Artificial Intelligence and Operations: A Foundational Framework of Emerging Research and Practice

Tinglong Dai* Jayashankar M. Swaminathan†

*Carey Business School, Johns Hopkins University, Baltimore, Maryland 21202; Data Science and AI Institute, Johns Hopkins University, Baltimore, Maryland 21218, dai@jhu.edu (Corresponding Author)

†Kenan–Flagler Business School, University of North Carolina at Chapel Hill, North Carolina 27599, msj@unc.edu

---

Abstract. Artificial intelligence (AI) is poised to reshape operations across industries. Yet its real-world impact reveals a jagged and uneven implementation frontier. To make sense of this emerging landscape, we develop a foundational framework that synthesizes research and practice at the intersection of AI and Operations Management (OM), anchored in three interdependent pillars: (1) AI for OM, (2) OM for AI, and (3) Human–AI Interaction. First, AI for OM analyzes how AI enhances core operational processes, including design, procurement, production, and delivery. Second, OM for AI argues that scaling AI safely and effectively stands to benefit from core OM principles, including workflow design, capacity management, process control, drift detection, and continuous improvement, all of which are central to AI development and deployment. Third, Human–AI Interaction emphasizes the role of trust, incentives, and organizational design in mediating how humans and machines learn from and collaborate with each other. This triadic framework provides a foundation for organizing research on AI and OM and offers practical guidance for integrating AI into business and societal systems.

Key words: Artificial intelligence, operations management, supply chain, human-AI interaction

History: Received: August 2025; accepted: December 2025 by Kalyan Singhal after two revisions.

---

1. Introduction

Operations Management (OM) has continually adapted to major technological and social shifts, each reshaping how organizations create, deliver, and capture value (Lepore 2009; Singhal and Singhal 2022). The First Industrial Revolution introduced factory production; the Second added mass production and scientific management. The late 20th century brought quality management and lean production (Womack et al. 1990), while the new millennium saw globally dispersed supply chains and mass customization enabled by IT and quantitative models (Lee 2004; Swaminathan and Tayur 1998). Following that, sustainability and resilience have come to the forefront of OM, expanding the traditional cost-quality-speed paradigm (Lee and Tang 2018; Kleindorfer et al. 2005). Most recent OM research continues this trajectory, increasingly engaging with transformative technologies and the grand societal challenges they pose (Flammer and Loch 2024; Swaminathan 2025). At its core, OM's history reflects a continual *alignment* of day-to-day operations with the era's dominant priorities and technologies—be it steam power, assembly lines, computing, or globalization. Looking back, the lesson is clear: success comes from aligning an organization's resources, processes, and incentives with its strategic objectives and customer needs. OM is, in essence, the science of alignment.

Artificial Intelligence (AI) has experienced a parallel revolution. The term "AI" itself was coined in 1955 (McCarthy et al. 2006), as pioneers like Allen Newell and Herbert Simon built programs (e.g., the Logic Theorist) that could mimic aspects of human problem-solving (Simon 1987). Early AI systems in the 1960s–70s were largely *rule-based*, encoding expert knowledge into `IF-THEN` decision rules. A classic example is MYCIN, a 1970s medical expert system that could diagnose infections from symptoms and lab results—a tour de force of that era's AI (Shortliffe 1976). These systems demonstrated the potential of "handcrafted" intelligence but were limited in scope and brittle outside their narrow domain. By the 1980s, a divergence had appeared: operations scholars focused on highly structured optimization and analytical models, whereas AI researchers pursued heuristic and symbolic reasoning. In a prescient commentary, Herbert Simon lamented that "after about 1960, AI and OR went their separate ways; whole new generations of scientists trained in each of these disciplines were largely unacquainted with the techniques provided by the other" and he called for renewed collaboration between the two fields (Simon 1987). Notably, Simon's call—"two heads are better than one"—urged that the computation and algorithmic prowess of AI be married with the modeling and problem-structuring genius of the operations scholarship.

AI soon entered a learning-centric phase. The late 1990s and 2000s saw the dominance of *learning-based* AI: algorithms that improve through data. Machine learning eclipsed expert systems, thanks to increasing computing power and massive datasets (Katz 2012). Predictive models and neural networks began to outperform human-crafted rules in domains like forecasting and image recognition. OM researchers started to take notice, as these data-driven methods offered new ways to predict demand, personalize services, or detect quality issues. Yet integrating machine learning into day-to-day operations proved nontrivial; predictions are useful only if embedded into operational decision-making (Simchi-Levi 2014). This gave rise to the mantra "from predictive to prescriptive analytics" (Bertsimas and Kallus 2020), emphasizing that the true value of AI lies in informing or automating decisions, not just generating predictive insights. In the 2010s, deep learning and advanced algorithms pushed AI to new frontiers: IBM's Watson beat human champions on Jeopardy (Ferrucci et al. 2010), AlphaGo mastered the game of Go through reinforcement learning (Silver et al. 2016), and autonomous vehicles—remotely monitored by humans—began to navigate our streets (Metz et al. 2024).[1]

Each breakthrough carried implications for OM: if AI could drive a car or defeat a Go master, why not manage a warehouse or optimize a factory schedule? Indeed, reinforcement learning started to be applied to inventory control and dynamic pricing. However, initial enthusiasm sometimes outpaced reality; studies note mixed results when deploying such AI in real-world supply chains, pointing to a gap between algorithmic advances and practical impact (Gijsbrechts et al. 2026). These observations echo a broader "AI paradox": despite enormous potential, the productivity gains from AI have been slow to materialize at the economy-wide level (Agrawal et al. 2022; Brynjolfsson et al. 2017). Explanations include implementation lags and organizational frictions, highlighting that technology alone is not enough; it must be effectively aligned with business processes and human workflows.

Today, we are on the cusp of a new AI era, driven by generative AI (GenAI) and large language models (LLMs). A defining moment came in late 2022 with the public release of OpenAI's ChatGPT, an AI assistant capable of natural conversation, coding, and answering broad knowledge queries. In less than two months, ChatGPT reached 100 million users, illustrating how rapidly AI can diffuse into practice. Meanwhile, foundation models—general-purpose models trained on terabytes of data (Bommasani et al. 2021)—are enabling AI to perform tasks never before automated, from drafting engineering designs to managing customer service chats. These GenAI systems represent a qualitative leap: they are not narrowly programmed for a single task but rather can be adapted to myriad tasks with relatively little extra data. For OM, this raises exciting possibilities. One can

[1] In this paper, we use "AI" to refer to technologies that perform tasks typically requiring human intelligence, and we distinguish three forms most relevant to operations: (i) predictive AI, which learns patterns from data to forecast, classify, or detect anomalies; (ii) prescriptive (decision automation) AI, which recommends and/or executes actions—often by combining learning with optimization or reinforcement learning—to automate or support decisions; and (iii) generative AI, which produces novel content or designs. For clarity, our use of "AI" includes decision automation methods when they integrate learning and optimization, rather than conventional optimizers.

envision LLM-based assistants for supply chain managers that automatically analyze procurement contracts or recommend factory scheduling adjustments by synthesizing operational data and best practices (Simchi-Levi et al. 2026). At the same time, these powerful AI tools pose new challenges: they can make errors or "hallucinate" (Chomsky et al. 2023), they disrupt traditional job roles and have the potential to cause massive job losses (VandeHei and Allen 2025), and their effective use may require rethinking processes (Cachon 2026). In a nutshell, AI has evolved from a set of specialized tools to a pervasive, general capability—one that could transform OM, *if* we properly understand the synergies and trade-offs involved (Tayur 2026).

Today, AI's emergence as both a disruptor and a tool for OM prompts us to revisit OM's alignment logic. By "alignment logic," we refer to OM's fundamental approach of designing systems so that capacity matches demand, supply chains match product and market characteristics, incentives match performance goals, and so on. AI can threaten this alignment if introduced carelessly—imagine an AI-driven demand forecast that is exquisitely accurate but is coupled with an inflexible production system that cannot adjust, leading to misalignment. Conversely, AI can greatly enhance alignment by providing better information and control; for example, an industrial AI system that optimizes energy usage aligns operational efficiency with sustainability goals.

In many ways, AI is both a stress test and a beneficiary of OM principles. It is a stress test because it changes the art of the possible, requiring organizations to realign workflows and decision rights. At the same time, AI is a beneficiary of alignment because deploying AI effectively demands careful operational design: data must be aligned with training needs, algorithms must be aligned with business objectives, and AI decision-making must be aligned with human values and ethics (a point the AI safety community terms the "AI alignment problem"). The late Herbert Simon's vision of collaboration between AI and operations (Simon 1987) is more pertinent than ever (Wiberg et al. 2025). It suggests that realizing AI's promise in business and society lies in integrating AI into OM's alignment framework. Indeed, early evidence from industry indicates that firms succeeding with AI are those that re-engineer their processes and upskill their workforce in tandem with AI adoption, rather than treating AI as a plug-and-play tool (Agrawal et al. 2022). The need for a new paradigm bridging AI and OM is clear.

A tighter interplay between AI technologies and OM logic is opening new research frontiers. Rather than viewing AI purely as a predictive tool or OM merely as an efficiency science, we see them as parts of a virtuous cycle: AI enhances prediction and autonomy, while OM ensures these capabilities create value through effective system design. We frame this convergence around three pillars: (1) AI for OM, (2) OM for AI, and (3) Human–AI Interaction.

The first pillar, *AI for OM*, focuses on using AI to tackle OM challenges: designing products and processes, managing supply chains, scheduling production, and delivering services. Tools like machine learning, computer vision, reinforcement learning, and GenAI are now applied across the operations value chain to improve speed, cost, quality, and flexibility. We structure this pillar around the Design–Procure–Produce–Deliver (DPPD) framework: in *Design*, AI supports generative design and forecasting; in *Procure*, it enables autonomous sourcing and data-driven supplier selection; in *Produce*, it powers robotics, predictive maintenance, and quality assurance; and in *Deliver*, it optimizes routing, logistics, and real-time monitoring. Ultimately, performance depends on aligning AI not just with data but also with human workflows, incentives, and decision rights.

The second pillar, *OM for AI*, flips the lens: it considers how OM can inform the development, scaling, and governance of AI systems. Here, data pipelines and model deployments become operational processes to be managed. OM insights into modular design, process flow, and resource allocation can guide how firms source training data, choose between internal development and external procurement, or streamline continuous model updates.

The third pillar, *Human–AI Interaction*, recognizes that decision-making does not occur in a vacuum. Humans must interpret, trust, and sometimes override AI output. This pillar draws from behavioral OM to examine how cognitive load, algorithm aversion, and organizational incentives shape AI use. Poor interface design, misaligned incentives, or unclear accountability can blunt AI's value. In contrast, hybrid systems, where roles, rules, and responsibilities are purposefully defined, enable AI and human judgment to complement each other.

Table 1 Three Pillars at a Glance.

Pillar	**Scope and representative examples**	**Meaning and relationships**
AI for OM	**Design:** generative design for products and services; **Procure:** supplier selection and dynamic sourcing; **Produce:** predictive maintenance, scheduling, robotics; **Deliver:** routing, fulfillment, virtual agents for service delivery.	Uses AI to enhance/automate tasks across DPPD. Creates demands on lifecycle operations (OM for AI) and depends on human oversight/adoption (Human–AI Interaction).
OM for AI	Data pipelines, labeling, training, deployment, monitoring/retraining, governance, and ethical oversight.	Applies OM principles to the design, scaling, and governance of AI systems; enables DPPD applications that incorporate human collaboration and social processes.
Human–AI Interaction	Trust calibration, override policies, social norms, routines, ownership, reskilling, and change management.	Behavioral/organizational factors that mediate AI success; link DPPD outcomes (AI for OM) to lifecycle processes (OM for AI).

Our triadic framework offers a lens on AI–OM co-evolution. *AI for OM* examines how AI applications may affect operational tasks—and under what conditions they improve, leave unchanged, or worsen performance. *OM for AI* concerns how operations principles can support the creation, scaling, and governance of AI systems, including potential failure modes. *Human–AI Interaction* studies how humans work with AI, including when trust is earned or eroded, how reliance is calibrated, and when overrides are warranted. The pillars are interdependent: applications generate lifecycle demands and governance questions, and human factors condition outcomes across both. Throughout, we treat effects as contingent on context, data, incentives, and design rather than assuming net benefits a priori.

The time is ripe to define a new research paradigm for the co-evolution of AI and OM; this article outlines this emerging paradigm. First, we examine AI for OM (Section 2), surveying applications to core operational problems and highlighting implementation challenges. Second, we explore OM for AI (Section 3), showing how operations principles can address AI deployment issues and productivity paradoxes. Third, we explore Human–AI Interaction (Section 4), examining when human-AI collaboration enhances versus hinders operational performance. Section 5 outlines several future research directions for OM scholars. Section 6 concludes with managerial implications. Throughout the article, we use cautious language (e.g., "can" and "may") to reflect heterogeneous and context-dependent effects of AI in operations.

2. AI for OM

AI is permeating the full spectrum of operations, from the way products are conceived to how they are delivered to customers. We organize this section along the Design–Procure–Produce–Deliver (DPPD) framework, examining how AI reshapes each stage of the operations lifecycle, treating performance effects as empirical: AI can improve efficiency or quality in some settings, have no effect in others, or introduce new frictions; accordingly, we urge attention to be paid to the underlying mechanisms, relevant boundary conditions, and the possibility of mixed evidence.

AI comes in many forms and shapes. We now indicate the predominant AI type in each subsection: Section 2.1 (generative AI) for design; Section 2.2 (mix of predictive and prescriptive AI) for procurement; Section 2.3 (predictive AI) for production and maintenance; and Section 2.4 (prescriptive/decision automation AI) for delivery, logistics, and control.

2.1. Design

AI transforms the design of new products and services through generative tools and data-driven innovation. In generative design, engineers specify high-level goals and constraints, and AI algorithms autonomously generate and evaluate thousands of alternatives. This process often uncovers novel, high-performance designs that human engineers might not have considered. For instance, General Motors partnered with Autodesk to design a seat bracket, using AI to explore over 150 permutations and ultimately producing a 3D-printed component that consolidated eight parts into one. The final design was 40% lighter and 20% stronger than the original (Danon 2018), illustrating AI's ability to achieve outcomes beyond conventional design limits. When combined with additive manufacturing, generative design enables the production of complex geometries, such as organic lattice structures, that were previously infeasible. The result is faster prototyping cycles and components that are lighter, stronger, and more cost-effective. The same logic extends to intangible offerings: generative AI models can propose novel drug molecules (Chatterjee and Dai 2025) and can even help redesign educational curricula (Fazlollahi et al. 2023).

Beyond generative design, AI enables a data-driven approach to exploring and refining products. Machine learning models analyze customer feedback, usage patterns, and market trends to surface latent needs and guide iterative development. GenAI can generate design assets, prototypes, or code, speeding up early-stage exploration and expanding the range of possibilities. This AI-augmented search process echoes long-standing OM insights: diverse teams produce higher-quality ideas in complex, cross-functional settings (Kavadias and Sommer 2009), and well-structured feedback loops can accelerate development without sacrificing quality (Loch and Terwiesch 1999). Across industries ranging from automotive to consumer electronics, AI is shifting design from a human-driven trial-and-error process to a collaborative search for optimal solutions. Designers increasingly act as problem framers, specifying objectives and constraints, while AI explores the solution space and surfaces innovative candidates (Marion et al. 2024).

Whereas the promise is clear, AI in design brings challenges. AI-generated outputs can be difficult to interpret or manufacture without additional processing. Engineers must ensure that AI's proposals meet real-world requirements, such as safety and regulatory standards, that may not be captured in the algorithm's objective function. Biases or omissions in training data or constraints can also lead to suboptimal or infeasible designs. Finally, as AI takes a larger role in creative processes, designers must learn to collaborate with algorithms—maintaining critical thinking, ensuring feasibility, and validating AI-proposed solutions (Dai and Taylor 2025).

2.2. Procurement

Procurement stands to gain substantially from AI through improved supplier selection, automated negotiation, and smarter spend analytics. Traditionally, sourcing managers evaluate suppliers and negotiate contracts largely based on experience and static criteria. AI can augment these tasks by rapidly processing vast amounts of supplier data, such as pricing, quality, reliability, and financial risk, by automating routine communication. For example, natural-language chatbots can handle initial RFQs (requests for quotations) and supplier inquiries, scaling up the number of suppliers considered and the speed of information exchange (Van Hoek et al. 2022). AI systems can also predict supplier risks (e.g., likelihood of delay or failure) by learning from historical performance data, news feeds, and even macroeconomic indicators, thus helping procurement officers proactively mitigate supply chain disruptions.

A key development in procurement is AI-assisted decision support in supplier negotiations. Cui et al. (2022) study a B2B platform where the buyer sometimes requested quotes via a chatbot (automation) and sometimes also signaled that an algorithm would select the winner based on price. When no algorithmic evaluation was mentioned, suppliers quoted higher prices to the chatbot than to human buyers, exploiting "naïve" automation. By contrast, when the message stated that a smart system would evaluate quotes, suppliers lowered prices and competed more aggressively. The value thus comes from pairing automation with transparent algorithmic selection: replacing human interaction with a bot alone can backfire. Managers should implement AI to streamline communication *and* credibly signal a meritocratic, data-driven selection process.

Beyond negotiations, AI is transforming procurement through applications in spend analytics, contract management, and supplier auctions. Machine learning algorithms can categorize and analyze organizational spending patterns to identify maverick expenditures, consolidation opportunities, and optimal ordering schedules (McKinsey 2024). Natural language processing enables the automated review of contract terms at scale, facilitating the detection of unfavorable clauses and compliance issues (SAP 2025). In supplier auctions, reinforcement learning agents have been developed to learn bidding strategies, dynamically adjusting bids to achieve cost savings. These agents can simulate numerous auction scenarios to refine their tactics, a task impractical for human negotiators (Wang et al. 2017). AI-powered procurement systems can reduce processing costs by up to 30% and enhance supplier compliance rates by 15–30% (C-Suite Strategy 2024).

Integrating AI into procurement carries risks. Vendor lock-in may arise if training data are limited to incumbent suppliers, leading algorithms to favor familiar vendors and suppress competition. Opaque selection criteria can also undermine managers' ability to interpret or defend decisions, eroding credibility and trust. From the supplier side, algorithmic processes may appear transactional, alienating partners that compete on service or quality rather than price. Mitigating these risks requires multi-attribute evaluation (price, quality, sustainability), human oversight of strategic purchases, and transparent communication about AI's role in decision-making.

2.3. Production

Within production operations, AI has enabled notable gains in efficiency, quality, and flexibility. A leading application is predictive maintenance (Olsen and Tomlin 2020). By analyzing sensor data on vibration, temperature, and pressure, algorithms can forecast equipment failures and trigger proactive interventions. Relative to run-to-failure or fixed-interval regimes, predictive maintenance reduces unplanned downtime and lowers costs. Empirical studies suggest breakdowns fall by up to 50 percent and equipment lifespans extend by 20–40 percent (Nucleus Research 2023). These gains reflect AI's ability to detect subtle, high-dimensional patterns—such as nonlinear interactions of load and temperature—that operators often miss. Major manufacturers now deploy industrial AI platforms to monitor assets like jet engines and turbines, dispatching crews only when needed (Siemens 2024; Wuest et al. 2020). This shift from reactive to predictive regimes both cuts costs and stabilizes schedules by allowing downtime to be planned during off-peak periods.

Another way in which AI shapes production is quality assurance. Computer-vision models trained on large labeled image sets now inspect items at speeds and resolutions humans cannot match, enabling semiconductor fabs to flag sub-micron wafer defects in real time (Saqlain et al. 2020). Similar architectures detect fabric flaws on textile lines and reject non-conforming parts upstream, cutting rework and scrap. In process industries, multivariate learning systems ingest high-frequency sensor streams to predict whether a pharmaceutical batch will meet purity targets well before completion, giving managers time to intervene (Tulsyan et al. 2018). Across services, language-model monitors parse live call-center dialogues, surfacing conversations where sentiment deteriorates so supervisors can recover the interaction, paralleling the role of automated vision on the factory floor. Collectively, these applications push defect rates toward Six-Sigma benchmarks by extending inspection from a discrete post-production gate to an always-on predictive layer, a progression made possible by advances in deep learning (LeCun et al. 2015).

Perhaps the most visible impact of AI in production is the rise of robotics and automation. Industrial robots have long been used for repetitive tasks, but they are becoming smarter and more autonomous thanks to AI. Modern robots use computer vision and reinforcement learning to handle variability in tasks—for example, a robotic arm equipped with an AI vision system can pick and sort mixed items, not just identical parts, by learning to recognize objects on the fly. Collaborative robots ("cobots") work alongside humans on assembly lines, using AI to sense human presence and adapt their motions safely. The adoption of industrial robots continues to accelerate globally, with the operational stock reaching 3.9 million units by the end of 2022, according to the International Federation of Robotics (2023). Looking ahead, the IFR projects that annual robot installations will grow at a compound annual growth rate (CAGR) of 7%, potentially reaching over 700,000 new installations annually by 2026. AI plays a dual role here—both in the robots themselves (for perception, decision-making, and learning new tasks) and in the coordination of robotic fleets. Operations managers increasingly rely on AI-based scheduling and control algorithms to orchestrate dozens or even hundreds of robots in a facility, assigning tasks, preventing bottlenecks, and dynamically reallocating machines as priorities change.

AI's impact on production is not limited to manufacturing; it extends across service operations. A survey of 2,547 U.S. hospitals found 65% use AI—mainly for inpatient risk prediction (92%), outpatient follow-up (79%), and scheduling (51%)—with adoption skewed toward hospitals in systems and those with high margins, while rural and underserved hospitals lag behind (Nong et al.

2025). Retailers use AI for demand forecasting, marketing, and inventory optimization, reporting measurable gains in productivity and margins (McKinsey 2024). Access to a GenAI-based conversational assistant increases customer service agent productivity by 15%, primarily by helping less experienced agents learn from the best practices of top performers; experienced agents see little productivity gain and a modest decline in quality (Brynjolfsson et al. 2025).

Integrating AI into production systems presents challenges. Retrofitting legacy lines with sensors and connectivity entails sizable fixed costs, so productivity gains from shop-floor AI often arrive only after a protracted implementation lag (Brynjolfsson et al. 2017). Realizing those gains also hinges on human capital: technicians must learn to interpret algorithmic alerts, and even a single visible error can trigger algorithm aversion (Dietvorst et al. 2015), a topic we elaborate on in Section 4. Safety and uptime impose additional discipline. Without continuous monitoring, hidden technical debt accumulates; a misclassified maintenance prediction or a control glitch in a collaborative robot can idle an entire line (Sculley et al. 2015). This is where OM principles are applicable (see Section 3). Robust fail-safes—human vetting during ramp-up, redundant control loops, and climate-aware model refresh—are integral to sustaining reliable, bias-resilient AI at scale.

2.4. Delivery

The final stage of the DPPD framework is delivery: getting products or services to the end customer. Here, too, AI is a game-changer, driving efficiencies in logistics, personalization in services, and new distribution models. In supply chain logistics, firms are deploying AI for dynamic routing and distribution network optimization.[2] These approaches are conceptually aligned with a rich literature on e-commerce and supply chain management (Cachon and Netessine 2006; Swaminathan and Tayur 2003). Machine learning models ingest traffic data, weather forecasts, and delivery histories to optimize delivery routes in real time, cutting transit times and fuel usage. For instance, UPS's ORION system (while rooted in operations research) has been AI-augmented to better predict package flow and adjust routes daily, reportedly saving millions of miles driven (Fickenscher 2021). AI can solve complex routing problems that were traditionally intractable; reinforcement learning algorithms, for example, have been developed to tackle vehicle routing with thousands of stops, learning heuristics that outperform classical methods on large-scale instances (Wang et al. 2021). These advances translate into faster deliveries and lower logistics costs.

Autonomous delivery is an active frontier for AI in operations. Firms are piloting self-driving trucks, sidewalk robots, and drones for last-mile service (Stewart 2017). AI provides perception, navigation, and control. Workhorse's HorseFly secured its first commercial purchase orders in 2023 for

[2] These applications qualify as AI because they leverage learning-based heuristics and/or adaptive feedback to drive optimization at scale. This integrated view is consistent with the long-standing aspiration to combine artificial intelligence and operations research in decision systems (Simchi-Levi et al. 2025; Wiberg et al. 2025).

a truck-integrated drone that serves remote drops while the driver continues the route (Workhorse 2023); the "flying sidekick" model coordinates truck–drone tasks via optimization to reduce time (Murray and Chu 2015). Pilot programs, such as UPS's tests of truck-launched drones, have provided early evidence of efficiency improvement (Stewart 2017). Ground robots show similar promise: Starship Technologies reports over 8 million autonomous deliveries across more than 150 locations in six countries (Oitzman 2025).

AI helps optimize service delivery across a range of sectors. In ride-hailing, AI systems match riders and drivers and set dynamic prices, executing thousands of dispatches and millions of price changes per minute (Yan et al. 2020). These tools balance demand and supply, cut waits, and prevent peak-time breakdowns. In the case of healthcare, the FDA-cleared LumineticsCore, for instance, autonomously analyzes retinal images to detect diabetic retinopathy, triaging patients in real time without human oversight (FDA 2018). By reallocating diagnostic effort—clearing patients without pathology and referring only those who require specialist evaluation—LumineticsCore increases screening capacity and clinic productivity. In a randomized controlled trial, the system enabled a 40% increase in the number of patients seen per hour by retina specialists (Abràmoff et al. 2023).

The benefits of AI in delivery include increased speed, higher asset utilization (trucks driving full or on optimized routes), and more personalized service. Customers increasingly expect rapid and convenient delivery, and AI is a key enabler in meeting these expectations in a cost-efficient fashion. Furthermore, AI's predictive capabilities allow contingency planning: logistics AI tools can forecast demand surges (say, holiday shopping or regional events) and proactively reroute shipments or reposition inventory before bottlenecks form. This predictive logistics reduces the incidence of stockouts or delayed deliveries, enhancing reliability.

Implementing AI in delivery faces several hurdles. Autonomous vehicles and drones must meet strict safety and regulatory standards; even small perception or navigation failures can erode trust and stall approval (Stanley et al. 2024). For example, in October 2023, California regulators suspended driverless permits for GM Cruise following a pedestrian injury, prompting a nationwide pause and multi-agency reviews (Jin and Shepardson 2023).Integration is nontrivial: AI must interoperate with legacy routing, warehouse, and carrier systems, while supply chains remain exposed to shocks—natural disasters, factory fires, and geopolitical events—that disrupt even sophisticated networks (Simchi-Levi et al. 2015); for a synthesis of how AI can bolster resilience in this environment, see Cohen and Tang (2024). Adoption also turns on worker acceptance: employees may initially view AI as surveillance rather than support (Dai and Tayur 2022). Effective programs pair better tooling with training that demonstrably reduces workload and planning stress, building acceptance and a more adaptive culture (Lanteri et al. 2023).

Taken together, AI is integrating operations by dissolving boundaries across design–procure–produce–deliver. Real-time delivery data feed back into design—for example, optimizing packaging for AI-managed warehouses—creating adaptive supply chains and accelerating OM's shift to data-driven decision making (Mišić and Perakis 2020). Learning systems anticipate demand, schedule maintenance, and reroute logistics in real time, offering potential productivity gains (Agrawal et al. 2024). Looking ahead, multi-agent systems that span functional and firm boundaries will matter, coordinating actions among heterogeneous AI and human agents (Swaminathan et al. 1998).

3. OM for AI

Just as AI is reshaping OM practice, OM principles are key to the effective development, deployment, and scaling of AI. We conceptualize the AI lifecycle—data, training, deployment, monitoring and retraining, and governance—as a set of interdependent operations that can either support or undermine system performance. OM tools can help manage the attendant risks and trade-offs.

In this section, we explore how the OM mindset and toolkit—honed over decades to improve processes, quality, and efficiency—can be applied to the lifecycle of AI itself. This perspective includes managing data as an operational input, streamlining the model development pipeline, ensuring the quality of AI outputs through rigorous monitoring (akin to quality control), and designing processes for continuous improvement and governance of AI systems. In essence, building and running AI systems is an operations process—often termed "MLOps" (Machine Learning Operations) in industry (Treveil et al. 2020)—and it can benefit from OM insights.

The challenges in managing AI systems differ from those of earlier technological waves such as RFID, the Internet of Things, or blockchain. AI performance shifts with changes in the data environment, making continuous monitoring, recalibration, and retraining essential to prevent model drift. These systems also operate with a level of autonomy and opacity unseen in prior technologies: their algorithms can uncover patterns and relationships that even their creators cannot easily interpret, underscoring the need for human oversight, audit trails, and fail-safes. A defining feature of generative AI models is their *emergent capabilities*—skills or behaviors that arise from scale or complexity rather than design intent (Berti et al. 2025). Such capabilities can yield surprising breakthroughs but also unanticipated errors. Moreover, modern AI development is intensely data-dependent, and outcomes are only as sound as the quality of the information they rely on, creating new pressure points in data collection, labeling, and feedback loops. The resulting governance challenges are thus distinct: issues of bias, safety, and compliance may evolve as the system learns, requiring oversight mechanisms that balance innovation with accountability.

Recognizing these unique challenges, we organize the discussion into two parts: first, viewing the AI development and deployment pipeline through an OM lens (focusing on productivity and scalability), and second, applying OM principles of quality, compliance, and continuous improvement to govern AI systems responsibly. Throughout, we highlight concepts like the AI flywheel—the self-reinforcing cycle of data and improvement—as well as emerging frameworks such as "Six Sigma for AI" that explicitly bridge OM and AI practice.

3.1. Managing the AI Lifecycle as an Operational Process

Training a small AI model can be handled as a technical task, yet building and operating today's large-scale AI systems is fundamentally an *operational* endeavor, replete with workflows, bottlenecks, and feedback loops. A standard AI lifecycle proceeds from data acquisition and preprocessing through feature engineering, iterative training, validation, and deployment, followed by continuous monitoring in which new data feed improvements back into the pipeline (Treveil et al. 2020). This sequence mirrors a manufacturing line (Hopp and Spearman 2011): raw materials (data) flow in, undergo transformation (training), emerge as a product (predictive capability), pass quality checks (validation), reach end users (deployment), and receive post-delivery support (monitoring and updates). Viewed through this lens, OM tools—bottleneck management, capacity planning, and continuous improvement—become indispensable at every stage of the AI workflow.

First, a data pipeline mirrors a supply chain: data must be sourced (from internal systems or external providers), cleaned and verified (ensuring quality input), and staged for use in model training. Timely, high-quality inputs matter because defective or stale data propagate errors, degrading model output (Treveil et al. 2020). OM principles amplify this point: sharing accurate, up-to-date information can raise system performance as much as inventory reductions in a two-level supply chain (Gavirneni et al. 1999; Lee et al. 2000). Lean principles further recommend "pulling" data just in time, reducing storage and transformation waste (Hopp and Spearman 2011). Yet the pipeline's dominant bottleneck remains cleansing and verification; surveys indicate that practitioners still devote a majority of effort to this step (Lohr 2014). Standardizing formats and automating validation therefore play the same role as setup-time reduction on the factory floor, accelerating flow without sacrificing quality. These pipelines are not purely technical. Manual data labeling and annotation are still essential. OM can also help optimize these social operations through workflow design and incentive alignment.

Next, model development and training constitute a production–optimization problem. Each training pass resembles a production lot: the choice of mini-batch size and retraining frequency parallels the classical lot-sizing trade-off between setup and holding costs. Graphics processing units (GPUs) and cloud instances function as servers in a queueing network, so Little's Law links job arrivals, resource capacity, and training throughput (Harchol-Balter 2013; Little 1961). The principal bottleneck is the experimentation–validation loop; recent advances mitigate this constraint by distributing computation across large clusters (Dean et al. 2012) and by employing AutoML to navigate the hyperparameter space efficiently (Feurer et al. 2015).

Tech firms have embraced this operational lens (Sendas 2025). Distributed training treats computation as a parallelizable assembly line, reducing cycle time. MLOps platforms—standardized toolchains for managing data, model training, and deployment—mirror lean production systems, minimizing rework and latency. Notably, Google's internal studies revealed the "technical debt" in AI systems: glue code, brittle dependencies, and undocumented pipelines that accrue like interest over time. In their seminal paper, Sculley et al. (2015) liken this debt to a high-interest credit card: without engineering discipline, maintenance costs spiral. Addressing this debt requires structured engineering practices, including modularization, documentation, and standardization, that are both foundational software principles and operational strategies across a wide range of industries (Swaminathan 2001; Bhatia and Swaminathan 2025).

Once an AI model is trained, its deployment into operational settings raises classic implementation challenges well studied in OM. Rather than a discrete event, deployment resembles initiating a new production process. Piloting via A/B testing or "shadow mode" allows for parallel performance evaluation and trust building, analogous to running a new assembly line alongside the incumbent system (Kleinberg et al. 2017; Sculley et al. 2015). Ramp-up and changeover management become salient, especially when transitioning decision-making authority from humans to algorithms. High-throughput environments, such as real-time fraud detection, introduce capacity planning problems: ensuring inference latency meets service level agreements calls for the application of queueing theory, simulation, and adaptive provisioning. Autoscaling in cloud services, where server instances flex in response to workload, mirrors classic inventory control problems with trade-offs between holding costs and stockouts (Dean et al. 2012; Jordan and Mitchell 2015).

Post-deployment, AI enters a continuous improvement loop that resembles an operations feedback system. The "AI flywheel" refers to a self-reinforcing dynamic where better model performance generates more usage, yielding more data, which then fuels further improvement. Gurkan and de Véricourt (2022) formalize this effect and emphasize the design of contracts and pricing schemes that accelerate data acquisition and model refinement. Operationalizing the flywheel requires structured learning processes, including bandit algorithms and adaptive experimentation, to navigate trade-offs between exploration and exploitation. In essence, learning must be embedded into the operational architecture, a principle reminiscent of Deming's continuous improvement cycle (Deming 2000). As AI becomes central to service delivery, managing the flow and quality of data becomes as critical as managing physical inventory or labor inputs.

Finally, AI itself has significant energy and infrastructure demands. As Fransoo et al. (2026) emphasize, the "supply chain for AI," spanning semiconductors, data centers, and grid infrastructure, is as fragile as it is capital-intensive. Building AI capacity requires long lead times: over a year for EUV tools, up to six months for wafers, and several years for grid connections. Data centers alone may need an additional 130–240 GW of power by 2030—equivalent to multiple national grids. Without careful management, these constraints risk triggering bullwhip cycles of scarcity and excess. Sustaining AI growth depends on strengthening these physical supply chains.

In sum, the AI lifecycle is an operations problem. Treat data as inventory, model development as production, deployment as service delivery, and learning as continuous improvement. Failures—stalled pilots, nondeployed or degrading models—typically reflect operational mismanagement: brittle pipelines, unclear ownership, and weak monitoring. Success follows process discipline: applying OM principles to reduce waste, relieve bottlenecks, and institutionalize feedback, thereby enabling faster innovation and more scalable, resilient deployments.

3.2. Quality, Compliance, and Continuous Improvement in AI Systems

As AI systems become increasingly embedded into day-to-day operations, ensuring their quality, safety, and compliance becomes a strategic challenge. This challenge echoes the quality revolution in manufacturing, which gave rise to Total Quality Management and Six Sigma. Today, we face an analogous need to operationalize quality assurance for AI. OM offers a rich heritage in quality control and continuous improvement that can be adapted to this new context. The emerging notion of a "Six Sigma for AI" calls for disciplined, data-driven processes. As with manufacturing quality, maintaining AI performance is not a one-off effort but a continuous operational endeavor.

A core OM tenet is that performance must be measured before it can be improved (Taylor 1911). In manufacturing, this tenet underpins tools such as control charts and process capability indices like C_p and C_{pk} (Montgomery 2009). For AI systems, progress similarly begins with identifying the right performance metrics—extending beyond accuracy to include false positive and false negative rates, latency, and fairness. Once these metrics are specified, statistical process control methods can be adapted to monitor model behavior over time. For instance, deviations in click-through rates on recommendation systems may signal model drift or reduced relevance. AI infrastructure at firms such as Google operationalizes these principles through dashboards and alert systems that continuously evaluate prediction stability, latency, and input distributions (Breck et al. 2017).

Root cause analysis in MLOps is often distributed across tools for anomaly detection, data validation, and monitoring. Systems like TensorFlow Data Validation (TFDV) operationalize root cause analysis by enabling early detection of schema violations, feature skew, and drift between training and serving data (Breck et al. 2019). Such systems are designed to generate high-precision alerts and contextual diagnostics to support human debugging. For example, in Google's production pipelines, TFDV helped uncover missing features in a recommender system, resulting in measurable improvements in business outcomes. These mechanisms support structured post-mortems by tracing failures to their origin in upstream data generation or configuration errors.

Continuous improvement in AI systems can be structured around the Plan-Do-Check-Act (PDCA) cycle, a foundational framework in operations and quality management (Deming 2000). In this framework, a team first identifies a performance objective (Plan), implements a change (Do), evaluates the outcome (Check), and either deploys the update or iterates further (Act). This disciplined cycle underlies emerging MLOps practices, which emphasize frequent monitoring, retraining, and deployment. Leading tech firms routinely update models in production, echoing the short feedback loops of lean manufacturing (Hopp and Spearman 2011). Yet AI introduces additional complexity: performance depends not only on code but also on data, which may shift over time. The resulting phenomenon of data drift parallels the variability of raw materials in manufacturing. As a response, organizations often adopt retraining protocols—such as monthly updates and performance-based triggers—that mirror scheduled recalibrations in industrial operations.

The logic of structured, bounded iteration is reflected in certain regulatory frameworks. The FDA's Predetermined Change Control Plans (PCCPs) operationalize this approach by requiring manufacturers to pre-specify update protocols, including verification and validation methods, acceptance criteria, and implementation procedures, to ensure device safety and effectiveness (FDA 2024). For example, a cardiology AI system might be authorized to incorporate 5,000 new patient cases every six months, with mandatory performance testing to confirm accuracy and consistency across populations before implementation. If modifications adhere to the authorized PCCP, no further regulatory review is needed. As Lai et al. (2025) demonstrate, such predictable processes can enhance compliance by aligning regulatory oversight with the operational needs of maintaining high-performing AI systems.

What might a "Six Sigma for AI" look like? In manufacturing, Six Sigma is aimed at reducing defect rates below 3.4 per million opportunities using the DMAIC framework: Define, Measure, Analyze, Improve, and Control (Linderman et al. 2006). This structured methodology enabled organizations to achieve highly reliable outcomes by iteratively diagnosing and eliminating sources of variation. A similar logic applies to high-stakes AI systems, where defects may involve materially harmful failures—such as false negatives in cancer screening or critical misclassifications in autonomous driving. Each step in DMAIC maps naturally to machine learning: Define failure modes, Measure baseline error rates, Analyze root causes, Improve through data or model refinement, and Control performance through ongoing monitoring. These practices are increasingly embedded in MLOps workflows. Red-teaming exercises, akin to quality audits, stress-test models against edge cases and adversarial inputs, paralleling destructive testing in industrial quality control (Perez et al. 2022). As in manufacturing, such disciplined routines serve to reduce unexpected failure rates and build trust in automated decision-making (Amodei et al. 2016).

Continuous improvement must also operate at the system and organizational levels. OM emphasizes documentation, standardization, and knowledge sharing to sustain process improvements. In the AI domain, this principle manifests through reusable model libraries, standardized post-mortem protocols, and shared best-practice templates. Scalable ML platforms such as Uber's Michelangelo exemplify this approach: each AI model is treated as a production process, with defined stages for data management, training, deployment, and monitoring. The system enables distributed teams to build, evaluate, and maintain models at scale, while also supporting feature sharing and reproducible pipelines across teams—closely mirroring the coordination found in factory production lines (Hermann and Del Balso 2017).

Organizations are forming AI governance boards to oversee compliance, risk, and accountability—analogous to quality councils in manufacturing. These boards vet model deployments, review testing protocols, and mandate rollback plans. OM adds procedural discipline by formalizing deployment workflows; standardized documentation and pre-launch checklists mirror aviation and healthcare practices, where structured protocols reduce errors and improve safety (Gawande 2011).

Recent deployments underscore ongoing safety and compliance challenges in AI-enabled operations. In 2024, a Canadian tribunal held an airline responsible for misinformation produced by a customer-service chatbot (Yagoda 2024), clarifying that automated agents do not displace organizational accountability and that escalation paths and audit trails must be explicit.

In summary, OM principles will play an increasing role in AI quality assurance and compliance. The rigor that OM brought to factories is now essential to realizing AI's promised operational benefits.

4. Human-AI Interaction

Even as algorithms improve, operations will remain fundamentally human for the foreseeable future. Across factory floors, managerial decision support, and customer-facing services, outcomes depend on how well humans work with information technology (Bolton and Katok 2008; Lurie and Swaminathan 2009). We now discuss how transparency and control, social norms and routines, and perceived ownership shape adoption and performance; miscalibration (over- or under-reliance) can undermine both. We then destill design principles to help improve joint human–AI performance.

4.1. Cognitive Load, Biases, and Trust

One fundamental consideration in human-AI interaction is the cognitive load imposed on users. Whereas AI systems can mitigate this burden by automating routine tasks, poorly designed interfaces may inadvertently increase mental effort, resulting in information overload and diminished performance. Cognitive Load Theory highlights the importance of reducing extraneous load—unnecessary complexity in information presentation—to enhance comprehension and task efficiency (Sweller 1988). Empirical studies support this view: in a recent clinical trial, the integration of AI-generated draft responses into physician workflows reduced self-reported task load and emotional exhaustion, despite no measurable change in reply time, suggesting a subjective alleviation of cognitive burden (Garcia et al. 2024). Design strategies such as progressive disclosure and visual hierarchy further enhance usability by aligning with human cognitive capacities (Sunil 2025).

A related phenomenon is *automation bias*, the human tendency to overtrust automated systems and accept AI-generated recommendations without sufficient scrutiny, even when those recommendations are flawed (Skitka et al. 1999). This bias has been observed across domains. In aviation, for example, overreliance on autopilot systems has contributed to accidents. The 2013 crash of Asiana Airlines Flight 214 offers a stark illustration: pilots, depending excessively on automation, failed to monitor airspeed adequately, leading to a fatal outcome (Greve 2014). Behavioral experiments reinforce this pattern, showing that individuals often follow AI advice even when it conflicts with their own assessments, particularly when the AI presents its recommendations with authority (Logg et al. 2019). Mitigating automation bias requires careful human oversight and ensuring that operators are trained to recognize and understand AI's limitations (Cummings 2004).

On the flip side of over-reliance is *algorithm aversion*, a human tendency to distrust or underutilize algorithms after seeing them err, sometimes even when the algorithm is on average more accurate than the human. Dietvorst et al. (2015) show that individuals who see an algorithm make a minor mistake often prefer their own judgment, even when the algorithm is objectively superior. This dynamic manifests in operations when users abandon decision support tools after a single failure. Part of this reaction stems from an emotional response to perceived loss of control; part reflects a misunderstanding of statistical error. Recent work indicates that user trust depends not only on stated performance metrics but also on observed outcomes during use. Yin et al. (2019) find that laypeople's trust in a model is shaped by both its advertised accuracy and its observed accuracy during live tasks. Controlled exposure, such as pilot deployments with parallel human and AI decisions, can help users contextualize algorithmic fallibility, compare against their baseline performance, and build calibrated trust over time. Consistent with this view, Snyder et al. (2025) show that operational gains hinge on both whether workers defer to algorithmic advice and how quickly they do so; higher load and higher algorithm quality increase reliance and speed.

Automation bias and algorithm aversion both stem from a deeper challenge in human–AI interaction: establishing calibrated trust. Trust must match the system's actual capabilities, neither too high nor too low (Sagona et al. 2025). Explainable AI (XAI) is one proposed solution, making algorithmic outputs interpretable for users—for instance, by highlighting key sensor inputs or trade-offs in scheduling decisions. Studies indicate that well-designed explanations can improve understanding and trust (Ribeiro et al. 2016), but explanations must be carefully tailored: overly complex ones can overwhelm, while superficial ones may appear arbitrary (Miller 2019). Yet Rudin (2019) warns that post hoc explanations of black-box models can be misleading and unreliable, especially in high-stakes contexts. For high-stakes decisions, interpretable models (i.e., those transparent by design) offer a safer and more trustworthy foundation than explanations bolted onto opaque systems. In addition to trust calibration, human–AI interaction is shaped by social norms, organizational routines, and perceived ownership of decisions; these factors influence whether humans defer to or override AI recommendations.

Cognitive workload constrains the number of AI-assisted processes a human can effectively supervise. This supervisory capacity is captured by the notion of *fan-out*—the maximum number of agents a single operator can manage (Goodrich and Olsen 2003). Formally, define F as the operator's fan-out capacity. Let T_N denote the *neglect time* (the duration an agent can function without attention before performance degrades) and T_I the *interaction time* (the duration required for a human intervention). A simple proxy is

$$F = \frac{T_N}{T_I},$$

so that longer neglect times or shorter interaction times increase the number of agents that can be supervised effectively.

This framework links naturally to queueing theory. Model the operator as a server and agents as stochastic sources of service requests. If requests arrive at rate λ and the operator resolves them at rate μ (with $\mu \approx 1/T_I$), then traffic intensity is $\rho = \lambda/\mu$. Stability requires $\rho < 1$; as $\rho \to 1$, expected waits and backlogs grow rapidly, elevating cognitive load and failure risk. In multi-operator settings ($M/M/k$), stability requires $\lambda < k\mu$, with service-level targets (e.g., $\Pr\{\text{wait} < \tau\}$) guiding staffing.

Two design levers follow. First, scope and thresholds: limit per-operator fan-out (or shed load dynamically) so that realized λ keeps ρ well below one. Adaptive intervention thresholds—such as batching alerts or escalating only when m concurrent anomalies occur—smooth arrivals and prevent alert storms. Second, cycle-time reduction: reduce T_I through better tools (pre-filled context, one-click rollbacks), training, and user interfaces that minimize extraneous load. Both levers increase effective F and preserve headroom for shocks.

In practice, teams can instrument three metrics on the supervision console: (i) alert arrival rate λ (per minute), (ii) median interaction time T_I (hence μ), and (iii) concurrency (in-queue or in-service). Control policies such as andon-style escalation, progressive disclosure, and priority classes keep ρ buffered (e.g., $\rho \leq 0.7$ under normal load), with preapproved surge modes during incidents (Dai et al. 2011). In other words, fan-out must be sized to human attentional bandwidth.

In short, cognitive and behavioral factors shape how AI is adopted in operations. Effective systems must manage cognitive load, avoid both overreliance and excessive skepticism, and maintain appropriate human oversight. Behavioral OM offers tools to design such systems, by accounting for biases and cognitive limits in decision-making.

4.2. Liability, Accountability, and User Behavior

Human use of AI reflects not only cognitive considerations but also incentives and liability. In high-stakes settings, decision makers weigh the legal consequences of accepting or rejecting algorithmic advice, not just its predictive accuracy (Price et al. 2019). These liability pressures mean that the social value of AI depends on institutional rules that assign responsibility. Aligning accountability with human values and objectives is therefore critical for its effective deployment.

In healthcare, liability regimes shape clinicians' use of assistive AI: to limit exposure, physicians may overuse AI in low-uncertainty cases and underuse it when uncertainty is high (Dai and Singh 2025). Fairness-oriented penalties can similarly discourage use for disadvantaged groups (Luan et al. 2024). Accountability design thus matters: punitive rules deter adoption, whereas safe harbors and integration into clinical guidelines promote it. Comparable attribution patterns appear in manufacturing and transportation, where automation failures often shift blame to human operators—especially when systems are anthropomorphized—weakening trust and collaboration (Dawtry and Callan 2024). Legal evidence likewise underscores the need to assign responsibility clearly when AI informs decisions (Cohen et al. 2023).

This interplay of liability and behavior extends to organizational decision-making. Firms may avoid deploying accurate and potentially fair AI tools—such as in credit underwriting or autonomous logistics—due to uncertainty over legal responsibility or regulatory exposure. Conversely, some firms invoke AI to offload blame for unpopular actions, attributing decisions to algorithmic processes (Binns 2018). These behaviors reflect how unclear accountability can both discourage valuable technology adoption and invite opportunistic blame-shifting. Such dynamics drive organizational demand for explainability and logging: when AI systems record decision rationales, operators are more likely to trust and engage with them, knowing that responsibility can be traced. The presence of a logging mechanism—analogous to a flight recorder in aviation—can both reassure users and promote more deliberate interaction.

From an OM (especially behavioral OM) perspective, the integration of machine learning and behavioral science provides actionable insights for aligning individual behavior with system-wide optimal AI use through careful incentive and interface design (Davis et al. 2024; Lu and Tomlin 2025). Physicians may underutilize AI in complex cases unless malpractice standards and training reduce liability concerns, while in aviation, the introduction of cockpit automation has created challenges at the human-machine interface, where highly reliable automated systems can alter pilot roles and decision-making processes (Wiener 1989). Similarly, loan officers may discount algorithmic risk assessments unless performance management emphasizes outcome consistency over isolated errors. To address these challenges of suboptimal AI utilization, organizations are increasingly investing in comprehensive training programs focused on human-AI partnership, with research indicating that resistance to AI adoption often stems from unfamiliarity with technologies and skill gaps (Deloitte 2024). Leading technology companies have formed consortia specifically focused on upskilling workers for AI collaboration, with members committing to train over 95 million individuals in AI-related skills over the next decade, recognizing the urgency of building workforce capabilities that complement rather than compete with automated systems (IBM 2024).

Organizational implementation strategies must account for the temporal dynamics of AI adoption within workforces. Successful deployment often follows a staged approach where early adopters serve as internal champions, demonstrating system value and building institutional knowledge before broader rollout (Rogers 2003; Singer et al. 2021). This social learning mechanism proves particularly important in operations contexts where cross-functional collaboration and employee involvement in system design significantly enhance technology acceptance compared to imposed solutions (Venkatesh et al. 2003).

The long-term sustainability of human–AI collaboration depends on designing systems that preserve rather than erode human expertise. Organizations that maintain meaningful human involvement while enhancing decision quality through AI augmentation achieve superior performance compared to full automation approaches (Chowdhury et al. 2022). Effective implementations create feedback mechanisms where human judgment informs algorithmic improvement, establishing complementary rather than substitutive relationships between human and machine capabilities (Parasuraman et al. 2000).

4.3. Designing Effective Human–AI Workflows

The goal of designing human–AI workflows is to harness the comparative strengths of humans and machines. This challenge spans OM, industrial engineering, and organizational design. Effective workflows must assign tasks, structure information flows, and define roles in ways that optimize system performance.

A core principle is comparative advantage. Machines excel at processing speed, scale, and consistency. Humans, in contrast, are better at contextual reasoning, ethical judgment, and adaptation (Autor 2015). Automation displaces labor on routine tasks but creates new opportunities in areas where humans retain an edge (Acemoglu and Restrepo 2018). The best workflows allocate high-volume, repetitive tasks to AI while reserving ambiguous or high-stakes decisions for humans. In customer service, AI handles initial inquiries, escalating complex cases to human representatives. In manufacturing, AI vision systems perform quality checks, with engineers reviewing exceptions and calibrating system thresholds through adaptive sampling strategies.

Interfaces are central to human-AI collaboration. Parasuraman et al. (2000) distinguish four automation functions—information acquisition, analysis, decision selection, and action implementation—each on a manual-to-automatic continuum. Effective interfaces avoid opacity: they expose confidence intervals, key drivers, and alternatives, and enable drill-down and selective overrides, improving trust and decisions. Transparency and control shape acceptance; complementing interpretability, a large-scale field experiment shows that disclosing chatbot identity and calibrating human-like cues shifts engagement and compliance, with implications for arrival, handling, and escalation rates (Xu et al. 2024).

Robust workflows also require exception handling. AI systems encounter edge cases, degraded inputs, and distributional shifts (Treveil et al. 2020). Well-designed processes specify when control shifts to humans, using fallback mechanisms inspired by poka-yoke principles. Queueing models help ensure sufficient standby capacity during intervention spikes. Autonomy should depend on task stakes and system reliability. As Agrawal et al. (2018) argue, AI reduces the cost of prediction, but *judgment*—deciding objectives and interpreting outputs—remains a human comparative advantage. AI can automatically handle routine decisions, such as reorder quantities or staffing allocations, within set bounds. Significant deviations from historical patterns trigger human review, akin to control limits in quality management. Over time, as performance data accumulates, these thresholds can be refined through simulation and experimentation.

Organizational structures must adapt (Agrawal et al. 2022). New roles, such as automation strategists and AI operations managers, become critical for monitoring system performance and resolving failures. Workflow designs often encounter unforeseen bottlenecks, requiring iterative refinement. Agile methods, time-and-motion studies, and variance analysis help diagnose problems and guide adjustments.

Social and ethical concerns also matter. AI systems risk undermining morale if perceived as opaque or biased. Human oversight, transparency, and appeal mechanisms can mitigate these risks. In customer-facing roles, keeping AI in the background can preserve the perception of human-centered service. Ultimately, successful AI integration requires workflows aligned not just with task efficiency but with user expectations and organizational values.

Finally, the future of work will be designed around Human-AI collaborative processes. It is important to consider the different aspects of such interactions based on whether those are internal (employee using AI) or external (customer using AI) since the critical focus areas will vary based on the type of system being considered (Swaminathan and Xu 2026).

Human–AI workflows, when thoughtfully designed, create synergies rather than simple automation. Brynjolfsson and Mitchell (2017) note that while machine learning is powerful for tasks with clear inputs, abundant data, and tolerance for errors, most occupations include tasks that resist automation. Achieving the performance premium from AI requires careful task design, robust exception handling, and continuous refinement. OM has a key role to play in this process, turning AI from a mere tool into a strategic asset.

5. Future Research Directions

Integrating AI and operations opens a broad research frontier. AI is a general-purpose technology that extends OM beyond traditional boundaries. A central theme is *complementarity*: prediction and automation should augment—not replace—human and organizational capabilities. Treating AI as a catalyst invites study of how OM problems are redefined as machine intelligence embeds in supply chains and services. The agenda demands bold ideas grounded in careful analysis; the avenues outlined below arise directly from the three pillars of our framework and highlight appropriate methodologies.

5.1. AI for OM

Claims that AI raises productivity and decision quality still outpace evidence; beyond a few high-profile generative-AI deployments, credible estimates remain sparse. OM scholars are well positioned to design field experiments and exploit natural experiments to identify causal effects, thereby expanding the empirical foundations of AI's operational impacts. Core tasks include mapping where and how AI scales value—when does adding AI to design lift productivity, how much does predictive maintenance extend asset life, and how does algorithmic advice reshape (and sometimes distort) human judgment? As richer operational data arrive, emphasis shifts to causal inference from observational settings, requiring quasi-experimental designs and identification strategies tailored to AI deployments (Einav and Levin 2014). At the theoretical level, AI-augmented operations move from static systems to adaptive processes that learn. This calls for models of manufacturing and services in which prediction and optimization are delegated to AI agents working with humans. Research questions include: What equilibrium and welfare properties emerge as supply chains self-reconfigure? How does convergence behave when choices evolve via reinforcement rather than repetition? Combining AI's pattern recognition with OM's alignment focus points to models with endogenous learning, stability and performance bounds, and human-in-the-loop constraints. Together, these empirical and theoretical directions address the need for rigorous methodologies—randomized trials, quasi-experiments, and analytic models—to validate operational gains.

5.2. OM for AI

AI lifecycles resemble supply chains: data are acquired, labeled, and transformed through training and deployment stages with feedback and quality checks. Research can model these pipelines as process networks, using queueing, inventory, and stochastic-control frameworks to optimize labeling, retraining, and compute allocation. Empirical analyses can quantify drift, retraining outcomes, and exception handling, while analytical work—drawing on principal–agent and mechanism-design models—can formalize incentives.

A related direction concerns governance and organizational design. Scalable, trustworthy AI depends on alignment with organizational goals. Building on work linking structure, innovation, and performance (Loch et al. 2025), we can view alignment as an adaptive process where safety, accountability, and learning are embedded in design. An important question relates to model drift and shifting norms. Another one relates to selective deployment of AI in business processes. Addressing such questions offers methodological avenues that span process simulation, modeling, and empirical analyses of AI pipelines.

5.3. Human–AI Interaction

Interaction is, by definition, reciprocal. People must trust AI outputs, but AI systems also rely on human prompts and feedback (Sagona et al. 2025). Understanding how this mutual adaptation shapes learning, tacit knowledge, and decision speed is an important research frontier. In clinical settings, for example, real-time decision support can improve diagnostic accuracy yet gradually erode physicians' experiential expertise—a dynamic observed as well in aviation and finance (Rajkomar et al. 2019). Moving forward requires shifting from static models of trust to dynamic frameworks that account for calibration, algorithm aversion, and over-reliance. Empirical work—ranging from behavioral and field experiments to qualitative case studies—can test how transparency, explanation, and feedback loops affect trust, performance and intervention behavior. Mixed-methods approaches, including ethnography and interviews, can further illuminate how social norms and perceived ownership influence overrides, adaptation, and sustained use. OM scholars should also engage with regulations such as the EU Artificial Intelligence Act (European Commission 2024), which call for proactive audits and human-centered governance. Bringing together behavioral insight and policy thinking through strong theory and credible evidence will be essential for making human–AI systems work—and for earning the trust they require.

6. Conclusion

AI and OM are reshaping each other. This perspective proposes a foundational triadic framework—AI for OM, OM for AI, and Human–AI Interaction—to structure their co-evolution. AI tools empower operations to enhance product design, optimize supply chains, automate production, and deliver services efficiently and flexibly. Yet successful deployment of AI systems critically depends on OM principles—process integration, continuous improvement, and rigorous governance methods such as Six Sigma-inspired approaches. Equally significant is human interaction: how people trust, adapt, and collaborate with AI is central to realizing potential benefits.

Complementarity emerges as a key insight. AI amplifies OM's alignment capabilities; OM provides systematic frameworks and human-centric considerations essential for AI development and deployment. Human–AI collaboration, rather than full automation, optimizes operational performance (Brynjolfsson and Mitchell 2017). An integrated approach—combining AI, workflow optimization, and behavioral insights—will define future operational excellence. This integration yields substantial benefits: enhanced efficiency, consistency, scalability, and innovation. Yet, risks are significant. Bias in AI data can perpetuate unfairness; integration difficulties cause many AI initiatives to stall. Human biases, notably algorithm aversion and automation overreliance, can undermine potential gains. AI governance lapses pose severe financial and societal risks, and AI-induced workforce displacement requires thoughtful policy responses.

In closing, AI's spread is reshaping OM, much as steam power, the moving assembly line, containerization, and the Internet did in earlier eras. Canonical OM objectives—efficiency, responsiveness, and accountability—remain vital, yet they now intertwine with novel tasks: orchestrating hybrid human–machine workflows, embedding operational metrics into algorithms, and ensuring quality when decisions are executed autonomously. Properly framed, OM can do more than absorb AI; it can help guide the technology's deployment, linking data pipelines, model governance, and human complementarity toward operational performance—when well designed and monitored. The alignment logic at the heart of OM offers that map: AI can expand the feasible set of operational choices, OM helps to discipline model development through process control and continuous improvement, and human–AI interaction shapes whether those choices are trusted and sustained. Realizing AI's potential therefore hinges not only on adoption but also on deliberate design, organizational alignment, and governance; OM helps identify when, where, and how AI improves performance—and when it does not.

References

Abràmoff MD, Whitestone N, Patnaik JL, et al. (2023) Autonomous artificial intelligence increases real-world specialist clinic productivity in a cluster-randomized trial. *npj Digital Medicine* 6(1):184.

Acemoglu D, Restrepo P (2018) The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review* 108(6):1488–1542.

Agrawal A, Gans J, Goldfarb A (2018) *Prediction Machines: The Simple Economics of Artificial Intelligence* (Boston, MA: Harvard Business Review Press).

Agrawal A, Gans J, Goldfarb A (2022) *Power and Prediction: The Disruptive Economics of Artificial Intelligence* (Boston, MA: Harvard Business Review Press).

Agrawal N, Cohen MA, Deshpande R, Deshpande V (2024) How machine learning will transform supply chain management. *Harvard Business Review* 102(2):128–137.

Amodei D, Olah C, Steinhardt J, Christiano P, Schulman J, Mané D (2016) Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565* URL http://dx.doi.org/10.48550/arXiv.1606.06565.

Autor DH (2015) Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives* 29(3):3–30.

Berti L, Giorgi F, Kasneci G (2025) Emergent abilities in large language models: A survey. *arXiv preprint arXiv:2503.05788* URL https://doi.org/10.48550/arXiv.2503.05788.

Bertsimas D, Kallus N (2020) From predictive to prescriptive analytics. *Management Science* 66(3):1025–1044.

Bhatia A, Swaminathan JM (2025) Measuring consistency in service delivery: Examining the effect of process standardization on hospital performance. *Production and Operations Management* URL http://dx.doi.org/10.1177/10591478251361985, ePub ahead of print.

Binns R (2018) Algorithmic accountability and public reason. *Philosophy & Technology* 31(4):543–556.

Bolton GE, Katok E (2008) Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing & Service Operations Management* 10(3):519–538.

Bommasani R, Hudson DA, Adeli E, et al. (2021) On the opportunities and risks of foundation models. Tech. Report Stanford CRFM Report, Center for Research on Foundation Models, Stanford.

Breck E, Cai S, Nielsen E, Salib M, Sculley D (2017) The ML test score: A rubric for ML production readiness and technical debt reduction. *2017 IEEE International Conference on Big Data*, 1123–1132 (IEEE).

Breck E, Polyzotis N, Roy S, Whang SE, Zinkevich M (2019) Data validation for machine learning. *Proceedings of the 2nd SysML Conference*.

Brynjolfsson E, Li D, Raymond L (2025) Generative AI at work. *Quarterly Journal of Economics* 140(2):889–942.

Brynjolfsson E, Mitchell T (2017) What can machine learning do? workforce implications. *Science* 358(6370):1530–1534.

Brynjolfsson E, Rock D, Syverson C (2017) Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. Working Paper 24001, National Bureau of Economic Research.

C-Suite Strategy (2024) AI in procurement: Revolutionizing strategic sourcing for today's c-suite. Online article, URL https://bit.ly/AiInProcureCSuite, accessed 31 May 2025.

Cachon GP (2026) The current and future impact of AI on supply chains. Cohen MA, Dai T, eds., *AI in Supply Chains: Perspectives from Global Thought Leaders*, 3–12 (Cham, Switzerland: Springer).

Cachon GP, Netessine S (2006) *Game Theory in Supply Chain Analysis*, 200–233 (INFORMS), ISBN 9781877640209, URL http://dx.doi.org/10.1287/educ.1063.0023.

Chatterjee C, Dai T (2025) How can AI speed life-saving cures to patients? *BioProcess International* 23(10):e1.

Chomsky N, Roberts I, Watumull J (2023) The false promise of ChatGPT URL https://bit.ly/

noam-chomsky-chatgpt-ai, accessed 31 May 2025.

Chowdhury S, Dey PK, Joel-Edgar S, et al. (2022) AI-employee collaboration and business performance: Integrating knowledge-based view, socio-technical systems and organisational socialisation framework. *Journal of Business Research* 144:31–49.

Cohen MA, Dahan S, Khern-am nuai W, et al. (2023) The use of AI in legal systems: Determining independent contractor vs. employee status. *Artificial Intelligence and Law* 1–30.

Cohen MA, Tang CS (2024) The role of AI in developing resilient supply chains URL https://gjia.georgetown.edu/2024/02/05/the-role-of-ai-in-developing-resilient-supply-chains/, accessed January 6, 2026.

Cui R, Li M, Zhang S (2022) AI and procurement. *Manufacturing & Service Operations Management* 24(2):691–706.

Cummings ML (2004) Automation bias in intelligent time critical decision support systems. *AIAA 1st Intelligent Systems Technical Conference*, volume 2, 557–562 (American Institute of Aeronautics and Astronautics).

Dai T, Singh S (2025) Artificial intelligence on call: The physician's decision of whether to use AI in clinical practice. *Journal of Marketing Research* 62(5):854–875.

Dai T, Sycara K, Lewis M (2011) A game theoretic queueing approach to self-assessment in human-robot interaction systems. *2011 IEEE International Conference on Robotics and Automation*, 3775–3780 (IEEE).

Dai T, Taylor T (2025) Designing enterprise AI systems: Hallucination, creativity, and moral hazard.

Dai T, Tayur SR (2022) Designing AI-augmented healthcare delivery systems for physician buy-in and patient acceptance. *Production and Operations Management* 31(12):4443–4451.

Danon B (2018) How GM and Autodesk are using generative design for vehicles of the future. ADSK News, URL https://bit.ly/gam-autodesk, accessed: May 31, 2025.

Davis AM, Mankad S, Corbett CJ, Katok E (2024) The best of both worlds: Machine learning and behavioral science in operations management. *Manufacturing & Service Operations Management* 26(2):345–359.

Dawtry RJ, Callan MJ (2024) Hazardous machinery: The assignment of agency and blame to robots versus non-autonomous machines. *Journal of Experimental Social Psychology* 111:104582.

Dean J, Corrado GS, Monga R, et al. (2012) Large scale distributed deep networks. *Advances in Neural Information Processing Systems*, volume 25, 1223–1231.

Deloitte (2024) State of Generative AI in the Enterprise 2024. Technical report, URL https://bit.ly/deloitte-genai-enterprise, accessed: May 31, 2025.

Deming WE (2000) *Out of the Crisis* (Cambridge, MA: MIT Press).

Dietvorst BJ, Simmons JP, Massey C (2015) Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General* 144(1):114–126.

Einav L, Levin J (2014) Economics in the age of big data. *Science* 346(6210):715–721.

European Commission (2024) Proposal for a regulation laying down harmonised rules on artificial intelligence (artificial intelligence act). URL https://artificialintelligenceact.eu/, accessed: May 31, 2025.

Fazlollahi AM, Yilmaz R, Winkler-Schwartz A, et al. (2023) AI in surgical curriculum design and unintended outcomes for technical competencies in simulation training. *JAMA Network Open* 6(9):e2334658.

FDA (2018) FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. URL https://bit.ly/fda-idx-dr, accessed: May 31, 2025.

FDA (2024) Predetermined change control plans for medical devices: Draft guidance for industry and FDA staff. Guidance document, URL https://bit.ly/fda-pccp, accessed 31 May 2025.

Ferrucci D, Brown E, Chu-Carroll J, et al. (2010) Building Watson: An overview of the DeepQA project. *AI Magazine* 31(3):59–79.

Feurer M, Klein A, Eggensperger K, et al. (2015) Efficient and robust automated machine learning. *Advances in Neural Information Processing Systems*, volume 28, 2962–2970.

Fickenscher L (2021) UPS upgrades ORION tool with real-time data to improve delivery routes. *Supply Chain Dive* URL https://bit.ly/ups-orion-route, accessed: May 29, 2025.

Flammer C, Loch CH (2024) Editorial statement—Sustainability. *Management Science* 70(4):vii–vii.

Fransoo JC, Peels R, Udenio M (2026) Navigating supply chain dynamics for sustained AI growth. Cohen MA, Dai T, eds., *AI in Supply Chains: Perspectives from Global Thought Leaders*, 37–54 (Cham, Switzerland: Springer).

Garcia P, Ma SP, Shah S, et al. (2024) Artificial intelligence–generated draft replies to patient inbox messages. *JAMA Network Open* 7(3):e243201.

Gavirneni S, Kapuscinski R, Tayur SR (1999) Value of information sharing in a two-level supply chain. *Management Science* 45(1):16–24.

Gawande A (2011) *The Checklist Manifesto: How to Get Things Right* (New York, NY: Metropolitan Books).

Gijsbrechts J, Boute RN, Van Mieghem JA, et al. (2026) AI in inventory management: The disruptive era of DRL and beyond. Cohen MA, Dai T, eds., *AI in Supply Chains: Perspectives from Global Thought Leaders*, 137–148 (Cham, Switzerland: Springer).

Goodrich MA, Olsen DR (2003) Seven principles of efficient human robot interaction. *Proceedings of the 2003 IEEE International Conference on Systems, Man and Cybernetics*, 3942–3948 (Washington, DC: IEEE).

Greve JE (2014) NTSB: Pilot of fatal asiana crash lacked 'critical manual flying skills' URL https://time.com/2917445/asiana-214-crash-pilot-ntsb/, accessed 31 May 2025.

Gurkan H, de Véricourt F (2022) Contracting, pricing, and data collection under the AI flywheel effect. *Management Science* 68(12):8791–8808.

Harchol-Balter M (2013) *Performance Modeling and Design of Computer Systems: Queueing Theory in Action* (Cambridge, United Kingdom: Cambridge University Press), URL http://dx.doi.org/10.1017/CBO9781139226424.

Hermann J, Del Balso M (2017) Meet Michelangelo: Uber's machine learning platform. Online article, URL https://ubr.to/4e0tHLV, accessed 30 May 2025.

Hopp WJ, Spearman ML (2011) *Factory Physics* (Long Grove, IL: Waveland Press).

IBM (2024) Leading companies launch consortium to address AI's impact on the technology workforce. URL https://bit.ly/ibm-consort, accessed: May 31, 2025.

International Federation of Robotics (2023) World robotics 2023: Industrial robots. URL https://bit.ly/

ifr23pdf, accessed: May 31, 2025.

Jin H, Shepardson D (2023) California sidelines GM cruise's driverless cars, cites safety risk. Reuters, URL https://bit.ly/GMCruiseCA.

Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. *Science* 349(6245):255–260.

Katz Y (2012) Noam chomsky on where artificial intelligence went wrong. *The Atlantic* URL https://bit.ly/chomsky-atlantic12, accessed: May 31, 2025.

Kavadias S, Sommer SC (2009) The effects of problem structure and team diversity on brainstorming effectiveness. *Management Science* 55(12):1899–1913.

Kleinberg J, Lakkaraju H, Leskovec J, Ludwig J, Mullainathan S (2017) Human decisions and machine predictions. *The Quarterly Journal of Economics* 133(1):237–293.

Kleindorfer PR, Singhal K, Van Wassenhove LN (2005) Sustainable operations management. *Production and Operations Management* 14(4):482–492.

Lai J, Xu LL, Fang X, Dai T (2025) Regulating adaptive new products: Can less oversight lead to better development practices?

Lanteri A, Cappuccio ML, Galliott JC, Eyssel F (2023) Successful AI adoption requires building employees' tolerance. https://bit.ly/lse-ai, accessed: May 31, 2025.

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444.

Lee HL (2004) The triple-A supply chain. *Harvard Business Review* 82(10):102–112.

Lee HL, So KC, Tang CS (2000) The value of information sharing in a two-level supply chain. *Management Science* 46(5):626–643.

Lee HL, Tang CS (2018) Socially and environmentally responsible value chain innovations: New operations management research opportunities. *Management Science* 64(3):983–996.

Lepore J (2009) Not so fast. *The New Yorker* 85(32):114–122, october 12 issue.

Linderman K, Schroeder RG, Choo AS (2006) Six sigma: The role of goals in improvement teams. *Journal of Operations Management* 24(6):779–790.

Little JDC (1961) A proof for the queuing formula: $L = \lambda W$. *Operations Research* 9(3):383–387.

Loch CH, Ladas K, Kavadias S (2025) Organizational culture, innovation, and competitive performance: A multilevel dynamic model. *Management Science* 71(11):9193–9212.

Loch CH, Terwiesch C (1999) Measuring the effectiveness of overlapping development activities in complex product development. *Management Science* 45(6):733–750.

Logg JM, Minson JA, Moore DA (2019) Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes* 151:90–103.

Lohr S (2014) For big-data scientists, 'janitor work' is key hurdle to insights. *The New York Times* URL https://bit.ly/data-janitor, accessed: May 31, 2025.

Lu T, Tomlin B (2025) Augmenting the operations manager with a prediction machine, available at SSRN: https://ssrn.com/abstract=5282711.

Luan S, Singh S, Dai T (2024) Algorithmic bias and physician liability. Working Paper 25-16, Johns Hopkins

University, available at SSRN: https://dx.doi.org/10.2139/ssrn.5046254.

Lurie NH, Swaminathan JM (2009) Is timely information always better? The effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes* 108(2):315–329.

Marion TJ, Srour M, Piller F (2024) When generative AI meets product development URL https://bit.ly/gai-pd, accessed 31 May 2025.

McCarthy J, Minsky ML, Rochester N, Shannon CE (2006) A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine* 27(4):12.

McKinsey (2024) Harnessing AI and analytics for advanced procurement strategies URL https://mck.co/3SO28vM, accessed 31 May 2025.

McKinsey (2024) LLM to ROI: How to scale generative AI in retail. URL https://bit.ly/llm2roi.

Metz C, Henry J, Laffin B, Lieberman R, Lu Y (2024) How self-driving cars get help from humans hundreds of miles away. *The New York Times* URL https://nyti.ms/3SO2aUq, accessed: May 31, 2025.

Miller T (2019) Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence* 267:1–38.

Mišić VV, Perakis G (2020) Data analytics in operations management: A review. *Manufacturing & Service Operations Management* 22(1):158–169.

Montgomery DC (2009) *Introduction to Statistical Quality Control* (Hoboken, NJ: John Wiley & Sons), 6 edition.

Murray CC, Chu AG (2015) The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery. *Transportation Research Part C: Emerging Technologies* 54:86–109.

Nong P, Adler-Milstein J, Apathy NC, Holmgren AJ, Everson J (2025) Current use and evaluation of artificial intelligence and predictive models in US hospitals. *Health Affairs* 44(1):90–98.

Nucleus Research (2023) Quantifying the value of predictive maintenance. Technical report, Nucleus Research, x98.

Oitzman M (2025) Starship Technologies surpasses 8M autonomous deliveries. URL https://bit.ly/starship8m, accessed: May 31, 2025.

Olsen TL, Tomlin B (2020) Industry 4.0: Opportunities and challenges for operations management. *Manufacturing & Service Operations Management* 22(1):113–122.

Parasuraman R, Sheridan TB, Wickens CD (2000) A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 30(3):286–297.

Perez E, Huang S, Song F, Cai T, et al. (2022) Red teaming language models with language models. *arXiv* URL http://dx.doi.org/10.48550/arXiv.2202.03286.

Price WNI, Gerke S, Cohen IG (2019) Potential liability for physicians using artificial intelligence. *JAMA* 322(18):1765–1766.

Rajkomar A, Dean J, Kohane I (2019) Machine learning in medicine. *New England Journal of Medicine* 380(14):1347–1358.

Ribeiro MT, Singh S, Guestrin C (2016) Why should I trust you?: Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data*

Mining, 1135–1144 (ACM).

Rogers EM (2003) *Diffusion of Innovations* (New York, NY: Free Press), 5 edition.

Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence* 1(5):206–215.

Sagona M, Dai T, Macis M, et al. (2025) Trust in AI-assisted health systems and AI's trust in humans. *npj Health Systems* 2:10.

SAP (2025) AI in procurement: A comprehensive guide. URL https://www.sap.com/resources/ai-in-procurement, accessed: May 31, 2025.

Saqlain M, Abbas Q, Lee JY (2020) A deep convolutional neural network for wafer defect identification on an imbalanced dataset in semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing* 33(3):436–444.

Sculley D, Holt G, Golovin D, et al. (2015) Hidden technical debt in machine learning systems. Cortes C, Lawrence ND, Lee DD, Sugiyama M, Garnett R, eds., *Advances in Neural Information Processing Systems 28 (NeurIPS 2015)* (Montréal, Canada: Curran Associates, Inc.).

Sendas N (2025) Operational excellence in MLOps. URL https://bit.ly/oe-mlops, accessed: May 31, 2025.

Shortliffe EH (1976) *Computer-Based Medical Consultations: MYCIN* (Amsterdam: Elsevier).

Siemens (2024) Generative artificial intelligence takes Siemens' predictive maintenance solution to the next level. URL https://bit.ly/gai-siemens, accessed: May 31, 2025.

Silver D, Huang A, Maddison CJt (2016) Mastering the game of Go with deep neural networks and tree search. *Nature* 529(7587):484–489.

Simchi-Levi D (2014) OM research: From problem-driven to data-driven research. *Manufacturing & Service Operations Management* 16(1):2–10.

Simchi-Levi D, Dai T, Menache I, Wu MX (2025) Democratizing optimization with generative AI. Working paper, Johns Hopkins University, URL http://dx.doi.org/10.2139/ssrn.5511218.

Simchi-Levi D, Mellou K, Menache I, et al. (2026) Large language models for supply chain decisions. Cohen MA, Dai T, eds., *AI in Supply Chains: Perspectives from Global Thought Leaders*, 93–104 (Cham, Switzerland: Springer).

Simchi-Levi D, Schmidt W, Wei Y (2015) From superstorms to factory fires: Managing unpredictable supply-chain disruptions. *Harvard Business Review* 92(1–2):96–101.

Simon HA (1987) Two heads are better than one: The collaboration between AI and OR. *Interfaces* 17(4):8–15.

Singer SJ, Kellogg KC, Galper AB, Viola D (2021) Enhancing the value to users of machine learning-based clinical decision support tools: A framework for iterative, collaborative development and implementation. *Health Care Management Review* 47(2):E21–E31.

Singhal K, Singhal J (2022) Technology and manufacturing-and-service operations since the industrial revolution. *Production and Operations Management* 31(12):4276–4282.

Skitka LJ, Mosier KL, Burdick M (1999) Does automation bias decision-making? *International Journal of Human-Computer Studies* 51(5):991–1006.

Snyder C, Keppler S, Leider S (2025) Algorithm reliance, fast and slow. *Management Science,* ePub ahead of print.

Stanley KD, Mignano J, Blumenthal MS (2024) Public welfare and emerging technology: Balancing interests in automated vehicle regulations across six jurisdictions. *Road Vehicle Automation 11*, 32–48 (Springer).

Stewart J (2017) A drone-slinging UPS van delivers the future. *WIRED* URL https://bit.ly/drone-ups, accessed: May 29, 2025.

Sunil G (2025) Cognitive load theory in UI design. *Aufait UX* URL https://bit.ly/UX-cog, accessed: May 31, 2025.

Swaminathan JM (2001) Enabling customization using standardized operations. *California Management Review* 43(3):125–135.

Swaminathan JM (2025) OM Forum—Operations management research: Relevance and impact. *Manufacturing & Services Operations Management* 27(1):1–7.

Swaminathan JM, Smith SF, Sadeh NM (1998) Modeling supply chain dynamics: A multiagent approach. *Decision Sciences* 29(3):607–632.

Swaminathan JM, Tayur SR (1998) Managing broader product lines through delayed differentiation using vanilla boxes. *Management Science* 44(12.2):S161–S172.

Swaminathan JM, Tayur SR (2003) Models for supply chains in e-business. *Management Science* 49(10):1387–1406.

Swaminathan JM, Xu Y (2026) Human–AI enabled supply chains. Cohen MA, Dai T, eds., *AI in Supply Chains: Perspectives from Global Thought Leaders*, 191–205 (Cham, Switzerland: Springer).

Sweller J (1988) Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12(2):257–285.

Taylor FW (1911) *The Principles of Scientific Management* (New York, NY: Harper & Brothers).

Tayur SR (2026) Can AI really transform real-world supply chains anytime soon? Cohen MA, Dai T, eds., *AI in Supply Chains: Perspectives from Global Thought Leaders*, 281–294 (Cham, Switzerland: Springer).

Treveil M, Omont N, Stenac C, et al. (2020) *Introducing MLOps* (Sebastopol, CA: O'Reilly Media, Inc.).

Tulsyan A, Garvin C, Ündey C (2018) Advances in industrial biopharmaceutical batch process monitoring: Machine-learning methods for small data problems. *Biotechnology and Bioengineering* 115(8):1915–1924.

Van Hoek R, DeWitt M, Lacity M, Johnson T (2022) How Walmart automated supplier negotiations. *Harvard Business Review* URL https://bit.ly/hbr-wmt, accessed: May 31, 2025.

VandeHei J, Allen M (2025) Behind the curtain: A white-collar bloodbath. *Axios* URL https://bit.ly/ai-job-axios, accessed: May 31, 2025.

Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology: Toward a unified view. *MIS Quarterly* 27(3):425–478.

Wang H, Zong Z, Xia T, et al. (2021) Rewriting by generating: Learn heuristics for large-scale vehicle routing problems, manuscript submitted to ICLR 2021; available at https://openreview.net/forum?id=xxWl2oEvP2h.

Wang Y, Liu J, Liu Y, et al. (2017) Ladder: A human-level bidding agent for large-scale real-time online auctions. *arXiv* URL https://arxiv.org/abs/1708.05565.

Wiberg H, Dai T, Lam H, et al. (2025) Synergizing artificial intelligence and operations research: Perspectives from INFORMS fellows on the next frontier. *INFORMS Journal on Data Science* URL http://dx.doi.org/10.1287/ijds.2025.0077, ePub ahead of print.

Wiener EL (1989) Cockpit automation. *Human Factors in Modern Air Transport Operations*, 433–449 (Amsterdam, The Netherlands: Elsevier).

Womack JP, Jones DT, Roos D (1990) *The Machine That Changed the World* (New York, NY: Free Press).

Workhorse (2023) Workhorse group announces first commercial drone purchase orders. URL https://bit.ly/workhorse23, accessed: May 31, 2025.

Wuest T, Kusiak A, Dai T, Tayur SR (2020) Impact of COVID-19 on manufacturing and supply networks—The case for AI-inspired digital transformation. Working paper, Johns Hopkins University.

Xu Y, Dai H, Yan W (2024) Identity disclosure and anthropomorphism in voice chatbot design: A field experiment. *Management Science,* ePub ahead of print.

Yagoda M (2024) Airline held liable for its chatbot giving passenger bad advice: What this means for travellers. Online article, URL https://bbc.in/49xDqsY.

Yan C, Zhu H, Korolko N, Woodard D (2020) Dynamic pricing and matching in ride-hailing platforms. *Naval Research Logistics* 67(8):705–724.

Yin M, Vaughan JW, Wallach H (2019) Understanding the effect of accuracy on trust in machine learning models. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12 (ACM).