

Human Robot Collaboration and AI-Enabled Warehouse Safety Management for Sustainable Logistics Operations in Saudi Arabia

Abstract

Robotic warehousing offers opportunities to reduce hazardous manual transport while introducing new risks at shared workstations, aisle intersections and digitally coordinated handovers. This critical integrative review evaluates how human–robot collaboration and artificial intelligence can support safer, more sustainable logistics operations, with Saudi Arabia as the proposed implementation context. Thirty DOI-indexed publications dated 2020–2025 were purposively examined across operational research, ergonomics, robotics, occupational safety and logistics management. Evidence was organised around exposure pathways, collaborative task allocation, predictive sensing, safety assurance, workforce acceptance and organisational governance. The synthesis shows that robotics may reduce walking and handling demands, but operational benefits cannot be equated with verified injury reduction. Adaptive speed control, transparent intention signalling, fatigue-aware scheduling and human-centred training appear promising when embedded within independently validated safeguards. Conversely, sensor uncertainty, deadline pressure, opaque algorithmic decisions and fragmented system ownership may transfer rather than remove risk. A layered safety architecture and staged evaluation framework are proposed for Saudi logistics facilities. The review concludes that sustainable robotic transformation requires the simultaneous measurement of injury exposure, employee wellbeing, operational quality, energy intensity and system recoverability, rather than optimisation of throughput in isolation.

Keywords: human–robot collaboration; autonomous mobile robots; warehouse safety; artificial intelligence; ergonomics; sustainable logistics; Saudi Arabia

1. Introduction

Warehouse automation is no longer confined to fenced industrial equipment. Autonomous mobile robots, mobile racks, collaborative manipulators and computer-vision systems increasingly operate near employees engaged in picking, replenishment, consolidation and dispatch. Their proximity creates a practical contradiction: collaboration can reduce burdensome travel and lifting, yet closer

human-machine contact introduces collision exposure, unpredictable motion and additional monitoring demands. Warehouse workers remain essential because many products, exceptions and customer requirements exceed the reliable capabilities of autonomous equipment. Accordingly, the pertinent research question is not whether humans should be replaced, but how their tasks can be redesigned without exchanging familiar hazards for poorly understood ones [3,10,11].

The Saudi setting sharpens this problem. The Kingdom's logistics strategy prioritises distribution connectivity, modern logistics facilities and service quality, while warehouse operators must accommodate heterogeneous products, multilingual workforces, variable automation maturity and demanding thermal conditions. These contextual features make the transfer of evidence from European and North American facilities conditional rather than automatic. An autonomous vehicle validated on an uncluttered test track may behave differently around temporary pallets, reflective wrapping, mixed vehicle traffic or inexperienced contract workers. Sustainable operations therefore require deployment rules that adapt international technical insight to local layouts, procurement arrangements and worker needs [2,30].

Existing scholarship supplies useful but fragmented knowledge. Studies of coordinated picking address throughput and robot scheduling [5,9,12]; ergonomic investigations identify posture and fatigue mechanisms [4,7,18]; robotics research examines localisation and obstacle avoidance [6,28]; and broader occupational reviews identify safety management concerns [8,16,17]. Technology-acceptance studies add the finding that the usefulness of robots depends partly on whether workers can predict, understand and challenge their behaviour [11,21]. Treating these literatures separately conceals the dependencies among operational efficiency, physical risk, psychological load and algorithmic control.

This review aims to develop an integrated, safety-first account of collaborative warehousing suitable for decisions in Saudi logistics operations. Its objectives are to identify distinct mechanisms through which robotic assistance changes occupational exposure; examine the technical and organisational controls needed for trustworthy AI-supported safety; compare evidence across field studies, laboratory research, simulations and reviews; and formulate measurable priorities for implementation and future investigation. Three questions guide the argument: Which hazards are reduced or redistributed through collaboration? Under what conditions can AI improve protection without undermining human authority? How should Saudi facilities evaluate joint safety, productivity and sustainability outcomes?

2. Review design and analytical approach

A structured critical integrative review was selected because the research question spans engineering performance, operational decision-making and worker wellbeing. Unlike a narrowly defined effectiveness review, this approach permits comparisons among controlled experiments, warehouse observations, optimisation models, systematic reviews and conceptual safety analyses. It remains a synthesis of published evidence, not a new field trial, quantitative meta-analysis or survey. Consequently, this article reports neither invented participant numbers nor artificial significance tests, pooled accident rates or search-flow totals.

The evidence base comprises thirty publications issued between 2020 and 2025, identified through targeted scholarly searches, publisher records, relevant citation trails and the supplied warehouse-automation article's bibliographic context. Search concepts included "collaborative warehouse", "human robot", "order picking", "autonomous mobile robot", "ergonomic workload", "safety", "collision avoidance", "trust" and "logistics sustainability". Studies were retained when their verified DOI resolved to a relevant publication and their reported methods or arguments contributed to one of the review questions. General robotics findings were included only where their mechanisms plausibly transferred to warehouse settings. The resulting set is purposive, not claimed to exhaust every indexed publication.

Each source was read against a common extraction logic: operating context, human role, robot function, research design, reported outcomes, safety mechanism and transferability limitations. Evidence was then compared across six analytical dimensions: task allocation, physical ergonomics, spatial interaction, algorithmic reliability, psychosocial effects and governance. Contradictory findings were retained rather than averaged away. Laboratory evidence was considered useful for isolating mechanisms but insufficient to predict site-wide accident rates; simulation findings were treated as conditional on model assumptions; and field observations were assessed for confounding and generalisability [10,19,22].

The method has explicit boundaries. No full-text institutional database census or reproducible exhaustive search export was available; therefore publication counts cannot establish the prevalence of findings in the global literature. Papers were not converted into a numerical quality score because sample sizes, methods and outcomes were highly heterogeneous. Instead, the synthesis differentiates observed workplace effects from theoretical propositions, and marks Saudi deployment recommendations as context-sensitive inferences. This transparency matters because

presenting conceptual frameworks as experimentally validated systems would overstate the present state of the evidence.

The comparison distinguishes outcomes by observation level. An optimised route is a machine-level outcome, a lower picking cycle time is a task-level outcome, and fewer recorded injuries is a population-level outcome. Moving between levels requires evidence rather than inference. The included resilience and automation reviews identify useful organisational capabilities, yet their findings cannot confirm safety improvements at individual workstations [13,14]. Similarly, a warehouse drone review addresses inspection and inventory possibilities, but airborne navigation introduces distinct overhead, charging and pedestrian-interface considerations [15]. These distinctions informed the synthesis: proposals were retained only when their operational mechanism and unresolved assumptions could both be stated explicitly.

3. Conceptual foundations of safe collaboration

Human-robot collaboration denotes coordinated task performance in which people and robotic systems share work goals and may occupy overlapping environments. It should be distinguished from simple coexistence, where tasks are independent, and from complete autonomy, where no employee is expected to intervene during normal execution. The distinctions matter for risk ownership: a mobile robot following a picker requires joint movement decisions, whereas a segregated conveyor primarily requires interface protection. Contemporary warehouse arrangements frequently combine these modes, causing safety boundaries to shift during replenishment, fault recovery and peak-demand adjustments [10,11,27].

A socio-technical perspective provides the strongest integrating explanation. Machinery, people, tasks, incentives and supervisory procedures form one operational system; modifying robot speeds or picking sequences changes human movement, attention and workload simultaneously. Sony and Naik argue that digital industrial transformation must be designed around interacting social and technical components [1]. Logistics-service-provider research similarly identifies leadership, resource constraints and organisational adaptation as determinants of successful digital change [2]. For warehouse safety, this means that protective control cannot be relegated to a robot vendor while shift scheduling, incentives and emergency procedures remain unchanged.

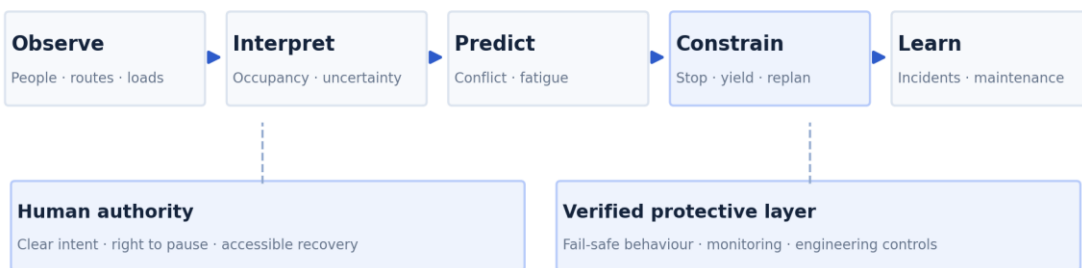
The principal causal pathways are exposure substitution, exposure redistribution and exposure amplification. Substitution occurs when robots remove repeated long-distance walking or transport. Redistribution arises when mobile platforms deliver more items per hour to stationary pickers,

potentially concentrating upper-limb repetition. Amplification can occur when high throughput targets increase pace, robots produce ambiguous right-of-way decisions or operators must recover frequent software faults. Evidence of lower transport time therefore does not establish lower musculoskeletal risk, and evidence of safer robot motion does not establish reduced cumulative fatigue [9,18,24].

Figure 1 illustrates the proposed safety architecture. Its defining distinction is between AI as a predictive advisory layer and certified safeguards as independently enforceable constraints. A model may estimate intent, congestion or fatigue, but safety-critical stopping and protective separation should not depend exclusively on an opaque probabilistic score. Operator override, physical emergency stopping, manual recovery and incident learning must remain accessible under communications loss or model drift. Trust should be calibrated to actual capability rather than manufactured through reassuring interface messages [17,21,23].

Human-centred warehouse safety architecture

Five linked functions, with independently enforceable protection



Conceptual synthesis; arrows denote operational dependencies, not measured causal effects.

Figure 1. Layered governance and control architecture for human–robot warehouse safety (original conceptual diagram).

4. Occupational risk pathways in collaborative warehouses

4.1 Physical contact, traffic and interaction geometry

Mobile robots change the geometry of warehouse risk. Human pedestrians, forklifts and autonomous vehicles encounter one another at aisle crossings, loading interfaces and blind corners where line of sight and braking room may be restricted. Risk cannot be expressed simply as the probability of contact; its severity also depends on relative velocity, payload mass, stopping distance, floor friction and opportunities for escape. Warehouse localisation research demonstrates that detecting the position of workers can support more selective protection than stopping an entire

robotic fleet, although sensing integrity becomes critical under occlusion or changing infrastructure [6].

Obstacle avoidance research describes methods for adapting robot paths in dynamic shared environments [28]. However, successful obstacle avoidance in a reported experiment does not prove robustness to damaged sensors, unusual clothing, transparent obstacles or crowded seasonal operations. Facility design should therefore make safe behaviour possible before adding sophisticated prediction: separate heavy-vehicle lanes, protect pedestrian crossings, manage temporary storage, prohibit unsafe shortcuts and preserve clear sightlines. Route optimisation may then operate within these bounded conditions. Where shared zones are unavoidable, robot behaviour should communicate a clear priority rule and provide sufficient time for human response [17,20].

4.2 Musculoskeletal exposure and fatigue

The ergonomic case for collaboration is persuasive but not unconditional. Assisted order-picking investigations find potential reductions in travel effort or handling burden, while posture-related benefits can be smaller and sensitive to shelf height, load weight and team configuration [10,19]. Experimental work on subjective exertion shows that item attributes and reach demands remain important even when warehouses deploy modern operating systems [20]. Designers who select robots solely through distance travelled per pick may therefore miss sustained shoulder elevation, trunk rotation, static standing and repeated grasping at concentrated workstations.

Fatigue-aware coordination provides an alternative optimisation objective. Sheu and Choi formulate order-fulfilment decisions that adapt robotic deliveries to human fatigue, while Niu and colleagues demonstrate how worker discomfort can enter multi-agent task assignment [4,9]. These are useful demonstrations of design logic, not licences to infer a clinically validated reduction in occupational disease. The difficult implementation question is how to observe fatigue accurately without intrusive surveillance. Practical pilots should use task-level ergonomic assessments and voluntary worker feedback, examine whether risks move between body regions and ensure that rest opportunities are not overridden by urgent dispatch demands.

4.3 Cognitive workload, trust and psychosocial conditions

Cognitive and social effects are easy to neglect because they are less visible than collisions. Workers may need to monitor robot status, interpret alerts, anticipate route changes and manage system failures while meeting ordinary picking quotas. A laboratory study of robot proximity

found that behavioural interference may appear even when measured stress changes are limited [23]. Acceptance research further suggests that perceived usefulness and predictability influence whether workers cooperate with automated systems [11]. These findings imply a need to measure attention demands, perceived control and confidence separately instead of assuming that a successful operational trial demonstrates comfortable working conditions.

Communication design is a direct safety intervention. Bhattathiri and colleagues evaluated robot intention displays at a simulated warehouse intersection and reported benefits for trust and task execution [18]. Their results support comprehensible signals but do not establish that any particular interface prevents real-world injury. Symbols, lights or spoken warnings should convey unambiguous direction, stopping state and intended manoeuvre; these signals also require testing with the languages and literacy levels present on site. Too many alarms can create habituation, while anthropomorphic presentation risks encouraging workers to treat machines as more competent than their limitations justify.

4.4 Evidence of system-level outcomes

Operational studies show that collaboration can improve certain measures of picking performance, although effect sizes differ across technology, product profile and layout. A field investigation of robot-assisted retail picking identifies substantial task-time benefits in suitable settings [22], whereas comparisons of human–robot and human–manual teams reveal more modest average gains and variable learning effects [24]. Amazon’s deployed robotic picking algorithm illustrates substantial possible reductions in robotic travel without demonstrating an independent safety outcome [12]. Taken together, these studies challenge the widespread assumption that productivity improvements and injury prevention arise through the same mechanism.

The broader labour-economics evidence also requires careful interpretation. Gihleb and colleagues associate industrial robot exposure with changes in occupational injury and physical work intensity, but their study examines economy-wide contexts rather than controlled Saudi warehouses [8]. Its results cannot be transplanted into a predicted local injury percentage. Better workplace evaluations would jointly record near misses, loss-of-separation events, ergonomic exposure, physical injury, psychological workload, throughput and unplanned stoppages. Table 1 identifies the measurement implications of the contrasting evidence and avoids ranking papers through unsupported numerical precision.

Table 1. Critical evidence comparison and implications for warehouse safety

Evidence stream	What the literature supports	What remains unresolved
Worker ergonomics [9,10,19,20]	Robot assistance may reduce travel or selected exertion; shelf position and task design remain influential.	Long-term injury outcomes, workload transfer and shift-wide recovery.
Shared-space interaction [6,17,18,28]	Localisation, cautious navigation and understandable intent signals support safer coordination.	Reliability under occlusion, congestion and degraded sensors.
Operational performance [12,22,24,26]	Picking efficiency is sensitive to layouts, algorithms, worker experience and task mix.	Whether faster operations independently lower exposure and injury.
Trust and adaptation [11,21,23,29]	Communication and decision transparency can affect trust and collaboration.	Calibration of trust, longitudinal effects and worker autonomy.
System sustainability [1,2,14,30]	Integration capability, recovery and change management influence durable adoption.	Saudi-specific comparative evidence on safety, costs and resource intensity.

4.5 A hierarchy of controls for changing risk

An evaluation should ask whether a hazardous interaction can be eliminated through layout or process redesign, followed by separation, speed restriction, sensing, warnings and administrative controls. AI is useful for adaptive prediction, but it generally occupies a dependent layer that requires reliable sensors and communication. At pedestrian crossings, a protected geometry can remain effective when a forecast fails; a screen displaying confidence scores cannot guarantee physical clearance. Drone inspection should likewise be assessed for overhead collision and battery risks rather than assumed harmless because it eliminates ladder use [15,17]. Trust-aware picking models suggest that feedback can improve collaboration, yet human psychological states are inferred imperfectly and should not dictate unsupervised protective thresholds [29].

5. AI-enabled safety management and assurance

5.1 Perception, prediction and decision boundaries

An AI-enabled warehouse safety system typically fuses observations from cameras, distance sensors, robot localisation, machine telemetry, wearable devices and warehouse-management events. Computer vision can help recognise pedestrian occupancy; motion prediction can identify possible path conflicts; anomaly models can flag unusual stopping patterns or sensor degradation. Yet each prediction is contingent on its training distribution. Low light, reflective packaging, dust, unusual pallet geometry or interrupted wireless signals can make an apparently accurate system unreliable at precisely the moment protection is needed. Sensor validation should consequently include adverse scenarios, not only average detection accuracy [6,17,28].

Prediction needs a disciplined handover into action. High-confidence warnings may trigger route replanning or controlled slowdown, but low-confidence or missing observations should also result in a safe operational response. Operators need visible explanations for stops and practical mechanisms to report false alarms, blind spots and unsafe routes. An incident that causes a robot to stop unexpectedly may be a successful protective action, a sensing defect or a production disruption; the distinction must be recorded for subsequent analysis. A robust architecture therefore combines real-time monitoring with conservative fallback states and independent protective functions rather than substituting machine-learning forecasts for basic safeguards [16,17].

5.2 Task allocation and human authority

Scheduling algorithms decide who performs which work, in what order and at what pace. This capability makes operational safety an explicit objective of AI optimisation. Multi-robot planning research shows the value of coordinating transport tasks across dynamic resources [3]; human-oriented models add fatigue, performance variability and worker comfort to these decisions [4,9]. A safety-aware allocator should balance physical capacity, travel demands, operator competence, recovery time and congestion while respecting hard constraints. A seemingly optimal assignment that repeatedly allocates heavy items to the same worker is not sustainable merely because total travel distance falls.

Human supervision requires more than a stop button. Employees must understand when robots will yield, how to suspend a task, who authorises resumption and how exceptions affect performance targets. Transparency research indicates that explanations can stabilise trust when autonomous behaviour diverges from user expectations [21]. A worker who pauses a robot after observing a

hazard should not be penalised through an opaque productivity metric. Equally, manual override must be controlled so that it cannot silently disable a critical protection. These design choices turn worker participation from a symbolic consultation into an enforceable operating condition.

5.3 Assurance, cybersecurity and change control

AI assurance must span design, commissioning and routine operation. Before commissioning, teams should define foreseeable misuse, pedestrian access, loading conditions, maintenance modes and failure combinations. Validation should include real people, cluttered aisles, shift changes and emergency recovery, with performance judged against predetermined safety criteria. After deployment, changes to algorithms, maps, wireless networks and robot payloads must trigger reassessment. The conventional acceptance test at installation is inadequate when software and warehouse layouts change continuously [16,17].

Digital interconnection introduces additional hazards. A compromised map, delayed location message or corrupted robot instruction may produce unsafe movement even when the mechanical system is sound. Facilities therefore need segregated operational networks, controlled access, authenticated software updates, secure logging and a tested manual operating plan. These are recommendations derived from the operational dependencies described in the reviewed literature, not claims that the included studies quantified Saudi cyber incidents. Asset owners, technology vendors and warehouse management must agree on responsibility for patching, sensor calibration and incident escalation before robots operate among workers [2,27].

5.4 Environmental and social sustainability

Sustainability should be measured as a joint outcome rather than inferred from automation. Mobile robots consume electricity and require batteries, charging facilities and replacement parts; nevertheless, efficient routing may reduce unnecessary vehicle movements, damaged stock and rework. Lower energy consumption per movement can be offset by higher order volume or excessive vehicle idling. A credible assessment should use energy per fulfilled order, lifecycle maintenance demand, inventory damage and recovery costs, alongside employee health measures. Saudi research on greener logistics provides contextual support for linking technological transformation to environmental and economic objectives, but does not establish that collaboration independently cuts emissions [30].

Social sustainability involves injury prevention, equitable skills development and meaningful input into work redesign. Wearable monitoring requires clear limits on collection, retention, worker

access and secondary use. Staff should be able to report safety concerns without retaliation or pressure to accept an unsafe pace. Supervisors need training in risk interpretation, not only dashboard use. New technical roles can support professional development, but benefits will be uneven if temporary staff and outsourced teams receive weaker induction. This makes contractor governance and accessible multilingual instruction essential elements of warehouse automation policy [1,2,27].

5.5 Evaluation of the entire safety chain

A validation exercise traces one credible event from detection through response and recovery. Consider a worker entering a crossing while an autonomous vehicle carries an obscuring load. The assessment should establish whether the sensors recognise the person, whether uncertainty is represented, whether braking performance remains adequate, whether intentions are intelligible and whether the supervisor can investigate an abnormal stop. Tests should be repeated after changes in lighting, route maps and load dimensions. Separate performance metrics are necessary for detection, decision delay, stopping reliability and worker understanding; combining them into a single accuracy score could mask a critical failure. Logged near misses should inform engineering changes before they are converted into individual disciplinary statistics. This systems perspective makes auditability central to both trust and learning.

6. Implications for Saudi logistics facilities

6.1 Deployment conditions and operational fit

Saudi warehouses are not a single technological category. Food distribution, e-commerce fulfilment, industrial spare parts, pharmaceuticals and bonded logistics differ in item handling, storage conditions and movement patterns. Heat-sensitive batteries, chilled storage, mixed manual and automated operations, and frequent reconfiguration may change sensor reliability and maintenance demands. These should be treated as site-specific assessment questions rather than universal national characteristics. A defensible pilot should begin with mapped pedestrian and vehicle flows, representative product profiles, incident history, job demands and the available technical competence of each shift.

The preferred first use case is not necessarily the task with the largest predicted labour saving. Repetitive cart movement on a clearly marked route may yield a simpler safety case than fully autonomous handling at crowded dispatch doors. Selection should consider traffic complexity, worker exposure, ease of segregation, equipment support and resilience during failure. Long-term

procurement ought to evaluate service documentation, integration interfaces, availability of replacement components and the ability of local teams to sustain the system. A demonstration with no credible maintenance or manual fallback plan creates technological dependence rather than operational resilience [2,3,14,26].

6.2 A staged implementation pathway

Figure 2 proposes four deployment gates: diagnose, design, validate and govern. At diagnosis, joint worker–engineer teams identify task boundaries and hazard scenarios. At design, the team selects operating modes, safety functions, route markings, communication signals and training materials. Validation tests credible abnormalities, including unexpected pedestrian entry and communications failure. Governance adds version tracking, periodic ergonomic review, incident learning and independent approval of material changes. Passing a gate requires written evidence, not solely a favourable vendor demonstration or the achievement of a productivity threshold.



Figure 2. Proposed safety-gated implementation and continuous evaluation pathway (original conceptual diagram).

The initial baseline should cover several representative shifts before implementation. Where available, direct measures should include pedestrian–robot interactions, manual handling, rest-break compliance, near misses, product damage, productivity and electricity use. After deployment, equivalent observation periods should account for demand mix, staffing, seasonality and shifts. Workers need to distinguish between hazards introduced by the equipment and hazards that were previously underreported. This paired operational design would be more informative than a one-time satisfaction survey, although causal attribution would still require credible comparison groups or staged roll-outs [10,22,24].

6.3 Workforce capability and governance

Operational competence includes recognising normal robot intentions, practising safe crossing, responding to degraded sensors and recovering from mispicks or blocked aisles. Supervisory competence includes investigating alerts, avoiding unrealistic pacing and judging whether performance gains depend on unsafe shortcuts. Technicians require diagnostic skills without uncontrolled bypass of protective functions. Training should be demonstrated through practical competency checks and repeated when routes or operating modes change. Interface messages should be field-tested with employees, including agency workers and staff whose strongest language is not English [11,18,25].

A clear accountability structure is indispensable for multi-tenant facilities and outsourced operators. The host warehouse should own the overall traffic plan and emergency arrangements; robot vendors should provide technical evidence and maintenance support; employers should supervise task allocation and worker induction; and a designated safety authority should approve operational changes. Contracts must clarify the ownership of risk data and responsibilities during shared-site incidents. Without this governance, individual subsystems may meet their local targets while the integrated warehouse remains vulnerable to unsafe handovers or unreported failures [2,16].

Table 2 translates the review into an implementation-oriented scorecard. It deliberately identifies indicators and decision rules rather than target percentages, because credible thresholds require a local risk assessment, worker consultation and equipment-specific validation. National ambitions for efficient logistics are compatible with this cautious approach: preserving operational continuity and human capability should be understood as conditions of competitiveness, not friction imposed on technology adoption. The proposed framework is intended for adaptation through Saudi site pilots, not as evidence that its effectiveness has already been demonstrated in the Kingdom.

Table 2. Recommended site-level indicators and assurance decisions

Domain	Candidate measure	Assurance decision
Collision exposure	Near misses, pedestrian–robot proximity events, protective stops	Verify sensor coverage and safe stopping under realistic failures.

Ergonomics	Observed postures, manual loads, repetition, recovery time	Compare exposures by body region and worker group.
Worker experience	Reported predictability, trust, workload, override accessibility	Investigate inconsistent comprehension and pressure to bypass controls.
Operational continuity	Orders/hour, picking errors, time to restore safe service	Reject efficiency gains dependent on unsafe shortcuts.
Environmental burden	Electricity per fulfilled order, battery maintenance, damaged stock	Assess absolute consumption and demand-adjusted intensity.
Governance	Training completion, maintenance defects, software changes, audit findings	Revalidate safety case after material equipment or process change.

7. Research priorities and limitations

The central research gap is the scarcity of studies observing physical safety, workload, trust and environmental outcomes together over sustained warehouse operation. Many reports concentrate on route efficiency or experimentally controlled interfaces; others describe ergonomic conditions without measuring system reliability. More rigorous evaluations should predefine primary safety outcomes, compare operations before and after deployment, and distinguish exposure to hazards from actual injury incidence. Because serious injuries are comparatively rare and confounded by reporting practices, near-miss measures and observed separation events may offer more sensitive intermediate outcomes when their definitions are stable [8,18,24].

Future Saudi research should use multi-site designs that represent different climates, inventory types, workforce arrangements and automation levels. A phased roll-out could compare matched aisles or shifts while recording seasonal demand and concurrent process changes. Researchers should assess subgroup differences without treating nationality, contract type or language as proxies for ability. Participatory design studies could investigate whether workers understand robot intentions, whether protective alerts are actionable and whether reporting systems capture quiet failures. Independent technical audits should test performance after model updates, altered routes and reduced sensor visibility [6,17,28].

Another unresolved question concerns which AI decisions should remain advisory. Predicting discomfort, forecasting pedestrian movement and reallocating picks all rely on uncertain behavioural models, whereas collision protection requires demonstrable performance under bounded conditions. Future work should test hybrid architectures that optimise operations inside hard safety constraints, with audited explanations and explicit fallback protocols. Researchers should publish failure modes and negative outcomes rather than only successful demonstrations. Comparisons should include operating energy, battery replacement, worker autonomy and recovery time so that environmental or economic gains are not purchased through hidden social costs [4,9,21].

The limitations of this review affect the strength of its recommendations. The thirty-source corpus was constructed purposively and is not an exhaustive inventory; differences in terminology may have left relevant contributions outside the set. Several included studies use simulations or limited laboratory samples, making their transfer to mixed-traffic warehouses uncertain. Most evidence originates outside Saudi Arabia, while the national adaptation proposed here remains a conceptual design. No injury reduction, emission saving or local return on investment is asserted as an established finding. These limitations identify testable propositions and prevent inappropriate certainty in procurement decisions.

8. Conclusion

Human–robot collaboration can contribute to sustainable warehousing when it removes burdensome transport, clarifies shared tasks and enables workers to retain meaningful control over changing operations. Yet robots also redistribute exposure toward crossings, concentrated picking stations, monitoring tasks and complex recovery work. The reviewed evidence therefore supports neither blanket technological optimism nor rejection of collaborative automation. Benefits depend on operating context, work design and the discipline with which safeguards are tested and maintained.

AI offers useful capabilities for sensing, route coordination, workload-aware scheduling and operational learning, but predictive sophistication cannot substitute for reliable protective functions, clear responsibility and conservative responses to uncertainty. Saudi logistics operators should begin with feasible use cases, evaluate authentic worker experiences, preserve safe manual alternatives and measure human, operational and environmental outcomes jointly. The proposed layered architecture and staged evaluation model provide a research-informed basis for this

process. Their effectiveness requires independent validation under actual Saudi warehouse conditions before any claim of improved safety or sustainability can be justified.

References

- [1] Sony, M.; Naik, S. (2020). Industry 4.0 integration with socio-technical systems theory: A systematic review and proposed theoretical model. *Technology in Society*, 61, 101248. <https://doi.org/10.1016/j.techsoc.2020.101248>
- [2] Cichosz, M.; Wallenburg, C.M.; Knemeyer, A.M. (2020). Digital transformation at logistics service providers: Barriers, success factors and leading practices. *The International Journal of Logistics Management*, 31, 209–238. <https://doi.org/10.1108/IJLM-08-2019-0229>
- [3] Bolu, A.; Korçak, Ö. (2021). Adaptive task planning for multi-robot smart warehouse. *IEEE Access*, 9, 27346–27358. <https://doi.org/10.1109/ACCESS.2021.3058190>
- [4] Niu, Y.; Schulte, F.; Negenborn, R.R. (2021). Human aspects in collaborative order picking: Letting robotic agents learn about human discomfort. *Procedia Computer Science*, 180, 877–886. <https://doi.org/10.1016/j.procs.2021.01.338>
- [5] Wang, Z.; Sheu, J.-B.; Teo, C.-P.; Xue, G. (2022). Robot scheduling for mobile-rack warehouses: Human–robot coordinated order picking systems. *Production and Operations Management*, 31. <https://doi.org/10.1111/poms.13406>
- [6] Popović, G.; Cvišić, I.; Écorchard, G.; et al. (2022). Human localization in robotized warehouses based on stereo odometry and ground-marker fusion. *Robotics and Computer-Integrated Manufacturing*, 73, 102241. <https://doi.org/10.1016/j.rcim.2021.102241>
- [7] Ciccarelli, M.; Papetti, A.; Scoccia, C.; et al. (2022). A system to improve the physical ergonomics in human–robot collaboration. *Procedia Computer Science*, 200, 689–698. <https://doi.org/10.1016/j.procs.2022.01.267>
- [8] Gihleb, R.; Giuntella, O.; Stella, L.; Wang, T. (2022). Industrial robots, workers' safety, and health. *Labour Economics*, 78, 102205. <https://doi.org/10.1016/j.labeco.2022.102205>
- [9] Sheu, J.-B.; Choi, T.-M. (2023). Can we work more safely and healthily with robot partners? A human-friendly robot–human-coordinated order fulfillment scheme. *Production and Operations Management*, 32, 794–812. <https://doi.org/10.1111/poms.13899>
- [10] Pasparakis, A.; De Vries, J.; De Koster, R. (2023). Assessing the impact of human–robot collaborative order picking systems on warehouse workers. *International Journal of Production Research*, 61, 7776–7790. <https://doi.org/10.1080/00207543.2023.2183343>

- [11] Jacob, F.; Grosse, E.H.; Morana, S.; König, C.J. (2023). Picking with a robot colleague: A systematic literature review and evaluation of technology acceptance in human–robot collaborative warehouses. *Computers & Industrial Engineering*, 180, 109262. <https://doi.org/10.1016/j.cie.2023.109262>
- [12] Allgor, R.; Cezik, T.; Chen, D. (2023). Algorithm for robotic picking in Amazon fulfillment centers enables humans and robots to work together effectively. *INFORMS Journal on Applied Analytics*, 53, 266–282. <https://doi.org/10.1287/inte.2022.1143>
- [13] Ferreira, B.; Reis, J. (2023). A systematic literature review on the application of automation in logistics. *Logistics*, 7, 80. <https://doi.org/10.3390/logistics7040080>
- [14] Colabianchi, S.; Bernabei, M.; Costantino, F.; et al. (2023). MARLIN method: Enhancing warehouse resilience in response to disruptions. *Logistics*, 7, 95. <https://doi.org/10.3390/logistics7040095>
- [15] Malang, C.; Charoenkwan, P.; Wudhikarn, R. (2023). Implementation and critical factors of unmanned aerial vehicle (UAV) in warehouse management: A systematic literature review. *Drones*, 7, 80. <https://doi.org/10.3390/drones7020080>
- [16] Giallanza, A.; La Scalia, G.; Micale, R.; La Fata, C.M. (2024). Occupational health and safety issues in human-robot collaboration: State of the art and open challenges. *Safety Science*, 169, 106313. <https://doi.org/10.1016/j.ssci.2023.106313>
- [17] Li, W.; Hu, Y.; Zhou, Y.; Pham, D.T. (2024). Safe human–robot collaboration for industrial settings: A survey. *Journal of Intelligent Manufacturing*, 35, 2235–2261. <https://doi.org/10.1007/s10845-023-02159-4>
- [18] Bhattathiri, S.S.; Bogovik, A.; Abdollahi, M.; et al. (2024). Unlocking human-robot synergy: The power of intent communication in warehouse robotics. *Applied Ergonomics*, 117, 104248. <https://doi.org/10.1016/j.apergo.2024.104248>
- [19] Zhang, M.; Marolt, J.; Bencak, P.; Grosse, E.; Lerher, T. (2024). Ergonomic evaluation of human–robot collaborative order picking: A combined laboratory and simulation study. *IFAC-PapersOnLine*, 58, 1042–1047. <https://doi.org/10.1016/j.ifacol.2024.09.137>
- [20] De Lombaert, T.; Braekers, K.; De Koster, R.; Ramaekers, K. (2024). What makes order picking so physically demanding? Ergonomic evidence from a large-scale lab experiment using subjective metrics. *IFAC-PapersOnLine*, 58, 181–186. <https://doi.org/10.1016/j.ifacol.2024.09.139>

- [21] Kox, E.S.; van den Boogaard, J.; Turjaka, V.; Kerstholt, J.H. (2024). The journey or the destination: The impact of transparency and goal attainment on trust in human-robot teams. *ACM Transactions on Human-Robot Interaction*, 14, article 23. <https://doi.org/10.1145/3702245>
- [22] Koreis, J.; Loske, D.; Klumpp, M. (2025). Together, we travel: Empirical insights on human-robot collaborative order picking for retail warehousing. *The International Journal of Logistics Management*, 36, 1–20. <https://doi.org/10.1108/IJLM-03-2023-0127>
- [23] Rinaldi, M.; Natale, C.; Fera, M.; et al. (2024). Evaluation of mental stress in human-robot interaction: An explorative study. *Procedia Computer Science*, 232, 726–735. <https://doi.org/10.1016/j.procs.2024.01.072>
- [24] Koreis, J. (2025). Human–robot vs. human–manual teams: Understanding the dynamics of experience and performance variability in picker-to-parts order picking. *Computers & Industrial Engineering*, 200, 110750. <https://doi.org/10.1016/j.cie.2024.110750>
- [25] Epe, M.; Azmat, M.; Islam, D.M.Z.; Khalid, R. (2024). Use of smart glasses for boosting warehouse efficiency: Implications for change management. *Logistics*, 8, 106. <https://doi.org/10.3390/logistics8040106>
- [26] Garg, V.; Maywald, J.D.; Naman, M. (2025). Optimising human-robot collaboration for efficiency in retail warehousing. *International Journal of Retail & Distribution Management*, 53, 1123–1138. <https://doi.org/10.1108/IJRDM-01-2025-0052>
- [27] Zheng, T.; Glock, C.H.; Neumann, P. (2025). Human robot collaboration in warehousing operations: A sociotechnical analysis. *IFAC-PapersOnLine*, 59, 2634–2639. <https://doi.org/10.1016/j.ifacol.2025.09.443>
- [28] Sousa, L.C.; Silva, Y.M.R.; Schettino, V.B.; et al. (2025). Obstacle avoidance technique for mobile robots at autonomous human-robot collaborative warehouse environments. *Sensors*, 25, 2387. <https://doi.org/10.3390/s25082387>
- [29] Liu, Y.; Guo, F.; Ma, Y. (2025). A Stackelberg trust-based human–robot collaboration framework for warehouse picking. *Systems*, 13, 348. <https://doi.org/10.3390/systems13050348>
- [30] Alasmari, T.; Alzahrani, A. (2025). Saudi Arabia’s shifts towards green and sustainable logistics: Bibliometric and machine learning-based insights and forecasts. *Journal of Cleaner Production*, 509, 145577. <https://doi.org/10.1016/j.jclepro.2025.145577>