Leveraging AI for Smart Manufacturing in SMEs

Productivity *Insights*

Vol. 7-3

Asian Productivity Organization

The Asian Productivity Organization (APO) is an intergovernmental organization that promotes productivity as a key enabler for socioeconomic development and organizational and enterprise growth. It promotes productivity improvement tools, techniques, and methodologies; supports the national productivity organizations of its members; conducts research on productivity trends; and disseminates productivity information, analyses, and data. The APO was established in 1961 and comprises 21 members.

APO Members
Bangladesh, Cambodia, Republic of China, Fiji, Hong Kong, India, Indonesia, Islamic Republic of Iran, Japan, Republic of Korea, Lao PDR, Malaysia, Mongolia, Nepal, Pakistan, Philippines, Singapore, Sri Lanka, Thailand, Turkiye, and Vietnam.

LEVERAGING AI FOR SMART MANUFACTURING IN SMEs

JUNE 2026 | ASIAN PRODUCTIVITY ORGANIZATION

PRODUCTIVITY INSIGHTS Vol. 7-3

Leveraging AI for Smart Manufacturing in SMEs

Dr. Huiyi Fan wrote this publication.

First edition published in Japan
by the Asian Productivity Organization
1-24-1 Hongo, Bunkyo-ku
Tokyo 113-0033, Japan
www.apo-tokyo.org

Designed by Convert To Curves Media

CONTENTS

CONTENTS

PREFACE

The Productivity Insights (P-Insights) series is an extension of the Productivity Talk (P-Talk) series, one of the APO Secretariat's digital information initiatives. Originally designed to maximize the full potential of the APO's digital outreach, the interactive, livestreamed P-Talks bring together practitioners, experts, policymakers, and ordinary citizens from all walks of life with a passion for productivity to share their experiences, views, and practical tips on productivity improvement.

With speakers from every corner of the world, the P-Talks effectively convey productivity information to APO members and beyond. However, it was recognized that many of the P-Talk speakers had much more to offer beyond the 60-minute presentations and Q&A sessions that are the hallmarks of the series. To take full advantage of their broad knowledge and expertise, the APO invites some to elaborate on their P-Talks, resulting in this publication. It is hoped that the P-Insights series will give readers a deeper understanding of the practices and applications of productivity.

INTRODUCTION

Manufacturing in Asia is rapidly evolving. Data and artificial intelligence are now used in daily production. Small experiments, such as a pilot that anticipates bearing failures or a vision model that detects surface scratches, are now considered standard work. The benefits are concrete: fewer unplanned stops, higher first-pass yield, and lower energy per unit. Yet many AI initiatives stall after the first proof of concept. The reasons are rarely algorithms alone. Success depends on pairing technology with sound operations, clear governance, and practical training for the people who will use the tools.

This report takes a plain-language approach for small and medium-sized enterprise (SME) leaders and policymakers. It condenses a broader comparative analysis into four practical strategies, or "plays," for applying AI in manufacturing. Each play is designed for resource-constrained settings, where firms may have limited budgets, limited technical staff, or limited time for experimentation. The emphasis is on doing the first project well, measuring the results that matter most, and then repeating successful practices from one production unit or work area in the next. Rather than presenting a long list of technical buzzwords, these plays focus on a small set of practical capabilities that have already shown value on the factory floor.

The four plays are (1) quick wins with AI visual inspection and predictive maintenance, (2) integration from single tasks to systems using a digital twin as the cyber-physical production system core, (3) human-in-the-loop operations that foster adoption and trust, and (4) data-secure collaboration across sites using federated learning.

This report highlights lessons that SMEs can implement next week, not next year, by referencing cases from the author's Productivity Talk (Asian Productivity Organization, 2025). It also offers a straightforward but thorough preparation method called the TOPS readiness check. Even for a small pilot, TOPS assists leaders in determining whether the technology, organization, people, and strategy are prepared. A 90-day lighthouse plan outlines how to demonstrate value, stabilize routine tasks, and determine whether to scale or cease. The chapter "Working with Partners: A Short Buyer's Guide" gives advice on collaborating with partners and vendors and summarizes typical pitfalls.

QUICK WINS: PREDICTIVE MAINTENANCE AND AI VISUAL INSPECTION

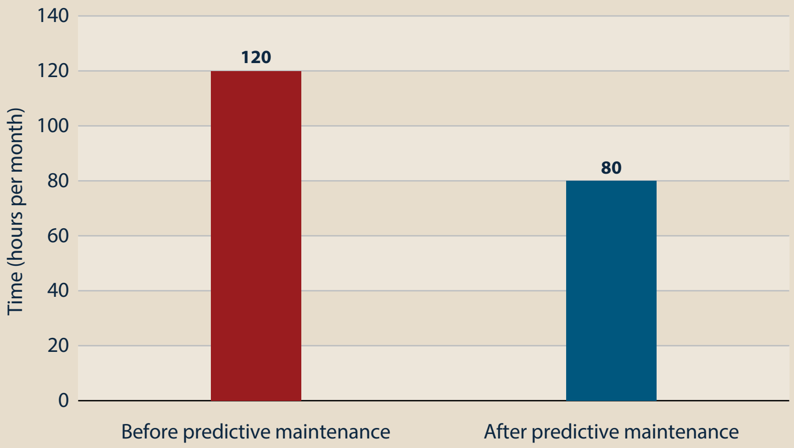
Why begin with predictive maintenance and AI visual inspection? Because they are easy to integrate into current workflows and provide value fast. Predictive maintenance forecasts failures early using sensor data such as temperature, vibration, current, and acoustics, while AI visual inspection analyzes camera images to detect manufacturing defects. For SMEs, the main challenge is often not the algorithm itself, but data readiness, labeling, and operational integration between alerts and frontline action. To illustrate the immediate operational impact, Figure 1 presents a comparative analysis of equipment downtime before and after the implementation of predictive maintenance. The baseline data (before implementation) represents the monthly average downtime recorded over a six-month historical period, while the post-implementation data reflects the average monthly downtime stabilized after a 90-day pilot phase. As shown in Figure 1, monthly downtime can decrease from 120 hours to 80 hours, representing a 33.3% reduction in unplanned interruptions. To ensure these gains are sustainable rather than one-off anomalies, we recommend confirming improvements over two consecutive four-week windows before scaling and mitigating lighting or camera drift with weekly calibration and back-labeling.

Business Case in One Paragraph

Assume that a surface-mount technology line or a computer numerical control cell experiences about 120 hours of unplanned downtime each month. Cutting that by one-third saves 40 hours. If each hour of downtime costs USD800.00 in lost output and overtime, the monthly benefit is USD32,000.00. Even after paying for sensors, a simple edge gateway, and a cloud service, the payback period can be under six months. The same logic applies to defect detection. If

FIGURE 1

MONTHLY EQUIPMENT DOWNTIME BEFORE AND AFTER PREDICTIVE MAINTENANCE IMPLEMENTATION (ILLUSTRATIVE).



Source: Author's compilation based on representative pilot case data.

first-pass yield rises from 95.5% to 98%, scrap and rework fall. For a line that ships USD1 million per month of product with a 10% gross margin, each percentage point of yield is worth USD10,000 per month. Appendix A provides a step-by-step worked example of this return-on-investment calculation that leaders can adapt to their own cost structures.

Minimal Stack That Works

A minimal stack includes one or two sensors per asset (or a small camera rig for quality control), a low-code data pipeline to stream and store time-stamped readings, a starter model (often an out-of-the-box anomaly detector or a transfer-learned classifier), and a dashboard that operators can read at a glance. Keep three alert levels (advisory, warning, stop). Every alert must map to a clear action: inspect, adjust, or call maintenance. If the system raises a false alarm, log it, fix the cause, and update the model on a weekly cycle. Appendix B describes predictive maintenance modeling approaches in greater detail, from physics-based thresholds to hybrid analytics.

Data You Need and What "Good" Looks Like

Data for at least a few weeks of normal operations, examples of faults or near misses (for predictive maintenance), and a few hundred labeled images per major defect category (for quality control) are needed. Lighting and camera distance should be fixed. For time-series analysis, ensure synchronized clocks, consistent sampling rates, and tag names that a technician can understand. If collecting historical data is difficult, start with simple rule-based triggers and evolve toward machine learning as labeled events accumulate. Appendix C provides a quality control playbook with practical guidance on image capture setup, labeling standards, and retraining loops.

Operator Experience Matters

The human-machine interface (HMI) should make the next best action obvious. Use large fonts, plain icons, and a short explanation of why the alert fired (e.g., "vibration spike above baseline for 5 minutes on spindle #2"). Add a feedback button so operators can mark false alarms. This human-in-the-loop feedback becomes training data for the next model version.

Case Snapshots

Rolls-Royce (n.d.) used AI-driven engine health monitoring to reduce the time taken to inspect aircraft engines. A 2025 Microsoft customer story reported that Rolls-Royce used AI and digital technologies to detect and prevent around 400 unplanned maintenance events annually. Foxconn Technology Group (2021) reported that its AI-based defect inspection technology reduced inspection manpower by 50% and that its AI solutions had improved reporting accuracy from 95% to 99% on production lines.

Common Pitfalls and How to Avoid Them

- **Overfitting to one machine:** Validate on a similar but different asset before declaring victory.
- **Dashboard sprawl:** Screens that resemble cockpits are ignored by operators. Less is better.
- **No maintenance playbook:** Users become irritated by alerts that do not have a standard response.

- **Camera drift:** Accuracy deteriorates in the absence of a weekly calibration schedule.

- **Hunting rare defects:** If examples are scarce, start with defect families (e.g., "scratch-like") and refine over time.

Measuring Success

Using the key performance indicator (KPI) template provided in Table 1, record both baselines and targets for downtime, mean time between failures, defect rate, first-pass yield, energy per unit, and overall equipment effectiveness. Define the baseline using the median from the most recent 8–12 weeks. Set targets such as a 10%–30% downtime reduction for predictive maintenance and a first-pass yield gain of 2.5 to 5 percentage points for quality control. Update weekly with a fixed scope (shift, machine, or product), publishing a weekly progress chart. Do not scale beyond the first cell until at least two consecutive four-week periods show sustained gains and operators sign off that the system helps their work.

TABLE 1

KEY PERFORMANCE INDICATOR TEMPLATE: ENTER BASELINES AND TARGETS FOR YOUR PILOT.

Key Performance Indicator	Baseline	Target
Downtime (hours per line per month)		
Mean time between failures (hours)		
Defect rate (%)		
First-pass yield (%)		
Energy per unit (kWh/unit)		
Overall equipment effectiveness (%)		

Source: Author.

Lessons for SMEs

- Pick one cell and one KPI class (predictive maintenance or quality control). Limit scope to two assets or one product family.

- Install only the minimum viable sensors and/or cameras. Stabilize lighting and mounting.

- Define three KPIs before day one: downtime (hours per line per month), defect rate (%), energy per unit.

- Map each alert level to a standard action. Capture operator feedback as training data.

- Aim for a 10%–30% improvement window in 90 days. Scale only after two stable cycles.

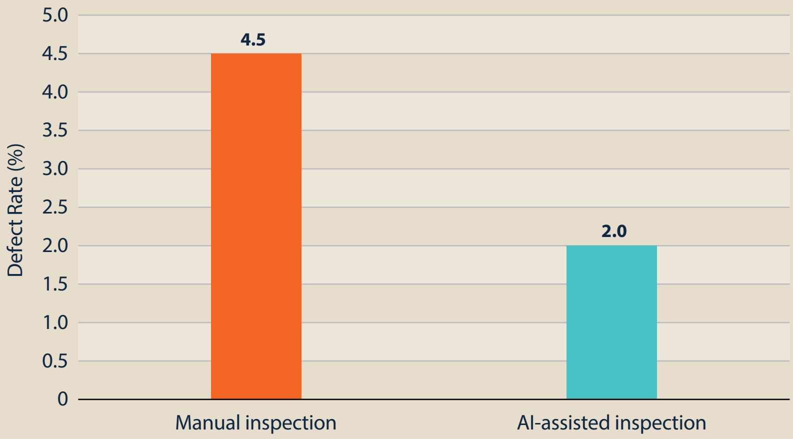
SCALE FROM TASKS TO SYSTEMS: BUILD A DIGITAL LOOP WITH DIGITAL TWINS AND CYBER-PHYSICAL PRODUCTION SYSTEMS

After quick wins, the next obstacle is fragmentation. You may have a predictive maintenance model here and a quality control model there, but planners still rely on spreadsheets and walk-arounds. System-level integration pulls these isolated pieces into a coherent loop. The digital twin serves as the core of a cyber-physical production system: a living model of the line or work cell that mirrors status, queues, quality signals, and planned work. For SMEs, the first twin can be small, comprising a bottleneck cell with two or three machines and a buffer. Although Figure 2 presents a simple quality-control example rather than the full digital loop, it illustrates the kind of measurable operational improvement that system integration is meant to support. As shown in Figure 2, the defect rate can fall from 4.5% under manual inspection to 2.0% under AI-assisted inspection, demonstrating how AI-based visual inspection improves defect detection consistency and reduces the number of faulty products that pass through the line.

In practice, when such inspection outputs are connected to a digital twin, machine signals pass through OPC UA or Modbus adapters to a work-cell-level model that tracks machine state, work in process, quality alerts, and production-order requirements in real time. This integrated data stream feeds dashboards, dynamic scheduling, and maintenance planning and supports automated updates to shop-floor standard operating procedures. Start with a small twin that can be templated and used to run what-if simulations before changeovers or planned maintenance.

FIGURE 2

AI VISUAL INSPECTION DEFECT RATE (ILLUSTRATIVE).



Source: Author.

What the First Twin Must Do

- Mirror the real-time state of machines (run, idle, fault)
- Show work in process and queue length
- Overlay quality alarms from the vision system
- Integrate work-order data so the system knows what should be happening

With these basics, supervisors can run "what-if" analyses (e.g., "If we change the shift plan or bring maintenance forward by two hours, what happens to production pace and delivery?"). Appendix D illustrates how digital twins can be used in daily huddles, weekly scheduling, and maintenance planning routines.

Data Model Basics

Create a simple tag dictionary that harmonizes sensor names and units; define the relationship between assets, stations, and products; and establish a common time reference so that events line up. Add a small library of reusable adapters (e.g., for Modbus or OPC UA) so the same twin template can be reused in another cell with minimal changes. Small manufacturers often underestimate the value of naming standards, but they are the difference between a twin you can maintain and one that breaks when a machine is replaced. Appendix E offers a data architecture quick start guide that details how to build a unified tag dictionary, establish time synchronization, and set up lightweight storage for pilot implementations.

Make or Buy?

Many SMEs can start with a commercial platform or an institute-supported tool that includes a 3D view, a run/idle/fault timeline, and basic simulation. Open-source options can lower license cost but may require more engineering time. The buying decision should be driven by the ability to reuse templates across cells, vendor support for legacy industrial communication protocols (e.g., Modbus RTU, PROFIBUS, or older proprietary interfaces commonly found in brownfield plants), and the cost of change when the line is reconfigured.

Operating Model

Determine who owns the twin. Assign a product owner on the site leadership team (not just IT) and create a small cross-functional squad (maintenance, production, quality, information technology or operational technology). Meet weekly. Keep a visible backlog of changes. Tie at least one management KPI (e.g., on-time delivery) to the twin's output so it stays relevant to the business.

Case Snapshot

Research institutes in the Asia-Pacific region, including the Industrial Technology Research Institute (Visual Components, 2016) in the Republic of China, have demonstrated flexible manufacturing systems, small flexible manufacturing cells, and cyber-physical manufacturing concepts that can inform modular SME adoption. A practical lesson from the literature is to start at the work cell level: work-cell-level digital twins have been developed as implementable entry points for manufacturing integration, root-cause analysis, and reconfiguration under frequent changeovers (Kibira et al., 2023; Detzner & Eigner, 2018; Leng et al., 2019; Leng et al., 2020).

Pitfalls to Avoid

- Building a perfect 3D visual without a solid data model.
- Treating the twin as a one-off project; it must be versioned and governed like software.
- Ignoring operators' mental models; if the screen does not match how the line is actually run, adoption will lag.

Lessons for SMEs

- Model one bottleneck cell first. Mirror status, queues, and quality alarms in real time.
- Integrate work orders so the twin understands intent (what should be happening).
- Create a tag dictionary and adapter library to reuse across cells and sites.
- Assign a business-side product owner and a cross-functional squad. Meet weekly.
- Pilot scheduling and maintenance windows in the twin before changing the real line.

PEOPLE DECIDE OUTCOMES: DESIGN HUMAN-IN-THE-LOOP OPERATIONS

The third play recognizes a simple truth: People decide whether AI succeeds. Operators, technicians, and supervisors must trust the system and find that it helps their work. Human-in-the-loop design treats AI as augmentation, not replacement. It focuses on clear interfaces, explainable alerts, and standard work that embeds new steps without adding cognitive overload.

Design Principles for Trust

- Keep the signal-to-noise ratio high, with no more than three alert levels and no flashing graphics.
- Show “why” in plain language (“temperature trend rising faster than yesterday’s baseline”).
- Default to human judgment in safety-critical moments.
- Provide a one-tap way to mark false positives.
- Rotate one operator per shift as an “AI champion” to collect feedback and coach peers.

Short, job-relevant microtraining on the line is more effective than generic classroom sessions. The dashboard itself should be used. Practice responding to each kind of alert as part of regular operations. Record that standard work, incorporate it into the onboarding process, and monitor adoption using a KPI (e.g., the percentage of alerts resolved within the allotted five minutes).

Adoption metrics include tracking opt-in rates for new features, tracking the rate of near misses reported by operators, and surveying operators every month with three questions: Does the system help you do your job? Are alerts clear? Would you recommend we use it on another line? The near-miss notes are extremely valuable because they frequently indicate the next model or HMI enhancement.

Safety and Compliance

Clearly state that AI systems do not supersede safety interlocks, lockout/tagout protocols, or regulatory checks. If a step is safety-critical, build the system so that it needs human approval before moving forward. This keeps accountability transparent and prevents a false sense of automation.

Culture Matters

Attitudes toward experimentation and hierarchy vary among nations and businesses. A courteous, iterative approach where frontline users see their input reflected in the next release builds trust quicker than any training deck. This is in line with human-centric ideologies such as Japan's Society 5.0 and the experience of lighthouse factories (WEF, 2023), which show that people embrace what they contribute to designing. Appendix F presents a change management cadence and incentive structure designed to accelerate operator adoption while respecting existing workplace culture.

Lessons for SMEs

- Collaborate with the shift team to co-design the HMI and conduct usability tests prior to launch.
- Limit the number of alert levels to three and give each alert a brief, straightforward explanation.
- Embed new steps in standard work and onboarding and track adoption as a KPI.
- Appoint shift "AI champions" and reward near-miss reporting that improves models.
- Never bypass safety and compliance; require human acknowledgment where needed.

GO FURTHER TOGETHER: FEDERATED LEARNING FOR CROSS-SITE LEARNING

Many SMEs cannot train robust models on their own data. Federated learning allows multiple factories to collaborate without sharing raw data. Each site trains a local model and sends only model updates to a neutral coordinator. The coordinator aggregates updates and sends back an improved global model. This fits scenarios where suppliers use similar equipment (e.g., motors, pumps) and face similar failure modes but cannot pool data for competitive or compliance reasons.

When Federated Learning Makes Sense

- Similar use cases across many sites
- Limited data at each site
- Strong reasons not to share raw data
- Access to a trusted hub (association, institute, or platform vendor) that can host the aggregator

A classic SME example is a shared predictive maintenance model for common motors or gearboxes across several suppliers in one cluster.

Minimal Governance to Do This Right

Establish a data schema (tags, units, sampling) that every site follows, define an evaluation set so improvements are measured apples to apples, specify how often model rounds occur (e.g., weekly), and agree on change control and rollback rules. Legal agreements should cover confidentiality, the use of updates for model improvement, and safeguards against model inversion attacks. Appendix G discusses privacy techniques, cross-site evaluation methods, and strategies for managing model drift in federated learning deployments.

Cost and Value

Federated learning does not remove the need for good local data. Sites must still collect, label, and validate. But by learning from broader patterns, each participant gets a stronger model, often with fewer false positives. In practice, the coordinator's costs (secure server, engineering time) can be shared as a service fee. For clusters supported by public institutes, this can be part of a regional competitiveness program.

A Day in the Life of a Federated Learning Round

On Monday, each site trains locally on last week's tagged data. On Tuesday, encrypted updates are uploaded to the coordinator. On Wednesday, the new global model is evaluated against the common test set. If metrics are better and within safety bounds, the model is distributed on Thursday. Friday is for operator feedback. Keep the loop simple, predictable, and transparent.

Lessons for SMEs

- Standardize tags, units, and sampling rates before the first federation round.
- Use a neutral coordinator (association, institute, or platform vendor) to host aggregation.
- Define a common evaluation set and publish weekly model metrics to all members.
- Clarify legal terms: confidentiality, update usage, rollback rules, and security controls.
- Share costs via a simple service fee; reinvest savings in better sensors and labeling.

TOPS READINESS AND A 90-DAY LIGHTHOUSE ROADMAP

Before scaling, run a quick TOPS (technology, organization, people, strategy) readiness check (Table 2). This is a short conversation that prevents expensive mistakes. *Technology* asks whether you can collect and move the right data securely. *Organization* checks whether a product owner has been named and a cross-functional squad has time allocated. *People* looks at training and adoption. *Strategy* confirms that the use case has a baseline, a target, and a decision rule for scale or stop.

TABLE 2

TOPS READINESS CHECKLIST FOR PILOTS.

Area	What Good Looks Like (Pilot)
Technology	Unified data tags; edge gateway; secure connectivity in the pilot cell
Organization	Named product owner; weekly stand-ups; cross-functional squad
People	Operator training plan; co-designed human-machine interface; adoption tracked
Strategy	Three key performance indicators with baselines; 90-day scope; budget and payback rule
Risk and security	Access control; data retention; supplier nondisclosure agreements
Governance	Change control; model versioning; audit trail
Scale plan	Criteria to expand to the second cell/site
Partners	Cloud/technology vendor and institute contacts on call

Source: Author.

Ninety-day Lighthouse Roadmap

A lighthouse project is a focused pilot designed to demonstrate practical value in a real production setting within a limited period and scope. Ninety days is long enough to capture variation in production and short enough to maintain urgency. The key is to timebox the scope, publish weekly progress, and make a scale-or-stop decision in week 12 based on measured changes in KPI performance from baseline and operator feedback.

Weeks 1–2: Confirm business case, select one cell, define three KPIs, and map data sources. Draft an installation and safety plan with maintenance and the team responsible for health, safety, and environmental compliance.

Weeks 3–6: Install and validate sensors or cameras. Set up gateways and secure connectivity. Create a basic dashboard. Collect and label initial data.

Weeks 7–10: Train and validate the starter model. Run side by side. Hold weekly feedback sessions with operators. Fix false alarms and stabilize standard work.

Weeks 11–12: Freeze the model and HMI for two weeks. Measure KPI deltas. Finalize documentation. Present results and decide whether to scale or stop.

Appendix H provides a sample 12-week progress report template showing how to structure weekly updates for the pilot team and leadership.

Working with Partners: A Short Buyer's Guide

Working with vendors and partners is unavoidable for most SMEs. A good buyer's guide helps SMEs avoid becoming dependent on a single vendor and reduces the risk of costly disappointment. Three questions matter most:

1. Can we export our data and models in open formats?
2. How easy is it to reuse templates across cells and sites?
3. What happens when the line changes, that is, how much of the configuration survives?

Checklist for Requests for Proposals

- Confirm support for common brownfield protocols (OPC UA, Modbus).
- Insist on role-based access control and audit logs.
- Ask for a model versioning scheme with rollback.
- Demand a clear separation between your data and the vendor's training data.
- Confirm that your engineers can add new tags and cameras without a service ticket.
- Ask for a short proof-of-value contract that ties payment to KPI improvement, not just milestones.

For Institutes and Public Programs Supporting SMEs

Encourage shared services (e.g., labeling, cybersecurity) and template libraries that clusters can reuse. Facilitate federated learning consortia for common equipment classes, with legal templates and a neutral coordinator.

RISK, SECURITY, AND MODEL MANAGEMENT

Risk and security cannot be afterthoughts. Threats range from obvious (malware on a gateway) to subtle (a model that degrades after a maintenance change). Practical controls include network segmentation, multifactor authentication, signed model packages, and a change log. Protect raw photos that could contain private product information for quality control systems. Logs that disclose production schedules should be kept confidential for predictive maintenance.

Model risk management defines when models can make suggestions and when they can make decisions. In safety-critical steps, require human acknowledgment. Log each model decision with its input snapshot and version. If a new version underperforms, roll back quickly. Keep an evaluation set from a “golden month” of production so you can compare apples to apples when models change.

POLICY ALIGNMENT AND ECOSYSTEM LEVERAGE

National strategies in Asia shape the context for SME adoption. The People's Republic of China's push to integrate AI into the real economy, the Republic of China's cluster-driven smart machinery strategy, Japan's human-centric Society 5.0, and the Republic of Korea's smart-factory goals with SME-oriented support all point to a common message: The ecosystem matters (State Council of the People's Republic of China, 2024; Ministry of Economic Affairs, n.d.; Cabinet Office, n.d.; Ministry of SMEs and Startups, 2024). SMEs benefit when policies provide shared services, training, and test beds that reduce the cost of first pilots and help scale successful use cases across a cluster. Table 3 provides a snapshot of selected national AI-for-manufacturing strategies, mapping core drivers, policy focuses, and SME entry points for benchmarking and for grant or consortium alignment. These strategies should be localized to fit brownfield constraints, supplier ecosystems, and the policy cadence of 2024–25.

TABLE 3

NATIONAL AI MANUFACTURING STRATEGIES (CONDENSED).

Economy	Core Driver	Policy Focus	SME Angle
People's Republic of China	Scale and technological capability	"AI Plus" across the real economy	Platform ecosystems; domestic capability building
Republic of China	Niche dominance and industrial clustering	Smart Machinery Promotion Program; cluster strength	Turnkey solutions; flexible cells
Japan	Aging workforce and quality enhancement	Society 5.0; human-centric automation	Collaborative robots (cobots); augmentation of skilled work

(Continued on next page)

(Continued from the previous page)

Economy	Core Driver	Policy Focus	SME Angle
Republic of Korea	SME competitiveness and manufacturing upgrading	Smart factories; data-driven and autonomous manufacturing support	Grants; training; SME-oriented toolkits

Note: SMEs = small and medium-sized enterprises.
Source: Compiled by the author based on State Council of the People's Republic of China, 2024; Ministry of Economic Affairs, n.d.; Cabinet Office, n.d.; and Ministry of SMEs and Startups, 2024.

For SME leaders, policy alignment means knowing which grants, tax incentives, and institute resources can be combined. For example, technical institutes may subsidize sensor kits and labeling, while export programs help package solutions for overseas customers once pilots succeed. The best way to capture value is to build a repeatable template in one cell, prove it in a second, and then export the know-how as a service from the cluster.

CONCLUSION

AI in manufacturing is no longer a distant promise. When paired with the right operating model, it is a set of practical tools that improve uptime, yield, and energy performance. The four plays in this report give SMEs a path to move from proofs of concept to reliable, scaled use. Start with a focused predictive maintenance or quality control pilot, integrate it into a small digital loop with a work-cell-level twin, design for human adoption, and collaborate securely with peers through federated learning when data is scarce.

Here are three closing recommendations:

1. Timebox pilots and publish weekly progress. Short cycles build trust.

2. Invest early in simple data standards (names, units, sampling). Standards reduce hidden costs later.

3. Treat people as the center of the system. Co-design interfaces and standard work with the shifts that will use them and reward near-miss reporting.

The factories that master these habits will not only capture the direct return on investment but also build a capability for continuous improvement that compounds over time.

Lastly, keep in mind that the experience is cumulative. Clean data, reusable tags, operator expertise, and templates are among the resources produced by each successful cell that will help the subsequent cell operate more quickly and affordably. In manufacturing, that is the true AI flywheel. The benefits extend beyond a single company to the entire ecosystem when clusters share those assets responsibly. This contributes to resilient, human-centered industrial growth and serves as a competitive strategy for Asian SMEs.

APPENDIX A

Worked Example: Return-on-investment Math You Can Run on a Napkin

Leaders frequently request a straightforward method for determining whether a pilot is worthwhile. The back-of-the-envelope method that follows is helpful.

1. Make a list of the primary cost drivers that the use case will impact. For predictive maintenance, these include collateral damage (scrap), emergency maintenance (rush parts, contractor rates), and unscheduled downtime (lost throughput, overtime). For quality control, they include line stoppages, returns, scrap and rework, and additional inspections.

2. For the first 90 days, estimate a realistic improvement range (not the best case). For predictive maintenance, assume 10%–30% fewer unscheduled downtime hours. For quality control, assume that 20%–50% fewer defects will pass undetected to later stages.

3. Make money from those deltas. For instance, a line generates 500 units per day at a margin of USD50. An hour of downtime results in the loss of 21 units, or USD1,050. If predictive maintenance reduces downtime by 40 hours per month, that would result in a saving of USD42,000 per month before emergency repair costs are deducted.

4. Deduct the following expenses: cloud/edge licenses (USD200–800/month), a gateway (USD300–900), sensors (USD80–300 each), and one person-month of engineering work for setup and handover. First pilots in many small and medium-sized enterprises have a payback period of two to four months and cost between USD8,000 and USD25,000.

APPENDIX B

Predictive Maintenance Modeling: From Thresholds to Trends and Hybrid Analytics

Not every pilot needs a deep learning model. Many wins come from hybrid analytics, combining physics-based thresholds with simple data-driven trend detectors. Begin with vendor manuals: If spindle vibration over X mm/s is abnormal, codify that as a Level 2 warning. Then add a trend rule: If vibration increases by 20% week over week for three weeks, create a Level 1 advisory.

As labeled events accumulate, experiment with anomaly detection (e.g., isolation forests) or forecasting models that predict when a variable will cross a threshold. Focus on interpretability: Show the variables that drove the alert so technicians know what to inspect.

Include maintenance context. If a part was replaced, reset or re-baseline the model. Record maintenance notes in a structured way so models do not misinterpret postrepair transients as faults. A three-field form (what changed, when, by whom) is sufficient for pilots.

APPENDIX C

Quality Control Playbook: Getting Image Capture and Labeling Right

Image quality determines model quality. Fix lighting first: Avoid mixed color temperatures and use diffused lighting to reduce glare on glossy parts. Lock camera exposure and white balance. Mount cameras rigidly; any wobble will cause blurs and lower accuracy.

Set up the capture area so that the field of view is constant in every picture. To stabilize placement, add conveyor gates or basic mechanical guides if parts are positioned differently. More often than not, a tiny adjustment increases accuracy more than any algorithmic change.

Create a small but clean dataset. For a first model, 300–500 images per major defect class and per normal class are usually sufficient. Label with simple, consistent rules. If unsure, create a “needs review” tag and resolve during weekly review. Consider class-imbalance handling (weighted losses) rather than over-collecting rare defects in early phases.

Start with a transfer-learned classifier or detector (i.e., a model adapted from a pretrained system rather than developed from scratch) and measure precision and recall, not just overall accuracy. For production, set a confidence threshold that favors recall when the cost of a missed defect is high and favors precision when false alarms are costly. Document the choice and revisit monthly.

Design the retraining loop. Every week, sample false positives and false negatives and add them, after human review, to the training set. Keep a frozen evaluation set (golden set) to track whether the model is actually improving rather than just memorizing new cases.

APPENDIX D

Digital Twins in Practice: Scheduling, Maintenance, and Continuous Improvement

Use the twin for three routines:

1. **Daily huddles:** Compare planned vs. actual output and decide small course corrections.

2. **Weekly scheduling:** Test alternative shift mixes or changeover sequences.

3. **Maintenance planning:** Simulate pulling forward a service window to reduce disruption.

Keep the twin honest. Once a week, run a reconciliation. Does the twin's throughput match actual throughput? If not, investigate data lags, missing downtime codes, or unmodeled constraints. A twin that drifts from reality will lose trust quickly.

Make improvement visible. When a bottleneck moves after a change, record it in a simple log (bottleneck history). Over months, this will become a learning asset for new supervisors and engineers.

APPENDIX E

Data Architecture Quick Start Guide for SMEs

To begin, create a unified tag dictionary with human-readable units, clear names (no cryptic codes), and sampling rates that correspond to process dynamics. Ad hoc notes are outperformed by a straightforward spreadsheet with four columns: tag name, unit, rate, and owner. Freeze it in week two and version-control it thereafter.

Establish time synchronization. If clocks drift, models drift. Use network time protocol on gateways and cameras. For vision systems, stamp images with both capture time and line position (e.g., part ID or station index).

Choose a lightweight storage approach. For time-series data, a compact historian (i.e., a lightweight system for storing time-based sensor and operational data) or a cloud time-series service is sufficient for pilot implementation. For images, store originals and downscaled thumbnails; keep a separate folder for labeled examples. Do not mix training and evaluation data in the same bucket.

Secure by default: Put gateways on a segregated network, require multifactor authentication for remote access, and log changes to configurations. A one-page “who can change what” policy avoids most surprises.

Finally, document the flow in one page: where data originates, how it is transformed, where it lands, and who uses it. Hang the diagram near the line so everyone can refer to it during stand-ups.

APPENDIX F

Change Management That Respects People and Accelerates Adoption

Adoption does not happen by memo. Use a simple cadence:

- Announce the pilot and its “why” at an all-hands meeting.
- Identify a small group of operators as co-designers.
- Run a hands-on dry run before the first shift.
- Conduct short daily check-ins during the first two weeks.
- Close with a retrospective that turns lessons into the next release plan.

Consider a light incentives program. Recognize the “AI champion” of the month, highlight teams with the most useful near-miss insights, and make improvements part of performance conversations without making them punitive.

Document two things: (1) the standard responses to each alert type, with photos or screenshots, and (2) a one-page troubleshooting guide. Paper copies near the line beat a buried intranet page.

APPENDIX G

Federated Learning: Privacy, Evaluation, and Model Drift

Privacy techniques to consider are secure aggregation (the server sees only the sum of updates), client-side differential privacy (noise added before sending), and strict access control for the coordinator. Simpler is often better for first rounds, so start with secure aggregation and strong operational controls.

Evaluation across sites is tricky. Maintain a shared, representative evaluation set and report the same metrics each round, for example, F1 score (precision-recall balance) for defect classification and mean absolute error for predictions of useful life remaining. Publish round-by-round charts for all participants.

If a site changes a supplier or material, its data distribution will shift. Encourage sites to monitor and log such changes and to re-baseline local models when needed. If a global model degrades for a subset of sites, consider clustering participants and training more than one federation.

APPENDIX H

What Weekly Progress Should Look Like (Sample 12-week Report)

Each Friday, publish a one-page note to the pilot team and leadership including the following:

- Key performance indicator snapshot (downtime, defect rate, energy per unit) vs. baseline and target
- Notable events (e.g., “false alarm on Station 2 fixed by camera calibration”)
- Operator feedback highlights (two to three quotes or notes)
- Risks or blockers (e.g., gateway instability, missing tags)
- Plan for next week (specific tasks with owners)

REFERENCES

Asian Productivity Organization. (4 March 2025). *Leveraging AI for smart manufacturing in SMEs* [Video]. YouTube. https://www.youtube.com/watch?v=51lXx4wBuHE

Cabinet Office. (n.d.). *Society 5.0*. Government of Japan. https://www8.cao.go.jp/cstp/english/society5_0/index.html

Detzner, A. & Eigner, M. (2018). A digital twin for root cause analysis and product quality monitoring. In D. Marjanovic, M. Storga, … N. Pavkovic (Eds.), *DS92: Proceedings of the DESIGN 2018 15th International Design Conference* (pp. 1547–1558). The Design Society. https://doi.org/10.21278/idc.2018.0418

Foxconn Technology Group. (21 January 2021). *Foxconn announces FOXCONN NxVAE, unsupervised learning AI technology*. https://www.foxconn.com/en-us/press-center/events/csr-events/533

Kibira, D., Shao, G. & Venketesh, R. (2023). Building a digital twin of an automated robot workcell. In *Annual Modeling and Simulation Conference 2023 (Proceedings)* (pp. 196–207). Society for Modeling and Simulation International. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=936285

Leng, J., Liu, Q., … Chen, X. (2020). Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. *Robotics and Computer-Integrated Manufacturing*, *63*, 101895. https://doi.org/10.1016/j.rcim.2019.101895

Leng, J., Zhang, H., … Zhang, D. (2019). Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop. *Journal of Ambient Intelligence and Humanized Computing*, *10*(3), 1155–1166. https://doi.org/10.1007/s12652-018-0881-5

Microsoft. (1 April 2025). *Rolls-Royce saves millions in cost avoidance with Microsoft Cloud for manufacturing*. https://www.microsoft.com/en/customers/story/23201-rolls-royce-azure-databricks

Ministry of Economic Affairs. (n.d.). *Industrial innovation and R&D: Promote the* “*Five Plus Two Industry Innovation Plan*”. Government of the Republic of China. https://www.moea.gov.tw/MNS/english/Policy/Policy.aspx?menu_id=32904&policy_id=13

Ministry of SMEs and Startups. (2024). *SMEs to soar new heights: MSS announced the new strategies for SMEs* [Press release]. Government of the Republic of Korea. https://mss.go.kr/site/eng/ex/bbs/View.do?bcIdx=1050186&cbIdx=244

Rolls-Royce. (n.d.). *Harnessing the power of AI to deliver more intelligent engine inspections*. https://www.rolls-royce.com/media/our-stories/discover/2021/intelligentengine-harnessing-the-power-of-ai-to-deliver-more-intelligent-engine-inspections.aspx

State Council of the People’s Republic of China. (2024). *Report on the work of the government*. https://english.www.gov.cn/news/202403/13/content_WS65f0dfccc6d0868f4e8e5079.html

Visual Components. (5 February 2016). *ITRI case study*. https://www.visualcomponents.com/case-studies/itri-case-study/

World Economic Forum. (2023). *Global Lighthouse Network: Adopting AI at speed and scale*. https://www.weforum.org/publications/manufacturing-lighthouses-and-the-path-to-impact-adopting-ai-at-speed-and-scale/

ABBREVIATIONS

AI	artificial intelligence
HMI	human-machine interface
KPI	key performance indicator
OPC UA	Open Platform Communications Unified Architecture
PROFIBUS	Process Field Bus
RTU	remote terminal unit
SMEs	small and medium-sized enterprises
TOPS	technology, organization, people, strategy
WEF	World Economic Forum

LIST OF TABLES

LIST OF FIGURES

ABOUT THE AUTHOR

Dr. Huiyi Fan
General Director of the Commerce Development Research Institute

Associate Professor of the Department of Business Administration at Chang Gung University
Republic of China

Watch the Productivity Talk on
Leveraging AI for Smart Manufacturing in SMEs