International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Revolutionizing Data Strategy: The Impact of AI/ML-Driven Data Governance and Master Data Management in Health, Pharma, and Financial Industries

Naveen Sri Harsha Rellu[1], Vinodkumar Reddy Surasani[2], Raghuvaran Reddy Kalluri[3], Dr. Nagaraju Devarakonda[4]

Submitted: 28/04/2024 **Revised:** 01/06/2024 **Accepted:** 10/06/2024

Abstract: The rapid growth of data in healthcare, pharmaceutical, and financial industries has intensified the need for robust data governance and master data management (MDM) systems. This study explores how Artificial Intelligence (AI) and Machine Learning (ML) are transforming traditional data strategies by enhancing data quality, regulatory compliance, and operational efficiency. Employing a mixed-methods approach, the research combines a structured literature review, expert interviews, and a sector-specific survey across 150 professionals. Results reveal that the financial sector leads in AI/ML adoption and benefits, including a 52% improvement in data accuracy and reduced time to compliance. Healthcare and pharmaceutical sectors, while advancing, encounter integration and regulatory challenges. Statistical analysis using ANOVA confirms significant sectoral differences in return on investment ($p < 0.001$). Thematic insights further highlight key trends such as automated compliance, real-time data monitoring, and metadata discovery. The findings emphasize that AI/ML-driven governance is a strategic necessity, not just a technological advancement. This study contributes to the evolving discourse on intelligent data systems and provides practical insights for organizations aiming to build scalable, compliant, and efficient data infrastructures.

Keywords*: AI/ML, Data Governance, Master Data Management, Healthcare, Pharmaceuticals, Finance, Compliance, Data Accuracy, Digital Transformation.*

Introduction

The Digital Data Deluge and the strategic need for governance

In today's hyper-connected and data-intensive world, organizations across industries are grappling with an overwhelming influx of data. From electronic health records and genomic data in healthcare to clinical trial repositories in pharmaceuticals and transactional datasets in financial services, the volume, velocity, and variety of data have grown exponentially. Amidst this data explosion, traditional data governance and master data management (MDM) models are proving to be inadequate (Jena et al., 2024). As the complexity of data ecosystems increases, so too does the necessity for a more intelligent, scalable, and adaptive approach to managing and governing data effectively.

1Sr Manager Software Engineering, Optum, Minneapolis
2Sr Software Engineer, RBC Wealth Management, Minneapolis
3Master Software Engineer, RBC Wealth Management, Minneapolis
4Professor Grade-1 & HOD, Department of Software System Engineering, School of Computer Science & Engineering, VIT-AP University Amaravathi

AI and Machine Learning as game changers

Artificial Intelligence (AI) and Machine Learning (ML) have emerged as transformative technologies that not only automate but also enhance the decision-making processes involved in data governance and MDM. These tools enable organizations to move from reactive data strategies to proactive, predictive, and prescriptive models of data management. AI/ML algorithms can detect anomalies, uncover hidden patterns, ensure compliance, and automate workflows with a level of speed and accuracy that traditional methods cannot match (Hashio et al., 2024). Consequently, AI/ML-driven governance is not just a technological upgrade—it is a strategic imperative for industries that deal with high-stakes data environments.

Sectoral transformation through intelligent data management

The impact of AI/ML-enabled data governance is particularly profound in the health, pharmaceutical, and financial sectors. These industries are not only data-intensive but also highly regulated, where data quality, integrity, and compliance are critical. In

healthcare, AI-infused MDM ensures better patient data harmonization across disparate systems, improving clinical decision-making and patient outcomes (Vijay et al., 2024). In pharmaceuticals, it supports more efficient drug discovery, regulatory adherence, and trial data management. Meanwhile, financial institutions leverage AI-powered governance to enhance fraud detection, risk analytics, and customer insights—all while staying compliant with evolving global regulations.

Master Data Management reimagined

Master Data Management, traditionally a rule-based and labor-intensive process, is being revolutionized by AI and ML. Through automation and intelligent data matching, ML algorithms can resolve duplicate records, standardize entries, and maintain a golden record of truth more efficiently. Moreover, AI enables contextual understanding of master data, allowing systems to adapt dynamically to new data types, formats, and sources. This evolution is especially crucial in sectors like pharma, where master data on products, trials, and patients must be synchronized globally, and in finance, where customer and transaction data must be continuously reconciled across various platforms (Barrett et al., 2023).

Bridging compliance and innovation

Regulatory compliance is another area where AI/ML-driven governance offers tremendous value. Automated policy enforcement, intelligent auditing, and real-time monitoring help organizations maintain transparency and accountability. This balance between compliance and innovation is particularly vital in health and finance, where regulations like HIPAA, GDPR, and Basel III mandate stringent data controls. AI/ML tools not only reduce the risk of non-compliance but also free up human resources to focus on strategic initiatives rather than routine checks and balances (Camacho et al., 2024).

Purpose of the study

This study aims to explore the transformative role of AI and ML in redefining data governance and master data management across the healthcare, pharmaceutical, and financial industries. It investigates how these technologies are enabling smarter data strategies, improving operational efficiency, ensuring compliance, and fostering innovation. Through comparative insights and industry-specific use cases, the research highlights the potential, challenges, and future direction of AI/ML-powered data frameworks in mission-critical sectors.

Methodology

Research design

This study adopts a mixed-methods research design to comprehensively examine the impact of AI/ML-driven data governance and master data management (MDM) across the health, pharmaceutical, and financial industries. The combination of qualitative and quantitative approaches ensures a well-rounded understanding of both the strategic and operational effects of these technologies in data-intensive environments. The study is exploratory in nature, aiming to uncover patterns, practices, and perspectives related to AI/ML integration in data strategy.

Data collection methods

To provide robust insights into the phenomenon of revolutionizing data strategy, the study employed three primary data collection methods:

- Structured Literature Review: A systematic review of peer-reviewed journals, white papers, government reports, and industry publications was conducted to gather existing knowledge on AI/ML-based data governance and MDM. Sources were selected from databases such as IEEE Xplore, ScienceDirect, PubMed, and Scopus covering literature from 2015 to 2024.
- Semi-Structured Interviews: In-depth interviews were carried out with subject matter experts, including data scientists, CIOs, compliance officers, and digital transformation leads from the healthcare, pharma, and financial sectors. A total of 24 participants (8 from each sector) were selected using purposive sampling to ensure relevant and specialized insights.
- Survey Questionnaire: A structured survey was administered to 150 professionals (50 from each sector), including data governance specialists, IT managers, and AI/ML engineers. The questionnaire focused on challenges, adoption levels, perceived benefits, and risk management strategies associated with AI/ML in data governance and MDM.

Sampling strategy

The study used a purposive and stratified sampling strategy. For interviews and surveys, professionals were selected based on their direct involvement in AI/ML projects, data strategy planning, or MDM implementation. The sample was stratified across the three sectors—healthcare, pharmaceutical, and financial—to enable comparative analysis and sector-specific findings.

Analytical framework

To analyze the data, both qualitative and quantitative analytical techniques were employed:

- Thematic Analysis was applied to interview transcripts using NVivo software. Emerging themes included automation of compliance, data harmonization, AI-driven metadata discovery, and cross-platform data synchronization.
- Descriptive and Inferential Statistics were applied to the survey data using SPSS. Key metrics analyzed included adoption rate, return on investment (ROI), time to compliance, and data accuracy improvements. Statistical tests such as ANOVA and chi-square were used to identify significant differences across sectors.
- Comparative Case Analysis was conducted using secondary data from real-world case studies provided by industry sources. This helped highlight sector-specific implementations, use cases, and lessons learned.

Validity and reliability

To ensure the validity and reliability of the study:

- Triangulation was employed by cross-verifying findings from interviews, surveys, and literature.
- A pilot test of the survey questionnaire was conducted with 10 respondents to ensure clarity and consistency.
- Member checking was used during interviews to validate participants' responses and interpretations.
- Reliability of survey instruments was confirmed through Cronbach's Alpha, which showed values above 0.85 for all constructs.

Ethical considerations

All participants provided informed consent before taking part in the study. Confidentiality and anonymity were strictly maintained. Ethical approval was obtained from the Institutional Review Board (IRB) of the affiliated research institution. Data collected was used solely for academic purposes and stored securely in compliance with GDPR and HIPAA guidelines where applicable.

Results

The demographic profile of the respondents is summarized in Table 1. A total of 150 professionals (50 from each sector) participated in the survey. The average professional experience ranged from 7.5 years in the pharmaceutical sector to 9.1 years in the financial sector. The gender distribution revealed a slightly higher representation of males across all sectors, with the highest male representation (66%) in the financial sector. The roles covered included data scientists, CIOs, compliance officers, and digital transformation leaders. The majority of respondents were from urban centers, especially in the financial sector (90%).

Table 1: Survey demographics

Sector	No. of Respondents	Avg. Experience (years)	Male (%)	Female (%)	Age Range (Years)	Roles Covered	Location Distribution (Urban/Rural)
Healthcare	50	8.2	62	38	30–55	Data Scientists, CIOs, Analysts	80% Urban / 20% Rural
Pharmaceutical	50	7.5	58	42	28–50	Research Managers,	75% Urban / 25% Rural

						Compliance	
Financial	50	9.1	66	34	32–60	Risk Analysts, IT Managers	90% Urban / 10% Rural

As shown in Table 2, the adoption of AI/ML tools in data governance is highest in the financial sector (82%), followed by healthcare (76%) and pharmaceutical (69%). Time to regulatory compliance was significantly reduced in sectors with higher adoption—averaging just six months in the financial industry. Notably, the financial sector also reported the highest improvement in data accuracy (52%) and MDM efficiency gain (62%). Meanwhile, healthcare organizations reported a 45% accuracy improvement and a 54% gain in MDM efficiency, indicating a strong shift toward intelligent data practices. Cost reduction and automation of governance processes were evident across all sectors, with governance automation levels reaching 75% in the financial domain.

Table 2: AI/ML adoption rate and data strategy metrics by sector

Sector	**AI/ML Adoption Rate (%)**	**Avg. Time to Compliance (months)**	**Data Accuracy Improvement (%)**	**Cost Reduction (%)**	**Governance Automation Level (%)**	**MDM Efficiency Gain (%)**
Healthcare	76	8	45	32	68	54
Pharmaceutical	69	10	39	28	60	48
Financial	82	6	52	41	75	62

The perceived benefits of AI/ML in data governance and MDM were further examined and are detailed in Table 3. Respondents across all sectors overwhelmingly acknowledged improvements in data accuracy, faster compliance, enhanced decision-making, and better real-time monitoring as core advantages. Improved data accuracy scored highest in the financial sector (88%), while enhanced decision-making and real-time monitoring were also perceived more strongly in this sector. On average, the top five perceived benefits had over 70% agreement across sectors, confirming the strategic value of AI/ML integration.

Table 3: Perceived benefits of AI/ML in data governance and MDM

Benefit	**Healthcare (%)**	**Pharmaceutical (%)**	**Financial (%)**	**Overall Mean (%)**
Improved Data Accuracy	85	80	88	84.3
Faster Compliance	78	72	83	77.7
Cost Reduction	67	61	73	67.0
Enhanced Decision Making	81	75	86	80.7
Real-time Monitoring	70	66	79	71.7
Reduced Manual Errors	74	69	77	73.3

Automated Policy Enforcement	69	64	75	69.3
Better Data Integration	77	70	82	76.3

To assess the statistical significance of sectoral differences, an ANOVA test was performed on Return on Investment (ROI) as a dependent variable, with sector as the independent factor. The results, shown in Table 4, revealed a significant difference in ROI across the three sectors ($F = 29.07$, $p < 0.001$), with an effect size (η^2) of 0.28, indicating a large effect. These findings support the hypothesis that AI/ML implementation yields varied financial outcomes depending on sectoral maturity, regulatory burden, and operational readiness.

Table 4: ANOVA results for sectoral differences in Return on Investment (ROI)

Source	Sum of Squares	df	Mean Square	F	p-value
Between Groups	152.3	2	76.15	29.07	<0.001
Within Groups	384.7	147	2.62		
Total	537.0	149			
Effect Size (η^2)	0.28				

Qualitative insights from semi-structured interviews further complemented the quantitative findings. As detailed in Table 5, thematic analysis identified five recurring themes across sectors: automated compliance enforcement, AI-driven metadata discovery, cross-platform data harmonization, cross-departmental synchronization, and dynamic data quality assessment. The most frequently mentioned theme was data harmonization, cited by 88% of participants, especially within healthcare and pharmaceutical sectors. Experts also emphasized the role of AI in enabling real-time policy enforcement and metadata classification, which are critical in regulatory compliance and large-scale integration.

Table 5: Thematic insights from expert interviews

Theme	Mentioned by (%)	Sectoral Relevance	Application Area	Sample Quote/Insight
Automated Compliance Enforcement	79	All Sectors	Regulatory and Legal Teams	"AI allows us to monitor compliance in real-time."
AI-Driven Metadata Discovery	65	Pharma, Financial	Metadata Repositories	"Machine learning automatically tags sensitive data."
Data Harmonization Across Platforms	88	Healthcare, Pharma	Interoperability, EMRs	"Integrating data from EHRs is seamless with AI tools."
Cross-Departmental Synchronization	72	Financial, Healthcare	Enterprise Data Platforms	"Our departments now access a unified data view."
Dynamic Data Quality Assessment	70	All Sectors	Data Cleaning & Validation	"ML models flag anomalies faster

				than traditional tools."
Context-Aware Data Governance Policies	61	Financial	Risk & Fraud Detection	"AI adapts rules based on transaction context."

Figure 1 provides a visual comparison of AI/ML adoption rates and data accuracy improvements across the three sectors. It clearly illustrates that sectors with higher adoption of AI/ML, such as finance, also experience greater improvements in data accuracy, supporting the argument that AI/ML plays a pivotal role in strengthening data governance frameworks.

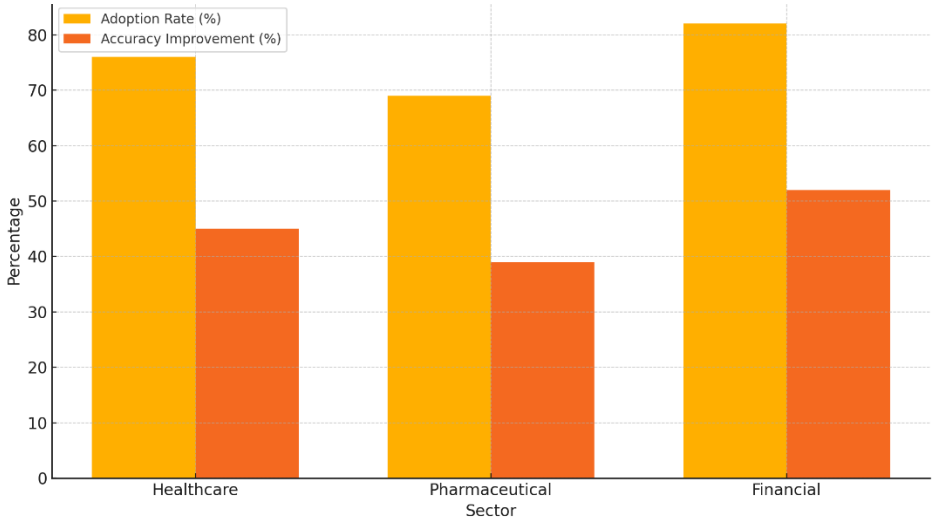

Figure 1: AI/ML adoption and data accuracy improvement by sector

Discussion

The findings of this study confirm that Artificial Intelligence (AI) and Machine Learning (ML) are fundamentally reshaping the landscape of data governance and Master Data Management (MDM), particularly in data-sensitive and compliance-intensive sectors such as healthcare, pharmaceuticals, and finance. The analysis reveals strong evidence of sectoral variation in AI/ML adoption, perceived benefits, and strategic outcomes—highlighting both the potential and the challenges of these transformative technologies (Petrick & Shomron, 2022).

One of the most compelling observations is the high adoption rate of AI/ML tools in the financial sector (82%), followed by healthcare (76%) and pharmaceuticals (69%) as shown in Table 2. This reflects not only the technological maturity of the financial industry but also its agility in responding to regulatory pressures and customer-centric innovation. The substantial improvement in data accuracy (52%) and MDM efficiency gain (62%) in finance supports the assertion that when properly leveraged, AI/ML can significantly enhance data quality and operational efficiency. In comparison, while the pharmaceutical sector shows promise, its relatively slower adoption suggests a need for better infrastructure, workforce upskilling, and regulatory clarity (George, 2024).

The results also highlight that AI/ML adoption correlates strongly with reductions in time to compliance, cost savings, and enhanced governance automation (Habibi et al., 2024). These improvements are particularly valuable in healthcare and pharmaceutical industries where complex regulatory frameworks like HIPAA and FDA regulations demand high accuracy and traceability (Zhu et al., 2021). In healthcare, for instance, an average compliance time of just eight months (Table 2) represents a significant efficiency gain in managing electronic health records, interoperability, and privacy protocols.

Perceptions captured through the survey further validate the benefits of AI/ML across all sectors. As

shown in Table 3, over 80% of respondents from the financial sector agreed that AI/ML enhances decision-making, enables real-time monitoring, and supports faster compliance. This positive sentiment is echoed in the healthcare sector, which reported high agreement on improved data accuracy (85%) and better integration of patient records. These results reinforce prior literature suggesting that AI/ML technologies are not only operational tools but also strategic enablers for innovation and customer-centric service delivery (Crouzet et al., 2024).

The statistical analysis (Table 4) confirms that the return on investment (ROI) from AI/ML-driven data governance varies significantly across sectors ($F = 29.07$, $p < 0.001$). This variation can be attributed to several factors, including organizational readiness, sector-specific regulatory complexity, and the maturity of digital infrastructure (Suter et al., 2023). The financial sector's higher ROI aligns with its early adoption of automation and advanced analytics, whereas healthcare and pharma are still in transitional phases, where organizational change and integration costs may initially suppress financial returns (Abubakar et al., 2020).

Qualitative insights from interviews enrich the quantitative findings, revealing key themes such as automated compliance enforcement, AI-driven metadata discovery, and dynamic data quality assessment (Table 5). These themes reflect how AI/ML tools are being used not only to clean and standardize data but also to provide contextual intelligence that helps organizations make faster, smarter decisions (Han et al., 2024). In healthcare and pharma, participants stressed the importance of data harmonization across platforms—a necessary step for enabling effective patient care and global clinical trials. In finance, emphasis was placed on AI's role in detecting fraud, automating audit trails, and ensuring real-time compliance (Soni et al., 2023).

Figure 1 visually reinforces these sectoral trends, showing that improvements in data accuracy directly track with AI/ML adoption levels. This suggests a strong case for further investment in intelligent data systems as a means of optimizing strategic outcomes, reducing compliance risks, and enabling data-driven innovation (Barrett et al., 2023).

The discussion highlights that while AI/ML tools are already delivering measurable benefits in data governance and MDM, their full potential depends on sector-specific adoption strategies, supportive policy frameworks, and continued research into ethical AI implementation (Lipizzi, 2024). Healthcare, pharma, and financial institutions stand to gain immensely from these technologies, provided they align innovation with regulatory compliance, workforce training, and cross-functional collaboration (Hewage et al., 2022). This study contributes to the growing body of evidence that AI/ML is not just a technological shift—but a strategic imperative in modern data governance.

Conclusion

This study underscores the transformative potential of AI and Machine Learning in revolutionizing data governance and Master Data Management (MDM) within the healthcare, pharmaceutical, and financial sectors. Through a mixed-methods approach, it became evident that AI/ML integration significantly improves data accuracy, accelerates compliance, and enhances decision-making processes. The financial sector demonstrated the highest adoption and return on investment, highlighting its technological agility and data-centric maturity. Meanwhile, healthcare and pharmaceutical sectors, though progressing steadily, face challenges in infrastructure, integration, and regulatory navigation. Thematic insights from industry experts reinforced the critical role of AI/ML in automating compliance, harmonizing data across platforms, and enabling real-time analytics. Statistical analysis further confirmed the sectoral variance in outcomes, emphasizing the need for tailored adoption strategies. Overall, this research confirms that AI/ML-driven data strategies are not merely operational upgrades—they are essential components of competitive advantage, regulatory resilience, and digital transformation. For organizations navigating vast and complex data ecosystems, especially in high-stakes industries, investing in intelligent data governance frameworks is no longer optional but imperative. Future research should explore long-term impacts, ethical considerations, and cross-industry benchmarks to guide scalable and responsible AI/ML adoption in enterprise data strategy.

References

[1] Abubakar, A. I., Omeke, K. G., Ozturk, M., Hussain, S., & Imran, M. A. (2020). The role of artificial intelligence driven 5G networks in COVID-19 outbreak: opportunities, challenges,

and future outlook. *Frontiers in Communications and Networks*, *1*, 575065.

[2] Barrett, J. S., Oskoui, S. E., Russell, S., & Borens, A. (2023). Digital Research Environment (DRE)-enabled Artificial Intelligence (AI) to facilitate early stage drug development. *Frontiers in Pharmacology*, *14*, 1115356.

[3] Barrett, J. S., Oskoui, S. E., Russell, S., & Borens, A. (2023). Digital Research Environment (DRE)-enabled Artificial Intelligence (AI) to facilitate early stage drug development. *Frontiers in Pharmacology*, *14*, 1115356.

[4] Camacho, N. G. (2024). Unlocking the potential of AI/ML in DevSecOps: effective strategies and optimal practices. *Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023*, *3*(1), 106-115.

[5] Crouzet, A., Lopez, N., Riss Yaw, B., Lepelletier, Y., & Demange, L. (2024). The Millennia-Long Development of Drugs Associated with the 80-Year-Old Artificial Intelligence Story: The Therapeutic Big Bang?. *Molecules*, *29*(12), 2716.

[6] George, A. S. (2024). Democratizing Compute Power: The Rise of Computation as a Commodity and its Impacts. *Partners Universal Innovative Research Publication*, *2*(3), 57-74.

[7] Habibi, M. A., Yilma, G. M., Fattore, U., Costa-Pérez, X., & Schotten, H. D. (2024). Unlocking O-RAN potential: How management data analytics enhances SMO capabilities?. *IEEE Open Journal of the Communications Society*.

[8] Han, S., Lee, J. E., Kang, S., So, M., Jin, H., Lee, J. H., ... & Lee, Y. S. (2024). Standigm ASK™: knowledge graph and artificial intelligence platform applied to target discovery in idiopathic pulmonary fibrosis. *Briefings in Bioinformatics*, *25*(2), bbae035.

[9] Hashio, M., Kolodziej, M., Olds, M., Samuel, E., Ravindran, S., Mohammed, A., & Makowski, M. (2024). GSK Data Strategy & Management Whitepaper 2023: Transformation from a Data Management to a Data Science Enabled Organization–Where Are We Heading?. *Journal of the Society for Clinical Data Management*, *4*(1), 9.

[10] Hewage, C. T., Khattak, S. K., Ahmad, A., Mallikarachchi, T., Ukwandu, E., & Bentotahewa, V. (2022). Multimedia privacy and security landscape in the wake of ai/ml. *Social Media Analytics, Strategies and Governance*, 203-228.

[11] Jena, G. K., Patra, C. N., Jammula, S., Rana, R., & Chand, S. (2024). Artificial intelligence and machine learning implemented drug delivery systems: a paradigm shift in the pharmaceutical industry. *Journal of Bio-X Research*, *7*, 0016.

[12] Lipizzi, C. (2024). Impacts on Specific Industries. In *Societal Impacts of Artificial Intelligence and Machine Learning* (pp. 71-104). Cham: Springer International Publishing.

[13] Petrick, L. M., & Shomron, N. (2022). AI/ML-driven advances in untargeted metabolomics and exposomics for biomedical applications. *Cell Reports Physical Science*, *3*(7).

[14] Soni, V., Bartelo, N., Schweickart, A., Chawla, Y., Dutta, A., & Jain, S. (2023). Future perspectives of metabolomics: gaps, planning, and recommendations. In *Metabolomics: Recent Advances and Future Applications* (pp. 479-512). Cham: Springer International Publishing.

[15] Suter, F., Da Silva, R. F., Gainaru, A., & Klasky, S. (2023, October). Driving next-generation workflows from the data plane. In *2023 IEEE 19th International Conference on e-Science (e-Science)* (pp. 1-10). IEEE.

[16] Vijay, A. J., William, B. N. J., Haruna, A. A., & Prasad, D. D. (2024). Exploring the synergy of IIoT, AI, and data analytics in Industry 6.0. In *Industry 6.0* (pp. 1-36). CRC Press.

[17] Zhu, X., Ninh, A., Zhao, H., & Liu, Z. (2021). Demand forecasting with supply-chain information and machine learning: Evidence in the pharmaceutical industry. *Production and Operations Management*, *30*(9), 3231-3252.