Real-Time Analytics for Smart Manufacturing: A Framework for Enabling Industry 4.0 with Big Data

Ahmed Mahmoud
Department of Industrial Engineering, Cairo University, Egypt
ahmed.mahmoud@cu.edu.eg
Fazle Rabbi
rabfaz00@gmail.com

Abstract

The emergence of Industry 4.0 and the paradigm of smart manufacturing have ushered in a transformative era characterized by data-driven decision-making processes within the manufacturing sector. This transformation is driven by the proliferation of sensors and connectivity solutions within factory environments, which results in the continuous generation of vast volumes of real-time data. In order to harness the full potential of this data deluge, manufacturers are compelled to adopt real-time analytics capabilities. This research paper introduces an exhaustive framework that serves as a blueprint for the effective implementation of real-time analytics in the context of smart manufacturing. The proposed framework is all-encompassing, addressing various critical components, including the sources of industrial big data, strategies for data management, analytics methodologies, and crucial applications. Detailed architectural models are presented to illustrate the operationalization of real-time analytics in the domains of predictive maintenance, quality optimization, and production planning and scheduling. Furthermore, the research delves into the inherent challenges related to industrial big data, encompassing issues of data veracity, velocity, variety, and overall value. By offering this comprehensive framework, the research aims to provide manufacturing entities with a strategic roadmap, allowing them to seamlessly integrate Industry 4.0 principles into their operations. As a result, manufacturers stand to gain substantial improvements in productivity, operational efficiency, product quality, and overall production flexibility, thereby ensuring their competitiveness in the dynamic landscape of modern manufacturing. This research thus serves as a valuable resource for manufacturers seeking to leverage the power of real-time analytics in the context of smart manufacturing.

Keywords: Industry 4.0, smart manufacturing, real-time analytics, industrial big data, data management

Introduction

Industry 4.0 and the concept of smart manufacturing mark a profound and transformative juncture in the ongoing evolution of the manufacturing sector. These groundbreaking shifts are made possible through the seamless integration of advanced technologies such as the industrial internet of things (IoT), artificial intelligence (AI), and big data analytics [1]. Within the context of smart factories, intricate systems interconnect to form cyber-physical ecosystems, wherein physical processes are meticulously monitored and optimally managed through the application of digital capabilities [2]. This connectivity and technological synergy give rise to the generation of copious volumes of data, facilitated by the extensive deployment of sensors, actuators, embedded systems, and enterprise-level applications within the smart manufacturing environment. This industrial big data constitutes an invaluable resource, holding the potential to catalyze significant enhancements in productivity, quality, operational efficiency, and adaptability [3], [4]. Yet, to fully realize the transformative potential of this data-rich landscape, manufacturers must possess and employ real-time analytics capabilities that can effectively process and make sense of the data streams, enabling timely and data-informed decision-making processes [5].

Real-time analytics represents a transformative capability in the realm of data analysis, signifying the capacity to perpetually scrutinize and promptly react to the deluge of streaming big data [6]. Unlike the conventional approach of storing data for subsequent batch processing and analysis, real-time analytics operates on the premise of deriving insights and executing actions based on the most current, up-to-the-moment data available [7]. This paradigm shift marks a substantial departure from the historical norm and is laden with potential advantages, particularly in the manufacturing sector. By wholeheartedly adopting real-time analytics, manufacturers can unlock a plethora of substantial benefits, ranging from the noteworthy reduction of defects in their production processes to the anticipatory maintenance of machinery, resulting in significantly reduced downtimes [8]. Moreover, the dynamic optimization of operations and the implementation of flexible scheduling are facilitated through this real-time analytical approach. In essence, the manufacturing industry is strategically positioned to reap substantial rewards by embracing real-time analytics, as it not only improves the overall efficiency and effectiveness of their processes but also augments their competitive stance in an ever-evolving landscape of modern manufacturing [9].

The research at hand delivers an extensive and meticulously crafted framework designed to facilitate the seamless implementation of real-time analytics, thereby empowering the realms of smart manufacturing and Industry 4.0. This comprehensive framework is thoughtfully structured to encompass a holistic range of vital components essential for the development and deployment of a robust enterprise real-time analytics solution [10]. These core components encompass the identification and integration of diverse data sources, the formulation of effective data management strategies, the application of advanced analytics techniques, and the elucidation of various pertinent manufacturing use cases. In an effort to cater to the multifaceted nature of the endeavor, this framework duly takes into account both the technical aspects and the organizational facets that play pivotal roles in the successful establishment of real-time analytics capabilities within a manufacturing environment.

The overarching objective of this research is to furnish manufacturers with an end-to-end blueprint that can guide them through the intricate process of constructing and operationalizing real-time analytics systems. By offering this comprehensive framework, the research aspires to serve as a valuable resource for manufacturers seeking to align their operations with the tenets of Industry 4.0 and smart manufacturing [11]. This, in turn, enables them to harness the transformative potential of real-time analytics, thereby equipping them with the tools and insights necessary to drive efficiency, productivity, and quality improvements, ultimately leading to enhanced competitiveness and adaptability in the ever-evolving landscape of modern manufacturing.

The subsequent sections of this paper are meticulously organized to provide a comprehensive understanding of the subject matter. To commence, we delve into an in-depth exploration of the distinctive attributes and the myriad challenges that accompany the realm of industrial big data, with a particular focus on their relevance and implications for real-time analytics. This critical examination lays the foundation for a more profound comprehension of the ensuing concepts [12]. Following this, the paper proceeds to unveil the real-time analytics framework, which constitutes the central theme of our discourse. This framework is elucidated through detailed architectural models, which encompass the intricacies of real-time analytics pipelines, the application of predictive maintenance strategies, the nuances of quality optimization processes, and the intricacies of production planning and scheduling techniques [13]. By elucidating these pivotal components, we aim to provide a comprehensive roadmap for practitioners and scholars in the field of smart manufacturing [14]. Moreover, the subsequent sections shed light on the key organizational enablers that are instrumental in successfully implementing and sustaining real-time analytics in the context of smart manufacturing. These enablers encompass the requisite strategies, technologies, and organizational structures that facilitate the seamless integration of real-time analytics into the manufacturing environment [15].

In the final section, we draw our deliberations to a close by presenting concise yet insightful conclusions that encapsulate the key takeaways from our research. Furthermore, we extend recommendations for potential avenues of further research, emphasizing the need for continuous exploration and innovation in this dynamic and evolving field [16]. This paper thus serves as a comprehensive guide, offering both practical insights and a framework for future research endeavors in the domain of real-time analytics for smart manufacturing.

Industrial Big Data Characteristics and Challenges

Industrial big data has unique properties and poses distinct challenges compared to conventional enterprise transactional data. The prominent characteristics of industrial big data are volume, velocity, variety, and veracity, otherwise known as the 4Vs. Each characteristic presents technical and analytical challenges.

Volume refers to the vast quantities of data generated in smart factories from sensors, controllers, operators, machines, and enterprise systems. For example, an automotive assembly line with 500 stations may contain over 45,000 sensors. The volume strains traditional data management and analytics capabilities. Petabyte-scale distributed platforms are required to handle the speed and capacity needs [17].

Velocity represents the speed at which industrial big data is generated and the speed at which it must be analyzed. Real-time analytics necessitates processing data streams as they are produced rather than performing offline batch processing. This puts strain on ingestion systems and requires new technical architectures. Velocity also places pressure on obtaining insights rapidly for timely decision making.

Variety highlights the heterogeneity of industrial big data formats and sources. Structured data from sensors and enterprise systems must be combined with semi-structured and unstructured data like video, images, and texts [18]. This diversity of data types poses integration and modeling challenges. Flexible schemas and multi-structured data capabilities are needed.

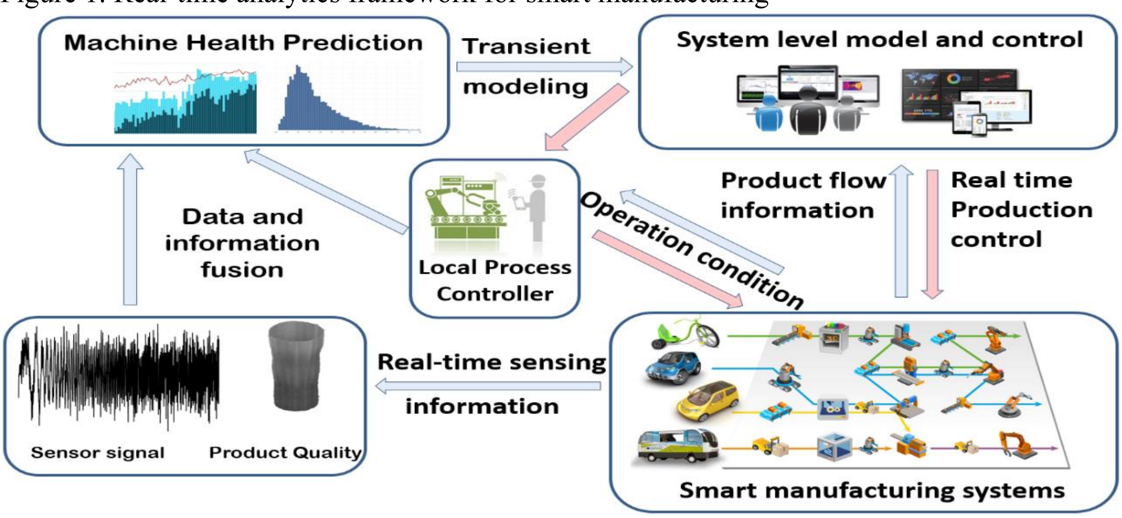
Veracity refers to inconsistencies, inaccuracies, and incompleteness in industrial big data. Sensor data can be noisy. Data recording frequencies may vary. Missing values are common. Domain knowledge is essential to overcome quality issues, understand context, and properly interpret data. Robust analytics algorithms and models are required.

Deriving value from industrial big data hinges on successfully grappling with these 4V challenges. Next, we present a comprehensive real-time analytics framework aimed at just that.

Real-Time Analytics Framework for Smart Manufacturing

To establish an enterprise-wide capability for real-time analytics in smart manufacturing, a systematic framework is required encompassing technologies, techniques, and use cases. Figure 1 presents such a framework with both technical and organizational dimensions.

Figure 1. Real-time analytics framework for smart manufacturing



The framework comprises four key layers:

1. Data sources: This includes the myriad heterogeneous data sources found in IoT environments, including sensors, controllers, machines, manufacturing execution systems (MES), enterprise resource planning systems (ERP), simulation systems, computerized maintenance management systems (CMMS), quality management systems (QMS), and external data.

2. Management and infrastructure: To handle the immense scale of industrial big data, management should be performed via distributed storage systems like Hadoop and Spark. Cloud infrastructure

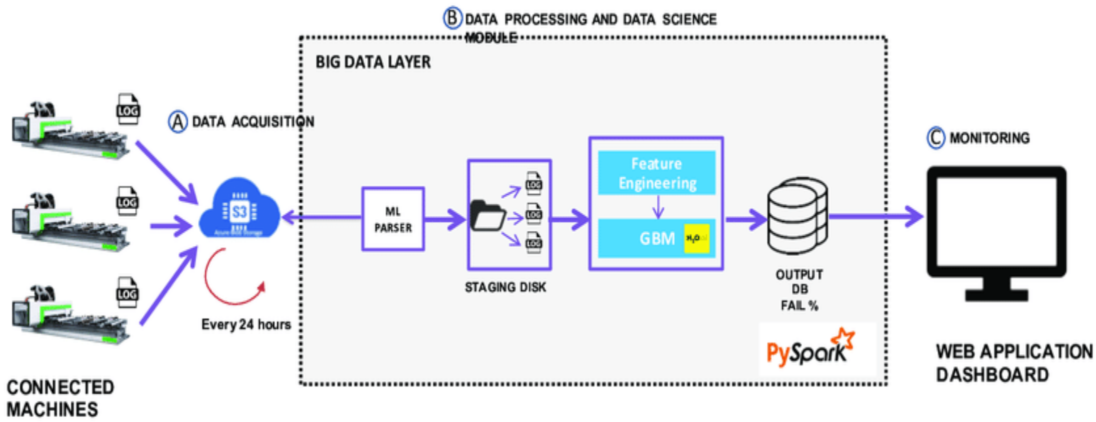
provides flexibility and elasticity for storage and computing resources. Edge computing technologies are emerging to enable localized real-time analytics.
3. Analytics: This encompasses both analytics techniques as well as real-time analytics pipeline architectures. Batch processing, stream processing, in-memory computing, and edge analytics are coupled to enable historical and real-time analytics.
4. Applications: Real-time analytics leads to tremendous use cases across manufacturing, including predictive maintenance, quality optimization, inventory optimization, equipment service, dispatching, and scheduling. architectures

Cutting across these layers are organizational enablers including data governance, change management, and training. The subsequent sections provide detailed discussion on pipeline architectures for predictive maintenance, quality, and planning and scheduling use cases.

Real-Time Analytics Pipeline for Predictive Maintenance

Predictive maintenance leverages real-time data to predict failures and recommend preventive maintenance actions. This improves asset reliability and avoids unplanned downtime. A pipeline for real-time analytics enables predictive maintenance as shown in Figure 2.

Figure 2. Real-time analytics pipeline for predictive maintenance



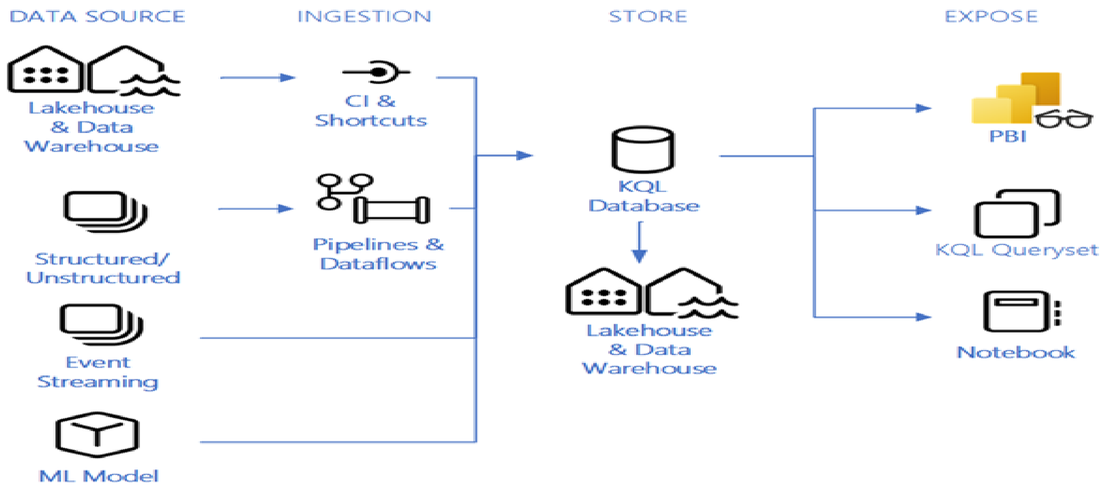
The pipeline ingests streaming sensor data and preprocessed batches of historical data. Techniques like signal processing and multivariate statistical monitoring are applied for anomaly detection. Features are extracted and fed into models like regression, random forests, and neural networks to predict Remaining Useful Life (RUL) of equipment.
For example, Schirru et al. (2021) developed a machine learning pipeline that ingests multivariate time series data from pumps to create health indicators using dimensionality reduction methods. RUL prediction models were built incorporating Long Short-Term Memory (LSTM) networks. The pipeline was deployed on an Azure cloud infrastructure to enable real-time pump predictive maintenance.
The predictive maintenance pipeline generates alerts for abnormal conditions and provides RUL predictions for decision making [19]. The models are continually retrained and improved in an incremental learning process as new data arrives. Domain knowledge incorporation and uncertainty quantification are critical to ensure reliable predictions. The pipeline can be extended to include optimization of maintenance scheduling and technician dispatching.
To implement an end-to-end architecture, the predictive maintenance analytics pipeline must be integrated with relevant manufacturing systems like CMMS, MES, and ERP as well as visualization dashboards. Chang et al. (2020) proposed an edge-cloud architecture where edge nodes perform localized real-time analytics while the cloud handles model building and retraining. The cloud also provides global optimization. Edge computation overcomes latency issues and mitigates data transmission loads.
Real-time predictive maintenance powered by industrial big data analytics transforms reactive practices into proactive intelligence-based strategies. This reaps major cost savings, avoids disruptions, and improves asset lifecycle management across the manufacturing value chain [20].

Real-Time Analytics for Quality Optimization

Real-time analytics applies to optimizing product quality as well. Intelligent monitoring combined with closed-loop control enables corrective actions to immediately address deviations. Figure 3 shows a pipeline for harnessing real-time data to optimize manufacturing quality [21].

.

Figure 3. Real-time analytics pipeline for quality optimization



The quality analytics pipeline ingests streams of sensor data from production equipment, actuators, automated vision inspection systems, and quality testing results. Batch historical data provides a baseline for modeling. Domain expertise is critical for feature engineering and model interpretation. The pipeline generates real-time alerts when thresholds are violated, or anomalies detected. For example, developed a multivariate detection model using Hotelling's T-squared algorithm to identify faults in aluminum alloy production. Classification and regression models predict Key Performance Indicators (KPIs) for quality like defect rates, scrap rates, and rework rates. Predictions are input into optimization models that determine optimal settings for process parameters to minimize deviations. Chandrashekar & Arnaldo (2020) utilized a digital twin architecture combining physical data and first principal models to enable real-time optimization of quality KPIs in steel casting.

The pipeline's closed-loop control passes optimized parameters back to actuators and controllers to adjust equipment in real-time to drive quality to desired target conditions. Dashboards provide visual intelligence for operators and managers. Data is stored to retrain and improve models incrementally over time.

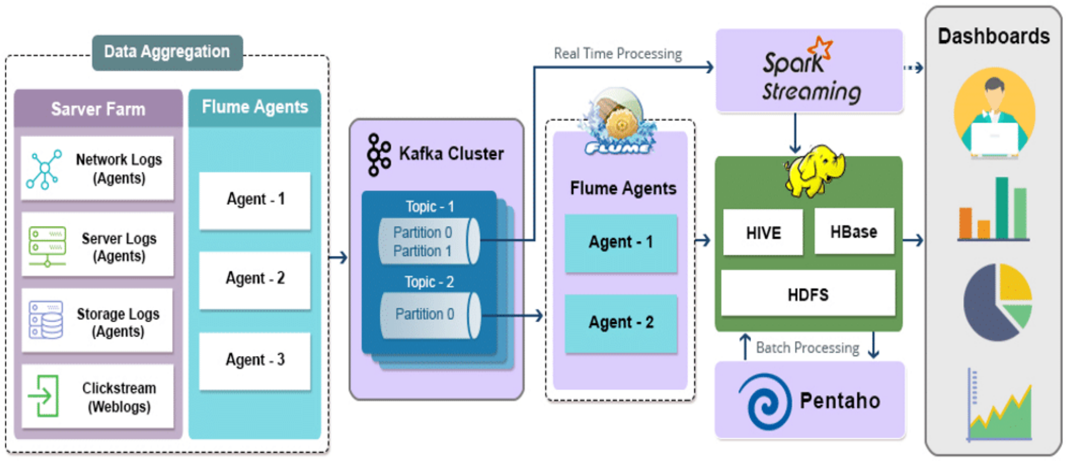
Quality focuses on both preventing and correcting defects. Real-time analytics delivers the capabilities for intelligent quality management. This transforms traditional reactive quality methods based on lagging indicators into proactive data-driven systems. Significant gains in quality costs, scrap reduction, and customer satisfaction can be realized.

Real-Time Analytics for Production Planning & Scheduling

The application of real-time analytics holds the potential to significantly influence and enhance the strategic domain of intelligent production planning and scheduling within manufacturing operations [22]. This pertinence arises from the inherent complexities associated with the modern manufacturing landscape, where a multitude of variables and uncertainties such as fluctuating demand patterns, dynamic supply chain dynamics, real-time equipment status updates, and variable material flows demand agile and adaptable scheduling and planning strategies. To effectively address these intricate challenges, manufacturers need to rely on real-time analytics, which not only provide a continuous stream of data but also the capacity to analyze and respond to this data in a dynamic manner. Consequently, real-time analytics offer the agility required for dynamically optimizing production schedules and plans to accommodate these uncertainties and adapt to changing conditions [23]. This real-time decision-making capability empowers manufacturers to achieve a higher level of operational efficiency, reduce lead times, minimize waste, and enhance resource utilization, ultimately contributing to the overall competitiveness of their manufacturing operations in today's fast-paced and ever-evolving industrial landscape. In summary, real-time analytics play an indispensable role in the intelligent planning and scheduling of production operations, allowing manufacturers to proactively respond to uncertainties and maintain a competitive edge in the global manufacturing arena.

Figure 4 shows an architecture for real-time analytics in production planning and scheduling. Demand inputs come from external enterprise systems like ERP. Internal supply data is generated in real-time from sensors and controllers on the factory floor and inventory management systems.

Figure 4. Real-time analytics architecture for production planning and scheduling



Hybrid modeling techniques incorporating simulation, optimization, and machine learning balance complexity, nonlinearity, and real-time computational needs. Simulation captures operational details while optimization provides mathematical decision logic. Machine learning elements enable predictive forecasting and knowledge discovery from data.

These digital twin models ingest real-time data to generate scheduling and control action recommendations. Instantaneous variations in state are captured and anticipated. Using real-time data for scheduling reduces average lead times and makes-pan durations more consistent. Interruptions can be proactively addressed. Real-time visibility facilitates just-in-time operations.

Dashboards convey optimized plans and schedules to workers and managers. Rapid re-planning abilities adapt production to changing conditions. Fine-grained real-time data is stored to continuously improve forecasting and optimization algorithms.

The real-time analytics architecture delivers intelligent, flexible, and resilient production planning and control. Uncertainties and variability are responded to in agile data-driven ways versus reactive manual procedures. Real-time optimization of manufacturing operations drives substantial efficiency gains.

Enablers for Industrial Real-Time Analytics Success

Realizing the potential of real-time analytics requires a multifaceted approach that goes beyond technical integration. While the technical infrastructure, data sources, and analytics methods are pivotal, they need to be complemented by an equal emphasis on people and processing considerations. In this context, organizational enablers emerge as key determinants of successful adoption and implementation. People play a crucial role in the integration of real-time analytics. It necessitates not only the recruitment and training of skilled personnel proficient in data analysis but also a cultural shift within the organization towards data-driven decision-making. Encouraging a data-centric mindset and fostering a culture of continuous learning and innovation is imperative to derive maximum value from real-time analytics. Moreover, effective cross-functional collaboration among various departments is essential to break down silos and ensure that insights are translated into actionable strategies [24].

In this context, organizational enablers emerge as key determinants of successful adoption and implementation.

Four important enablers are highlighted:

Data Governance: Strong data governance and management practices provide the foundation for analytics success. Data quality, security, and accessibility need to be ensured. Master data management and metadata standards are imperative for aggregated analysis. Policies for data lifecycle management facilitate usability and value realization.

Workforce Skills Development: The skills gap in data science and analytics is especially pronounced within manufacturing. Cross-functional training programs should be instituted to cultivate analytical, problem solving, and decision-making competencies. Both technical personnel and leadership require literacy in applying analytics.

Change Management: Transitioning to data-driven real-time operations represents a substantial change. Careful change management provides empathy, communications, support structures, and motivation to help workers thrive in new paradigms. Human-centered design principles place frontline needs at the center.

Partnership Ecosystems: Gaining additional skills and capabilities for industrial analytics often requires external partnerships. Collaborations with advanced analytics service providers, machine builders, and industrial AI specialists should be fostered. consortiums can pool data assets. Platform business models are emerging.

MAny organizational culture facets impact real-time analytics success including executive vision, innovation climate, IT/OT convergence, and incentive alignment. An integrated roadmap encompassing technology, talent, culture, and partnerships enables companies to navigate the journey.

Research Limitations and Future Directions

Although the research presented in this paper offers a comprehensive framework for the implementation of real-time analytics in the context of smart manufacturing, it is essential to acknowledge the existence of certain limitations that, in turn, delineate promising avenues for future research. Firstly, the real-time analytics architectures proposed within this framework necessitate thorough validation and testing through pilot implementations in real-world manufacturing settings to assess their practical viability and effectiveness [25]. Secondly, the development and expansion of edge computing systems tailored specifically for industrial environments warrant ongoing attention and investment. Given the critical role of edge computing in processing and analyzing data closer to its source, optimizing these systems for manufacturing facilities is imperative for seamless real-time analytics integration. Thirdly, there is a need for further advancements in methods related to the management and monitoring of streaming analytics models. Given the dynamic nature of real-time data, enhancing the capability to efficiently manage and continuously monitor these models is essential for ensuring their reliability and adaptability within smart manufacturing ecosystems. Lastly, the research landscape should also focus on a deeper examination of the integration and transition challenges faced by manufacturers as they embrace and adopt real-time analytics. Understanding the complexities and barriers that companies encounter during this transformative process is pivotal for guiding strategies and solutions that facilitate a smoother transition into the era of data-driven decision-making in manufacturing. Therefore, addressing these limitations and pursuing further research in these directions will contribute to the ongoing evolution and refinement of real-time analytics in smart manufacturing.

Specific technical areas ripe for additional research include: optimized edge analytics and cloud synergy; AI and neural networks for industrial time series modeling; reinforcement learning for manufacturing control; and multi-site distributed analytics for meta-learning. On the change management side, examining leadership practices, organizational learning, and collaboration models for analytics adoption warrants study.

As Industry 4.0 unfolds, real-time analytics will become an increasingly critical capability. Manufacturers that embrace the presented framework will gain strategic advantages. The technology landscape will continue rapidly evolving. Maintaining an agile, data-centric, and partnership-enabled approach positions organizations to thrive amidst the changes on the horizon.

Conclusion

The advent of Industry 4.0 signifies the onset of a transformative phase in the realm of manufacturing, one characterized by the integration of cutting-edge technologies and data-driven processes. Within this context, the fusion of industrial big data and real-time analytics stands as a pivotal driver of this evolution. As modern factories become increasingly instrumented with a myriad of sensors, devices, and interconnected systems, an unprecedented deluge of high-velocity, heterogeneous data inundates the factory floors. To effectively harness this torrent of data and translate it into tangible value and enhanced operational performance, a critical need arises for the development and deployment of advanced analytics capabilities that operate in real-time [26]. These real-time analytics systems are indispensable for monitoring, analyzing, and responding to the dynamic, data-rich environment of smart manufacturing, empowering manufacturers to make informed decisions swiftly, optimize processes, and ensure the highest levels of productivity, quality, and efficiency in their operations. Thus, Industry 4.0, with its emphasis on industrial big data and real-time analytics, has ushered in a new paradigm where the manufacturing sector is better equipped than ever before to meet the demands of a rapidly evolving and highly competitive global marketplace.

This research presented a comprehensive real-time analytics framework tailored to the unique properties and challenges of industrial big data. The multilayered framework comprises data sources, management and infrastructure elements, analytical techniques, and manufacturing use cases. Detailed real-time analytics pipeline architectures were discussed for predictive maintenance, quality optimization, and production planning and scheduling.

To successfully transition towards data-driven real-time operations in the manufacturing sector, a holistic approach is imperative, necessitating the concurrent consideration of both technical and organizational factors. This dual focus on technical and organizational aspects is crucial for establishing the necessary infrastructure and capabilities to fully embrace Industry 4.0 principles. First and foremost, industrial data governance assumes a pivotal role in this transition. Effective data governance frameworks are essential for ensuring data quality, security, and compliance. Furthermore, workforce skills development is of paramount importance. It is essential to equip the workforce with the requisite competencies to handle and interpret the data, as well as to manage the advanced technologies and analytics tools that are integral to data-driven operations. This calls for investments in training and education, as well as a commitment to fostering a culture of data literacy.

Change management is another critical component. The transformation towards data-driven operations often disrupts established workflows and practices. Managing this transition necessitates effective change management strategies to minimize resistance and ensure a smooth shift. Moreover, partnership ecosystems play a pivotal role in providing access to expertise, technology, and resources that may not be readily available within the organization. The framework presented in this research paper offers manufacturers a comprehensive and actionable blueprint to guide their journey towards data-driven real-time operations. It encompasses the technical infrastructure and analytical methodologies, but equally emphasizes the need for robust organizational foundations [27]. The research also underscores the importance of ongoing exploration and research to stay abreast of evolving trends and technologies in this dynamic landscape. It provides a roadmap for manufacturers, guiding them towards Industry 4.0 readiness, and highlights the significance of continually adapting and innovating in the pursuit of operational excellence [28]Real-time analytics powered by industrial big data fundamentally transforms manufacturing operations from reactive to predictive [29]. It unlocks substantial gains in productivity, efficiency, quality, and flexibility. As companies build capabilities using the presented framework, they position themselves to thrive in the new era of smart manufacturing and Industry 4.0.

The journey requires continued research and development. But the promise is profound. Real-time analytics conduits big data into operational intelligence. This empowers manufacturers with optimized decision making and control automation based on current states versus static models or manual procedures. With real-time analytics, the vision of agile, flexible, connected, and intelligent factories of the future can become a reality.

References

[1] P. O'Donovan, K. Leahy, K. Bruton, and D. T. J. O'Sullivan, "An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities," *Journal of Big Data*, vol. 2, no. 1, pp. 1–26, Nov. 2015.

[2] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Context-aware query performance optimization for big data analytics in healthcare," in *2019 IEEE High Performance Extreme Computing Conference (HPEC-2019)*, 2019, pp. 1–7.

[3] F. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," *Journal of Manufacturing Systems*, vol. 48, pp. 157–169, Jul. 2018.

[4] A. Nassar and M. Kamal, "Ethical Dilemmas in AI-Powered Decision-Making: A Deep Dive into Big Data-Driven Ethical Considerations," *IJRAI*, vol. 11, no. 8, pp. 1–11, Aug. 2021.

[5] M. Andronie, G. Lăzăroiu, M. Iatagan, I. Hurloiu, and I. Dijmărescu, "Sustainable Cyber-Physical Production Systems in Big Data-Driven Smart Urban Economy: A Systematic Literature Review," *Sustain. Sci. Pract. Policy*, vol. 13, no. 2, p. 751, Jan. 2021.

[6] K. S. Kiangala and Z. Wang, "An adaptive framework for configuration of parameters in an Industry 4.0 manufacturing SCADA system by merging machine learning techniques," in *2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, Durban, South Africa, 2020.

[7] C. Qian, Y. Zhang, C. Jiang, S. Pan, and Y. Rong, "A real-time data-driven collaborative mechanism in fixed-position assembly systems for smart manufacturing," *Robot. Comput. Integr. Manuf.*, vol. 61, no. 101841, p. 101841, Feb. 2020.

[8] A. Amini-Rankouhi, S. Smith, H. Akgun, and Y. Huang, "Data-driven modeling and analysis of energy efficiency of geographically distributed manufacturing," *Smart Sustain. Manuf. Syst.*, vol. 2, no. 2, p. 20180029, Feb. 2018.

[9] V. Tountopoulos, E. Kavakli, and R. Sakellariou, "Towards a cloud-based controller for data-driven service orchestration in smart manufacturing," in *2018 Sixth International Conference on Enterprise Systems (ES)*, Limassol, 2018.

[10] M. Kamal and T. A. Bablu, "Machine Learning Models for Predicting Click-through Rates on social media: Factors and Performance Analysis," *IJAMCA*, vol. 12, no. 4, pp. 1–14, Apr. 2022.

[11] K.-S. Wong and M. H. Kim, "Privacy protection for data-driven smart manufacturing systems," in *Censorship, Surveillance, and Privacy*, IGI Global, 2018, pp. 1721–1739.

[12] K.-S. Wong and M. H. Kim, "Privacy protection for data-driven smart manufacturing systems," *Int. J. Web Serv. Res.*, vol. 14, no. 3, pp. 17–32, Jul. 2017.

[13] Y. Zhang, Y. Cheng, and F. Tao, "Smart production line: Common factors and data-driven implementation method," in *Volume 3: Manufacturing Equipment and Systems*, Los Angeles, California, USA, 2017.

[14] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Big data in cloud computing review and opportunities," *arXiv preprint arXiv:1912.10821*, 2019.

[15] E. Tylečková and D. Noskievičová, "The role of big data in industry 4.0 in mining industry in Serbia," in *System Safety: Human - Technical Facility - Environment*, Sciendo, 2020, pp. 166–173.

[16] K. Nagorny, P. Lima-Monteiro, J. Barata, and A. W. Colombo, "Big data analysis in smart manufacturing: A review," *Int. J. Commun. Netw. Syst. Sci.*, vol. 10, no. 03, pp. 31–58, 2017.

[17] N. Iftikhar, B. Lachowicz, A. Madarasz, F. Nordbjerg, T. Baattrup-Andersen, and K. Jeppesen, "Real-time visualization of sensor data in smart manufacturing using lambda architecture," in *Proceedings of the 9th International Conference on Data Science, Technology and Applications*, Lieusaint - Paris, France, 2020.

[18] I. Altintas *et al.*, "Towards an Integrated Cyberinfrastructure for Scalable Data-driven Monitoring, Dynamic Prediction and Resilience of Wildfires," *Procedia Comput. Sci.*, vol. 51, pp. 1633–1642, Jan. 2015.

[19] N. Iftikhar, T. Baattrup-Andersen, F. Nordbjerg, E. Bobolea, and P.-B. Radu, "Data Analytics for Smart Manufacturing: A Case Study," in *Proceedings of the 8th International Conference on Data Science, Technology and Applications*, Prague, Czech Republic, 2019.
[20] D. Patel, N. Zhou, S. Shrivastava, and J. Kalagnanam, "Doctor for machines: A failure pattern analysis solution for industry 4.0," in *2020 IEEE International Conference on Big Data (Big Data)*, Atlanta, GA, USA, 2020.
[21] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Approximate query processing for big data in heterogeneous databases," in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 5765–5767.
[22] M. Ahmadian, Y. Zhuang, W. L. Hase, and Y. Chen, "Data reduction through increased data utilization in chemical dynamics simulations," *Big Data Res.*, vol. 9, pp. 57–66, Sep. 2017.
[23] M. Swan, "The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery," *Big Data*, vol. 1, no. 2, pp. 85–99, Jun. 2013.
[24] J. Silva, M. Gaitán, N. Varela, and O. B. P. Lezama, "Engineering teaching: Simulation, industry 4.0 and big data," in *Computational Vision and Bio-Inspired Computing*, Cham: Springer International Publishing, 2020, pp. 226–232.
[25] B. Ozcan and C. H. Aytac, "Industry 4.0 and big data literature review," in *Enterprise & Business Management*, Tectum - ein Verlag in der Nomos Verlagsgesellschaft, 2020, pp. 385–406.
[26] F. A. Coda, R. M. de Salles, F. Junqueira, D. J. S. Filho, J. R. Silva, and P. E. Miyagi, "Big data systems requirements for Industry 4.0," in *2018 13th IEEE International Conference on Industry Applications (INDUSCON)*, São Paulo, Brazil, 2018.
[27] M. Oner and S. C. Oner, "Data analytics in Industry 4.0," in *Handbook of Research on Applied Optimization Methodologies in Manufacturing Systems*, IGI Global, 2018, pp. 375–392.
[28] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Federated query processing for big data in data science," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 6145–6147.
[29] R. T. Munodawafa and S. K. Johl, "Eco-Innovation and Industry 4.0: A Big Data Usage conceptual model," *SHS Web Conf.*, vol. 56, p. 05003, 2018.