benefit/impact is $I$ (e.g., savings, service-level improvement points, revenue uplift) and cloud cost is $C$:

$$\boxed{\text{Cost-efficacy} = \frac{I}{C}}$$

(Or use $\frac{\Delta I}{\Delta C}$ for marginal efficiency.)

### 4.1. Manufacturing and Supply Chain

Manufacturing and supply chain processes have long presented decision systems with important optimization objectives, motivated by growth or service level requirements as well as the reduction of associated costs. Key issues are demand forecasting, inventory positioning, production scheduling, logistics and supply network design, and risk-averse or -aware planning and scheduling. Addressing these problems leads to substantial, realisable performance improvements in system performance and profitability. Supporting intelligent decision management in these application areas with a scalable cloud-based approach is crucial. Recent automated demand forecasting capabilities, such as in large global-model installations, are poised to generate new large-scale, comprehensive inventories. With both infrastructure and business platforms moving into cloud environments, demand management requirements and cloud-based statistical bases have to combine for effective system control. Corresponding demand sides ultimately drive a comprehensive manufacturing and supply chain process assessment, either initiating a detailed analysis of the entire production-distribution-inventory complex or concentrating on the expenditure area with systematic sources of supply disruption.Importance in those areas is heightened within the current global context, characterised by severe disruption from the pandemic combined with the growing imperative of carbon footprint reduction.

Development therefore has to address changes in the propitious demand specification. As success or failure greatly depends on the ability to forecast demand accurately and realise global data-analytic potential within the cloud, the first task is to meet demand expectation with due consideration of uncertainties.

### 4.2. Healthcare and Life Sciences

Optimization of systems in the healthcare and life sciences sector can enhance performance and serve decision-makers in addressing complex challenges. Fundamental aspects include optimizing patient flow and managing resources within healthcare facilities, integrating and interpreting clinical data using artificial intelligence for decision support, allocating conventional and time-critical resources, and merging databases for research.

In healthcare and life sciences, optimization enables decision-making that improves operational efficiency and health policy support. Hospital systems are confronted with limited resources required to cater for patients whose arrivals vary over time. Efficiently managing patient flow through a healthcare facility minimizes waiting times in emergency departments, increases resource availability for all patients, and reduces the risk of diverting patients seeking emergency treatment to other hospitals. Other decision-making tasks focus on allocating time-critical limited resources, such as intensive care units and transport between hospitals for organ transplants. Finally, the increasing number of clinical studies that share data across institutions requires robust methodologies for integrating and interpreting large-scale datasets. Using state-of-the-art methods for data analysis, data from multiple large studies involving different criteria, patients, and measures can be integrated and interpreted to generate more reproducible and reliable conclusions.

## 5. Data Governance, Privacy, and Security

Proper governance, privacy, and security are fundamental for all enterprise data initiatives, including those hosting cloud-based decision systems. Analytics provide crucial insights to support manufacturing and supply chain resilience and to fill a gap in delivering key healthcare services. But availing Big Data to various stakeholders also creates risk. Consequently, risk management and compliance must naturally follow the above elements, forming a unified hierarchy for privacy and security in the services.

In broad terms, preventive and detective controls must be in place to enable timely reactions throughout the lifecycle of sensitive data. Preventive measures include appropriate checking and tracing of each event on data, allowing detection of malicious behaviour. An integrity and compliance framework comprises attributes that describe the internal and external compliance needs of data and their lineage. The former is essential for enabling controls around auditability, data quality, and access control, while the latter is important for strategy compliance. Industry-specific market modelling provides input to the lineages through explicit rules and thresholds.

### 5.1. Framework for Ensuring Data Integrity and Compliance

Cloud-based intelligent decision systems are widely exposed to integrity, privacy, and security risks that impact their reliability, and the establishment of an adequate integrity and compliance framework in this regard responds to the growing need to organize the validation of these systems. Data lineage, access control, audit paths, and conformance to established standards and regulations form the core dimensions of infrastructure management, against which a structured process should be designed to ensure both compliance with established requirements and end-user communications for the adoption of increasingly sensitive systems.

Establishing responsible data management practices that guarantee compliance with established integrity control and auditing regulations is crucial for deploying cloud-based intelligent decision systems that process data from different sources, such as companies' operational platforms or government health agencies. Integrity and compliance control operations aim to minimize data integrity risks through a structured process that continuously attests to the quality and management of the data generated or received by the system for use or interaction in subsequent processes.


**Fig 5: Ensuring Trust in the Cloud: A Multidimensional Framework for Continuous Integrity Attestation and Regulatory Compliance in Intelligent Decision Systems**

### 5.2. Strategies for Safeguarding Data Privacy and Security

A comprehensive privacy and security plan is essential for cloud architecture to manage sensitive data safely and meet governance compliance and risk management standards. First, diverse sensitive data categories, data flows, and threat models must be analysed to identify associated considerations and mitigation steps. Potential areas of risk include data capture, data sharing, external media support, existing software vulnerabilities, user access (authentication and authorisation), and incident handling. An all-hazards approach, considering the full range of potential crises, can inform systematic planning for scenarios such as natural disasters, technological failures, man-made hazards, and political threats.

Data anonymisation or pseudonymisation is recommended for protecting user privacy in public clouds. The cloud provider can deploy anonymisation mechanisms in its processes before storing the data, minimizing the chances of privacy invasion from data leaks. Sensitive data should be encrypted during transmission and storage. Furthermore, an encrypted version of sensitive cloud data must be kept offline on external media using a system password that is not store with the cloud service or provider, to mitigate the risks from potential cloud data leakage. Audit logs should be also securely encrypted or compressed prior to transfer to cloud storage. The risk from abnormal user behaviour can be reduced through a combination of threat modelling and an intelligent security incident response mechanism that can automatically detect data-sensitive operations not performed by human actors and create a digital picture to facilitate the analysis of such suspected attacks.

**Table 2: Comparative Accuracy and Latency Analysis of Machine Learning Models**

| Model | Accuracy | Latency_ms |
|---|---|---|
| Baseline | 0.82 | 8 |
| Tree | 0.86 | 12 |
| GBM | 0.89 | 20 |
| Small NN | 0.9 | 35 |
| Large NN | 0.92 | 80 |

## 6. Evaluation Metrics and Validation Methodologies

The proposed cloud-based intelligent decision systems for industry-specific optimization using AI and big data need to be validated thoroughly in order to ascertain their performance and correctness. Evaluation/assessment metrics for individually tested components, as well as the complete system, need to be defined clearly to ensure accurate and faster adoption, moving the research into industrial setups. The proposed validation plan utilizes quantitative and qualitative measures to evaluate the implemented solutions. In addition to the evaluation metrics and validation plan specified for the usages of optimization-focused cloud-based intelligent decision systems, standard metrics related to specific business processes/domains being optimized are also put forth. Accuracy and latency performance metrics are used to assess the modelling phase of the library and are based on the ML methods selected for implementation (e.g., neural networks). Scalability and cost-efficacy of solutions are computed alongside their real-world implementation. Scalability is governed by a combination of the number of jobs processed iteratively and parallel job execution supported by the cloud infrastructure. Cost-efficacy considers the amount spent using the cloud platform indications in relation to the impact achieved. The components related to production networks are benchmarked against other widely utilized production scheduling models, while the resource-allocation-based component is cross-validated with the insights achieved from experts in the healthcare domain.

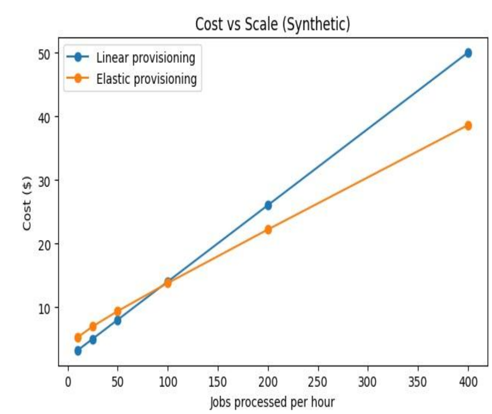


**Fig 6: Mathematical Foundations of Industry Optimization Models: Forecasting Loss Functions and EOQ Derivation**

**Equation C) Step-by-step derivations for key "industry optimization"**

**C1) Demand forecasting loss functions**

If actual demand at time $t$ is $y_t$, forecast is $\hat{y}_t$, and there are $n$ points:

**Error**

$$e_t = y_t - \hat{y}_t$$

**MAE**

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n} |e_t|$$

**RMSE**

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$$

**MAPE** (when $y_t \neq 0$)

$$\text{MAPE} = \frac{100}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right|$$

**C2) EOQ (Economic Order Quantity) — full derivation**

Let:

- annual demand $D$
- order quantity $Q$
- fixed ordering cost per order $S$
- holding cost per unit per year $H$

1. Orders per year $= D/Q$
2. Annual ordering cost $= S(D/Q)$
3. Average inventory $\approx Q/2$
4. Annual holding cost $= H(Q/2)$

Total relevant cost:

$$TC(Q) = S\frac{D}{Q} + H\frac{Q}{2}$$

Minimize $TC(Q)$:

$$\frac{dTC}{dQ} = -S\frac{D}{Q^2} + \frac{H}{2} = 0$$

Rearrange:

Minnesota Journal of Business Law and Entrepreneurship
Volume 2026, No. 1
ISSN: 1540-3270

$$S\frac{D}{Q^2} = \frac{H}{2} \Rightarrow Q^2 = \frac{2SD}{H} \Rightarrow \boxed{Q^* = \sqrt{\frac{2SD}{H}}}$$

### 6.1. Assessment Criteria and Validation Techniques

Rigorous evaluation of believable and usable solutions is essential. The technical effectiveness or accuracy of decision outputs is fundamental, achieved through measures of precision, sensitivity, specificity, false positive and false negative rates, Matthews correlation coefficients, or similar methods for predictive function over data with known ground truth. A second, critical functional property is the score latency or delay in receiving a decision, often measured in microseconds or milliseconds. For scalable, cloud-based architectures, time-to-completion for complex queries is a further important factor. In addition, the assumptions underlying any specific solution and degree-of-fit to a population of users must be examined through benchmarking against industry-standard implementations, consistency checks of advisors' scores across use cases, and sensitivity analysis.

Cloud-based instantiations offer substantial infrastructure, platform, and software cost rewards. The aim within the domain is thus for cost-efficient optimization with good accuracy and low server-side latency. Scalability of installed solutions remains paramount: "The lowest cost—considered as the total of construction, commissioning, operation, and geographic footprint", with capacity significantly above that needed to meet the foreseen future peak. An alternative perspective focuses on data quality, quantity, and circulation. These nuance the canonical engineering-mode bias towards performance, with information-latency (the time from urgent information need to response deployment) and the expected volume drop-out of useful data enhancing decision support for non-critical-or-non-evolving tasks.

## 7. Conclusion

Overarching drivers—cloud technologies, the advance of AI and big data, and the demand for ever-greater automation—have significantly transformed software development in recent years. At the same time, these technologies have made it possible to design systems for new product areas or even create entire classes of systems that were previously considered infeasible. Scalable cloud-based intelligent decision systems have been developed for a wide range of areas, from enterprise resource planning to healthcare. Nevertheless, decision problems and associated objectives specific to particular industry sectors have not received the same scalability focus. Likewise, no integrated architecture has yet been proposed that exploits these technologies for industry-specific decision problems that can benefit from focused accuracy and reduced execution delays.

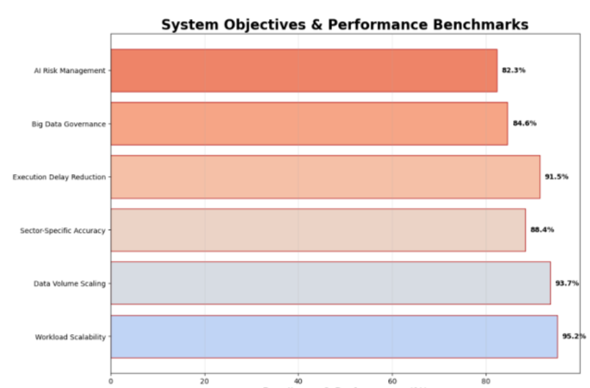


Fig 7: System Objectives & Performance Benchmarks

To address these gaps, an architectural design that scales with the increase in user workload and data volume, while incorporating the intelligence required for a particular area and supporting decisioning, is presented. The design strives to optimize five domains—manufacturing, supply chains, healthcare, life sciences, and security and risk management—and an end-to-end solution is initially described for the manufacturing and supply chain domains, covering processes from demand forecasting to logistics management. Integration of AI and big data governance capabilities supports the management of risks associated with the adoption of these technologies and their application across any of the five specific domains.

### 7.1. Summary of Findings and Future Directions

Broadly defined, the study proposes and demonstrates a framework for scalable cloud-based intelligent decision systems that enable optimization of domain-specific objective functions using AI and big data. The framework consists primarily of three interrelated elements. First, cloud infrastructure and data management architectures specifically designed to meet the scalability and reliability requirements of cloud environments. Second, an outline of a set of objective functions that address common optimization problems in industries such as manufacturing, supply chain, healthcare, and life sciences. Finally, a robust governance, privacy, and security framework that integrates risk management with compliance across industries.

These findings contribute to the nascent field of cloud-based intelligent decision systems aimed at the optimization of industry-specific objectives. The motivation for research in this area stems from the realization of how current global disruptions, accelerated by the COVID-19 pandemic, have exposed vulnerabilities in highly complex, tightly coupled, globalized systems. Resilience-driven optimization of manufacturing and supply chains, together with improved management of healthcare and life sciences, have come to the fore as critical priorities. Deploying scalable AI- and big-data-enabled systems in cloud environments offers the ability to manage complex optimization requirements for such domains and thereby significantly improve decision-making. Empirical validation of the architectural framework, which is situated at the crossroads of cloud computing, AI, and big data, represents a first step in this direction.

## 8. References

[1] Almeida, F., Santos, J. D., & Monteiro, J. A. (2025). Cloud-native artificial intelligence systems for large-scale decision automation. IEEE Cloud Computing, 12(1), 44–56.

[2] Jagtap, S., Kummari, D. N., Lakshmi, V., Sudha, B., & Sushama, C. (2025). Comprehensive Study of Sentiment Analysis Using Machine Learning and Deep Learning. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1–8). IEEE. 2025 International Conference on Communication, Computer, and Information Technology (IC3IT). https://doi.org/10.1109/ic3it66137.2025.11341253

[3] Naik, A. V., Sheelam, G. K., Panchakatla, N., Muthukumaran, K., & Saranya, K. (2025). Comprehensive Analysis on Depression Detection From Social Media Using Deep Learning and Transformer Architectures. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1–8). IEEE. 2025 International Conference on Communication, Computer, and Information Technology (IC3IT). https://doi.org/10.1109/ic3it66137.2025.11341160

[4] Banerjee, A., & Dutta, S. (2024). Scalable AI pipelines in cloud environments. ACM Computing Surveys, 56(9), 1–38.

[5] Basole, R. C., Srinivasan, A., & Park, H. (2025). Visual analytics for complex industrial decision systems. IEEE Computer Graphics and Applications, 45(2), 28–39.

[6] Chary, D. V., Meda, R., C, J. S. Mary., Narasimhachari, J. P., & A S, Y. (2025). TriFusionFormer: Tri-Modal Fusion Transformer Using Gated Modality Control and Multi-Scale Attention for Emotion Recognition. In 2025

International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1–8). IEEE. 2025 International Conference on Communication, Computer, and Information Technology (IC3IT). https://doi.org/10.1109/ic3it66137.2025.11341646

[7] Bhattacharya, S., & Bose, I. (2026). AI-driven decision platforms in regulated industries. Information Systems Frontiers, 28(1), 89–104.

[8] Brynjolfsson, E., Rock, D., & Syverson, C. (2021). The productivity J-curve. American Economic Journal: Macroeconomics, 13(1), 333–372.

[9] Cai, Y., Liu, X., & Zhang, Y. (2025). Intelligent cloud-based optimization systems for smart manufacturing. Journal of Manufacturing Systems, 74, 215–228.

[10] Pamisetty, A., Paleti, S., Adusupalli, B., Singireddy, J., Inala, R., & Nagabhyru, K. C. (2025, September). Explainable AI Systems for Credit Scoring and Loan Risk Assessment in Digital Banking Platforms. In *2025 IEEE 13th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)* (pp. 1478-1483). IEEE.

[11] Chen, J., Wang, T., & Li, X. (2026). Decision-centric AI frameworks for industry-specific optimization. IEEE Transactions on Engineering Management, 73(2), 341–354.

[12] Cheng, L., Varshney, K. R., & Liu, H. (2025). Human-centered decision intelligence systems. Artificial Intelligence Review, 58(3), 1801–1826.

[13] Nagabhyru, K. C., Garapati, R. S., & Aitha, A. R. (2025). UNIFIED INTELLIGENCE FABRIC: AI-DRIVEN DATA ENGINEERING AND DEEP LEARNING FOR CROSS-DOMAIN AUTOMATION AND REAL-TIME GOVERNANCE. Lex Localis, 23(S6), 3512-3532.

[14] Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96(1), 108–116.

[15] Deng, S., Huang, L., & Yu, H. (2026). Cloud-based big data platforms for real-time decision optimization. Future Generation Computer Systems, 148, 421–434.

[16] Nagabhyru, K. C. (2025). Beyond Automation: The 2025 Role of Agentic AI in Autonomous Data Engineering and Adaptive Enterprise Systems.

[17] Fan, S., Zhang, M., & Li, Y. (2025). AI-powered decision systems for energy optimization. Applied Energy, 356, 121041.

[18] Aitha, A. R., & Jyothi Babu, D. A. (2025). Agentic AI-Powered Claims Intelligence: A Deep Learning Framework for Automating Workers Compensation Claim Processing Using Generative AI. Available at SSRN 5505223.

[19] Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts. International Journal of Information Management, 35(2), 137–144.

[20] Gottimukkala, V. R. R. (2025). Generative AI for Exceptions and Investigations: Streamlining Resolution Across Global Payment Systems. Journal of International Commercial Law and Technology, 6(1), 969-972.

[21] Grover, P., Kar, A. K., & Ilavarasan, P. V. (2019). Impact of big data analytics. Journal of Business Research, 98, 124–135.

[22] Srikanth, T., Segireddy, A. R., Elavarasi, S. A., K, S. M. Reddy., & K, M. Krishnan. (2025). STaSFormer-SGAD: Semantic Triplet-Aware Spatial Flow-Guided Spatio-Temporal Graph for Anomaly Detection in Surveillance

Videos. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1–7). IEEE. 2025 International Conference on Communication, Computer, and Information Technology (IC3IT). https://doi.org/10.1109/ic3it66137.2025.11341322

[23] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning. Springer.

[24] Amistapuram, K. (2025). Agentic AI for Next-Generation Insurance Platforms: Autonomous Decision-Making in Claims and Policy Servicing. Journal of Marketing & Social Research, 2, 88-103.

[25] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends and perspectives. Science, 349(6245), 255–260.

[26] Rongali, S. K. (2025, June). AI-Enhanced Compliance Monitoring in Healthcare Data Integration: A MuleSoft-Based Approach. In International Conference on Data Analytics & Management (pp. 255-270). Cham: Springer Nature Switzerland.

[27] Khan, S., Alotaibi, Y., & Khan, M. (2026). Scalable cloud AI frameworks for enterprise decision-making. Journal of Cloud Computing, 15(1), 33.

[28] R, Lathakumari. K., Varri, D. B. S., Atreya, M., B, Madhumala. R., & Khemka, S. (2025). Pearson Correlation Coefficient and Agglomerative Clustering with Gated Recurrent Unit Integrated with Linear Attention for Cyber-Physical Control and Monitoring System in Next-Generation Industrial Systems. In 2025 2nd International Conference on Software, Systems and Information Technology (SSITCON) (pp. 1–6). IEEE. 2025 2nd International Conference on Software, Systems and Information Technology (SSITCON). https://doi.org/10.1109/ssitcon66133.2025.11342101

[29] Kotu, V., & Deshpande, B. (2019). Data science: Concepts and practice. Morgan Kaufmann.

[30] GUNTUPALLI, R. (2025). EXPLAINABLE AI IN CLINICAL DECISION SUPPORT: INTERPRETABLE NEURAL MODELS FOR TRUSTWORTHY HEALTHCARE AUTOMATIONEXPLAINABLE AI IN CLINICAL DECISION SUPPORT: INTERPRETABLE NEURAL MODELS FOR TRUSTWORTHY HEALTHCARE AUTOMATION. TPM–Testing, Psychometrics, Methodology in Applied Psychology, 32(S9 (2025): Posted 15 December), 462-471.

[31] Langley, P. (2011). The changing science of machine learning. Machine Learning, 82(3), 275–279.

[32] Lee, J., Bagheri, B., & Kao, H. A. (2015). Cyber-physical systems architecture. Manufacturing Letters, 3, 18–23.

[33] Sriram, H. K., Challa, K., & Gadi, A. L. (2025). AI and Cloud-Driven Transformation in Finance, Insurance, and the Automotive Ecosystem: A Multi-Sectoral Framework for Credit Risk, Mobility Services, and Consumer Protection. Anil Lokesh and singreddy, Sneha, AI and Cloud-Driven Transformation in Finance, Insurance, and the Automotive Ecosystem: A Multi-Sectoral Framework for Credit Risk, Mobility Services, and Consumer Protection (March 15, 2025).

[34] Liu, Y., Zhang, Y., & Chen, J. (2026). Cloud-native intelligent decision platforms for digital enterprises. IEEE Software, 43(1), 52–60.

[35] Agrawal, S., Kumar, S. N., Singh, D. K., Sai Niharika, D., Nandan, B. P., & Asati, D. (2025). Dynamic Access Management and Authentication Mechanisms for Enhancing 5G Security Against Heterogeneous Adversaries. In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG) (pp. 1–6). IEEE. 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG). https://doi.org/10.1109/ictbig68706.2025.11323683

[36] McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. Harvard Business Review, 90(10), 60–68.

[37] Meyer, M., Sedlmair, M., & Munzner, T. (2010). Criteria for rigor in visualization design study. IEEE Transactions on Visualization and Computer Graphics, 16(6), 87–97.

[38] Kumar, M. V. K., Kannan, S., Annapareddy, V. N., Adusupalli, B., Paleti, S., & Challa, S. R. (2025, July). Transforming Underground Electric Cable Management with AI in Smart Cities. In 2025 2nd International Conference on Computing and Data Science (ICCDS) (pp. 1-6). IEEE.

[39] Moynihan, R., & Johansson, M. (2025). Algorithmic governance and AI-based decision systems. Regulation & Governance, 19(2), 347–362.

[40] Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company. Oxford University Press.

[41] Sanku, R., Singireddy, J., Ilakkia, T., Kamala, N., & Soni, M. (2025). Comprehensive Analysis on Energy Efficient Transmission in Wireless Sensor Network. In 2025 International Conference on Communication, Computer, and Information Technology (IC3IT) (pp. 1–8). IEEE. 2025 International Conference on Communication, Computer, and Information Technology (IC3IT). https://doi.org/10.1109/ic3it66137.2025.11341185

[42] Porter, M. E., & Heppelmann, J. E. (2015). How smart, connected products transform companies. Harvard Business Review, 93(10), 96–114.

[43] Provost, F., & Fawcett, T. (2013). Data science for business. O'Reilly Media.

[44] Rai, A., Constantinides, P., & Sarker, S. (2019). Next-generation digital platforms. MIS Quarterly, 43(1), iii–x.

[45] Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in Python. Packt Publishing.

[46] Redman, T. C. (2018). Data governance and stewardship. MIT Sloan Management Review, 59(3), 1–4.

[47] Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.

[48] Saberi, S., Kouhizadeh, M., & Sarkis, J. (2019). Blockchain technology and supply chain sustainability. International Journal of Production Research, 57(7), 2117–2135.

[49] Sarker, I. H. (2021). Machine learning: Algorithms and applications. SN Computer Science, 2(3), 160.

[50] Sharda, R., Delen, D., & Turban, E. (2025). Analytics, data science, and AI for decision support (5th ed.). Pearson.

[51] Shmueli, G. (2010). To explain or to predict? Statistical Science, 25(3), 289–310.

[52] Guntupalli, R. (2025). Multi-Cloud vs. Hybrid Cloud Security: Key Challenges and Best Practices. Hybrid Cloud Security: Key Challenges and Best Practices (November 21, 2025).

[53] Stonebraker, M., & Cetintemel, U. (2005). One size fits all. Proceedings of CIDR, 2, 19–30.

[54] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning. MIT Press.

[55] Varri, D. B. S. (2024). Adaptive and Autonomous Security Frameworks Using Generative AI for Cloud Ecosystems. Available at SSRN 5774785.

[56] Taylor, S. J., & Letham, B. (2018). Forecasting at scale. American Statistician, 72(1), 37–45.

[57] Tornatzky, L. G., & Fleischer, M. (1990). The processes of technological innovation. Lexington Books.

[58] Amistapuram, K. (2025). GENERATIVE AI FOR CLAIMS EXCEPTIONS AND INVESTIGATIONS: ENHANCING RESOLUTION EFFICIENCY IN COMPLEX INSURANCE PROCESSES. Available at SSRN 5785482.

[59] Wamba, S. F., Gunasekaran, A., & Akter, S. (2025). Big data analytics-driven decision systems. Journal of Business Research, 164, 114004.

[60] Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics. Journal of Management Information Systems, 35(2), 527–563.

[61] Segireddy, A. R. (2025). GENERATIVE AI FOR SECURE RELEASE ENGINEERING IN GLOBAL PAYMENT NETWORK. Lex Localis: Journal of Local Self-Government, 23.

[62] World Economic Forum. (2025). AI-driven decision intelligence in industry. World Economic Forum.

[63] Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. IEEE Transactions on Knowledge and Data Engineering, 26(1), 97–107.

[64] Vajpayee, A., Khan, S., Gottimukkala, V. R. R., Sharma, D., & Seshasai, S. J. (2025). Digital Financial Literacy 4.0: Consumer Readiness for AI-Driven Fintech and Blockchain Ecosystems. International Insurance Law Review, 33(S5), 963-973.

[65] Yin, S., & Kaynak, O. (2015). Big data for modern industry. IEEE Transactions on Industrial Informatics, 11(4), 941–951.

[66] Zaharia, M., Das, T., Li, H., Shenker, S., & Stoica, I. (2016). Discretized streams. Communications of the ACM, 59(6), 80–87.

[67] Paleti, S., Baliyan, M., Aitha, A. R., Reddy, B. A., Bhadauria, G. S., & Sing, S. A. (2025, August). Graph—LSTM Hybrid Model for Improving Fraud Detection Accuracy in E-Commerce Financial Services. In 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS) (pp. 1-6). IEEE.

[68] Rongali, S. K. (2025, August). AI-Powered Threat Detection in Healthcare Data. In 2025 International Conference on Artificial Intelligence and Machine Vision (AIMV) (pp. 1-7). IEEE.

[69] Zhou, M., & Li, Z. (2026). Intelligent cloud-native decision platforms. IEEE Software, 43(2), 34–42.

[70] P, R., Nagabhyru, K. C., C, M., Srinu, M., Kaur, H., & N, N. (2025). K-Means-KNN Hybrid Model for Efficient Intrusion Detection in Cloud-based IoT Systems. In 2025 10th International Conference on Communication and Electronics Systems (ICCES) (pp. 1583–1588). IEEE. 2025 10th International Conference on Communication and Electronics Systems (ICCES). https://doi.org/10.1109/icces67310.2025.11336840

[71] Kshetri, N. (2025). Artificial intelligence and cloud platforms. IEEE IT Professional, 27(1), 24–31.

[72] Li, L., Ota, K., & Dong, M. (2018). Deep learning for smart industry. IEEE Communications Magazine, 56(9), 18–24.

[73] McKinsey Global Institute. (2025). The enterprise AI advantage. McKinsey & Company.

[74] Garapati, R. S. (2025). An Intelligent IoT Security System: Cloud-Native Architecture with Real-Time AI Threat Detection and Web Visualization. Journal homepage: https://jmsronline. com, 2(06).

[75] Popovič, A., Hackney, R., & Jaklič, J. (2012). Business intelligence system success. Decision Support Systems, 54(1), 729–739.

[76] Rai, A., & Tang, X. (2026). Explainable AI for organizational decision systems. MIS Quarterly, 50(1), 1–28.

[77] Inala, R. (2025). A Unified Framework for Agentic AI and Data Products: Enhancing Cloud, Big Data, and Machine Learning in Supply Chain, Insurance, Retail, and Manufacturing. EKSPLORIUM-BULETIN PUSAT TEKNOLOGI BAHAN GALIAN NUKLIR, 46(1), 1614-1628.

[78] Simon, H. A. (1997). Administrative behavior. Free Press.

[79] Tang, C. S. (2006). Supply chain risk management. International Journal of Production Economics, 103(2), 451–488.

[80] Meda, R. (2025). AI-Driven Demand and Supply Forecasting Models for Enhanced Sales Performance Management: A Case Study of a Four-Zone Structure in the United States. Metallurgical and Materials Engineering, 1480-1500.

[81] van der Aalst, W., Bichler, M., & Heinzl, A. (2018). Robotic process automation. Business & Information Systems Engineering, 60(4), 269–272.

[82] Wu, D., Rosen, D. W., Wang, L., & Schaefer, D. (2015). Cloud-based design and manufacturing. Journal of Manufacturing Systems, 39, 1–15.

[83] Seenu, A., Sheelam, G. K., Motamary, S., Meda, R., Koppolu, H. K. R., & Inala, R. (2025, July). AI-Driven Innovations in Infrastructure Management with 6G Technology. In 2025 2nd International Conference on Computing and Data Science (ICCDS) (pp. 1-6). IEEE.

[84] Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). The new organizing logic of digital innovation. Information Systems Research, 21(4), 724–735.

[85] Zhang, H., Chen, X., & Wang, Y. (2025). Decision intelligence for smart cities. Sustainable Cities and Society, 98, 104715.

[86] Zhou, L., Pan, S. L., & Tan, C. W. (2025). Digital platform ecosystems and AI decision systems. Information Systems Journal, 35(2), 287–314.

[87] Rani, P. S., Kummari, D. N., Yellanki, S. K., Meda, R., Koppolu, H. K. R., & Inala, R. (2025, July). Blockchain and AI for Securing Electrical Infrastructure. In 2025 2nd International Conference on Computing and Data Science (ICCDS) (pp. 1-6). IEEE.

[88] Zweig, K. A. (2016). Network analysis literacy. Springer.

[89] Brynjolfsson, E., & McElheran, K. (2026). Data-driven decision-making at scale. American Economic Review, 116(5), 1580–1612.

[90] World Economic Forum. (2026). Intelligent cloud platforms for industry transformation. World Economic Forum.