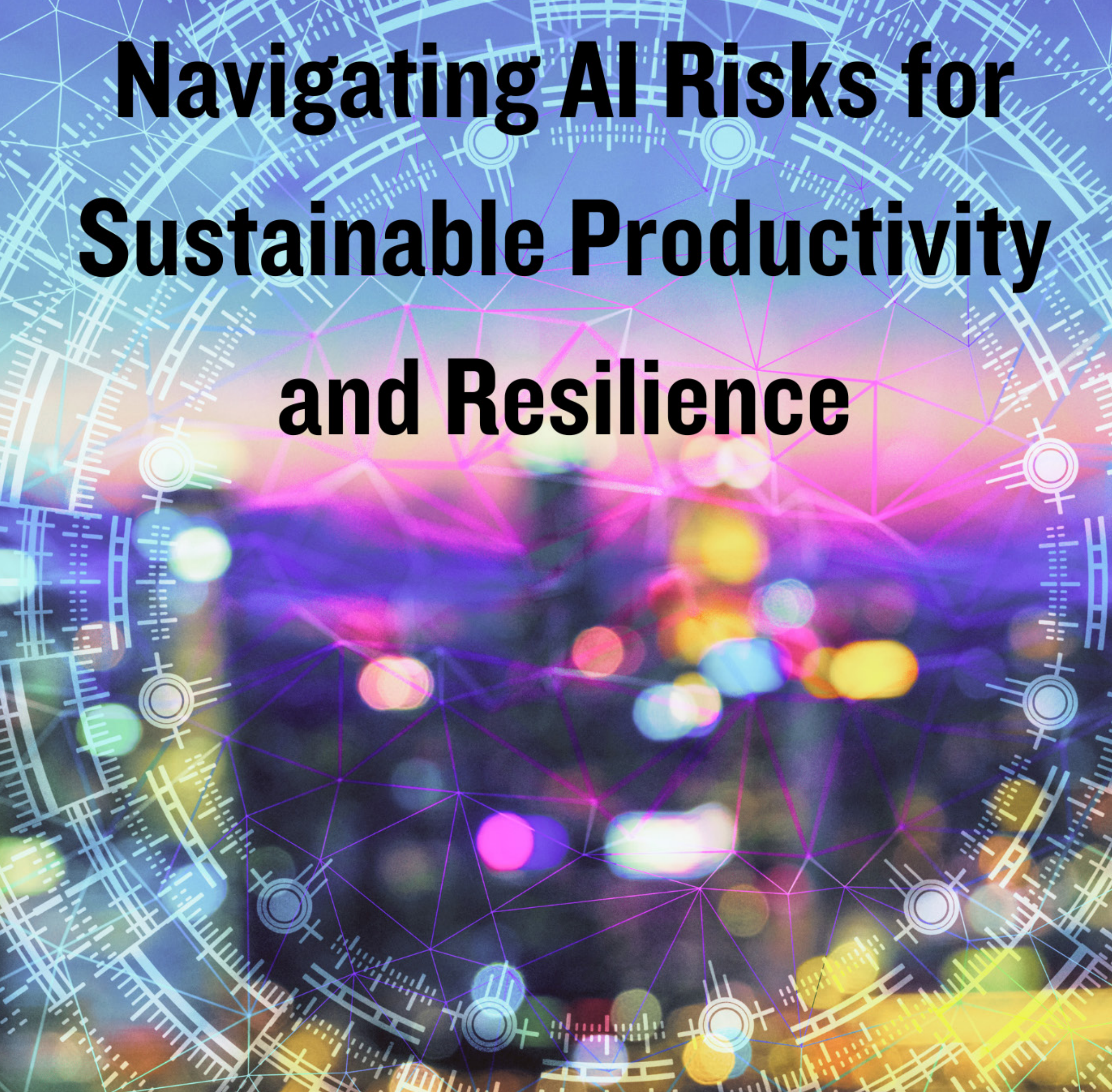
Navigating AI Risks for Sustainable Productivity and Resilience



Productivity *Insights*

Vol. 6–3

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, ROC, Fiji, Hong Kong, India, Indonesia, Islamic Republic of Iran, Japan, ROK, Lao PDR, Malaysia, Mongolia, Nepal, Pakistan, Philippines, Singapore, Sri Lanka, Thailand, Turkiye, and Vietnam.

Navigating AI Risks for Sustainable Productivity and Resilience

Dr. Dian Wirawan Tjondronegoro

PRODUCTIVITY INSIGHTS Vol. 6-3
NAVIGATING AI RISKS FOR SUSTAINABLE PRODUCTIVITY AND RESILIENCE

Dr. Dian Wirawan Tjondronegoro 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 Interbooks Co., Ltd.

CONTENTS

PREFACE

The Productivity Insights (P-Insights) series is an extension of the Productivity Talk (P-Talk) series, which is a flagship program under the APO Secretariat's digital information initiative. 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.

EXECUTIVE SUMMARY

In today's rapidly evolving landscape, artificial intelligence (AI) stands poised to transform the way organizations deliver value, manage risk, and foster resilience. However, while pilots and proofs of concept frequently generate excitement, many AI initiatives falter when exposed to shifting data, emerging regulations, or evolving stakeholder expectations. The gap between early promise and sustained impact often comes down to one simple reality: AI is not a point solution but rather a socio-technical capability that must be nurtured, governed, and continually refined.

This edition of P-Insights offers a concise, actionable blueprint that translates into three interlocking frameworks. **EPIC** ensures that your organization is ready for AI by building skills, forging partnerships, stabilizing data pipelines, and engaging communities. **TOAST** embeds governable, testable principles into every model, establishes a lean but powerful oversight board, wires continuous monitoring, and secures a public licence to operate. **RAIIF** closes the loop by watching for drift, making improvements via disciplined retraining and auditing, experimenting safely in governed sandboxes, and motivating every team member to surface early warnings.

Each framework aligns with global benchmarks, from the OECD Principles and NIST AI Risk Management Framework to ISO/IEC 42001 and the forthcoming EU AI Act, ensuring both compliance and competitive advantage. We have distilled our lessons into actionable and practical steps while illustrating them with real-world examples from health care, education, and government.

Whether you lead a clinical AI deployment or a public-sector service transformation, this report will help you turn abstract principles into concrete steps. Our aim is to spark "lightbulb" moments that you can put into practice immediately, advancing your next AI journey from an intriguing experiment to a trusted, high-impact workhorse.

INTRODUCTION

Artificial intelligence now threads through every major sector, from hospitals and classrooms to council chambers. It schedules surgical teams, marks essays at scale, balances city budgets, predicts flood peaks, and even helps draft parliamentary bills. When applied well, AI can be a powerful productivity lever: radiology backlogs shrink, lesson-planning time is cut in half, and citizen queries receive round-the-clock answers. Recent economic forecasts suggest that such gains could unlock trillions of dollars in new value over the next decade.

Yet every success story carries a cautionary echo. A machine learning model can drift when real-world data shifts, an AI algorithm trained on yesterday's patterns can suffer from hidden bias, and a "friendly" chatbot can become confused due to opaque logic. Left unmanaged, those slipups will erode trust, attract lawsuits, and wipe out the very efficiency gains leaders hoped for. Recent systematic reviews confirm the pattern: technical performance alone does not determine AI impact, as organizational culture, governance, and public acceptance are equally decisive (Dwivedi et al., 2021; Glikson & Woolley, 2020).

Why a Socio-Technical Lens Matters

The core problem is not that AI is mysterious code. The problem is that AI sits squarely at the intersection of technology, people, processes, and policy. Back in 2018, Taddeo and Floridi argued that AI must be treated as “a force for good” only if concrete ethical guardrails accompany the algorithms (Taddeo & Floridi, 2018). They pointed out that learning systems already shape daily practices, interactions, and the environment, so they must be steered, not merely built.

Regulators and standards bodies have since turned that call into action. The OECD Recommendation on AI (2019) lays down five value-based and five policy principles for “innovative, trustworthy AI that respects human rights and democratic values.” The U.S. National Institute of Standards and Technology followed with AI Risk Management Framework 1.0, a voluntary blueprint for embedding reliability, safety, and bias control into every stage from design to retirement (NIST, 2023). In 2023, the International Organization for Standardization released ISO/IEC 42001, the world’s first auditable management-system standard for AI, giving organizations a practical Plan-Do-Check-Act cycle across the lifespan of any AI model. And, in 2024, the Council of the European Union reached political agreement on the AI Act, the most comprehensive risk-tiered regulation to date, covering everything from high-risk medical devices to unacceptable-risk social scoring.

Taken together, these documents converge on a single insight: AI risk is a socio-technical challenge. To capture the upside, we must align algorithms with people, culture, workflows, and regulations, rather than just bolt them on and hope.

Three Frameworks, One Practical Blueprint

During the past decade, we have applied that socio-technical lens to health, education, and digital-government programs. The result is three complementary frameworks: EPIC, TOAST, and RAIFF.

TABLE 1

THREE COMPLEMENTARY FRAMEWORKS FOR A PRACTICAL BLUEPRINT

Framework	Core purpose	Key focus
EPIC (Education, Partnership, Infrastructure, Community)	Build organisational readiness	Skills & literacy, multi-stakeholder partnerships, robust data/ IT backbone, transparent engagement
TOAST (Trustworthy, Optimized, Adaptable, Socio-Technological)	Guide ethical, efficient integration	Trust-building, performance optimization, adaptability, socio-technical harmony
RAIIF (Responsible AI Implementation Framework)	Run an iterative co-creation cycle	Human-centred design, continual feedback and staged scaling with embedded ethics

These frameworks can turn AI risk into a competitive edge when organizations follow a disciplined loop:

1 Prepare → 2 Govern → 3 Adapt → repeat.

Preparation equips people and infrastructure; governance keeps systems safe, fair, and compliant; and adaptation ensures that models stay sharp as reality shifts. Each turn strengthens the next, forming a self-reinforcing engine of productivity and resilience.

This edition of P-Insights will translate this loop into three concrete, inter-dependent actions you can start taking this quarter:

1. **Prepare your organisation for AI**: Build skills, partnerships, data quality, and an open culture before the first line of code ships.
2. **Implement responsible AI governance**: Embed clear principles, cross-functional oversight, and transparent audit trails so every model remains aligned with your mission, laws, and public expectations.
3. **Continuously adapt and innovate**: Monitor, retrain, and improve AI (and the rules around it) in real time, turning learning into a lasting advantage.

Each section will discuss each action, blending frontline lessons with examples and checklists. Concise tables summarize readiness checkpoints, governance components, and adaptation tactics, while a single diagram shows how the three actions lock into a virtuous cycle. The objective is to provide a plain-language blueprint that is rooted in EPIC, TOAST, and RAIIF while also being tightly aligned with global benchmarks from the OECD, NIST, and ISO as well as the forthcoming EU AI Act for converting AI risk into sustainable productivity and resilience. Grab the parts that light a bulb, apply them immediately, and hopefully your next AI journey will advance from an intriguing pilot to a trusted workhorse.

PREPARE YOUR ORGANIZATION FOR AI

Many people who hear “artificial intelligence” still picture science-fiction robots. In reality, every AI success story we’ve examined begins with very ordinary tasks: training people, cleaning up data, finding expert partners, and explaining the project openly. If you skip that groundwork, even the smartest model falls over. However, if you do it properly, the same model slips straight into the workflow and starts making an impact from day one.

This first action therefore answers a single question:
How do we build an AI-ready organization before the first model ships?

We unpack the answer through the four **EPIC** pillars: **Education**, **Partnership**, **Infrastructure**, and **Community** (Tjondronegoro, 2024).

1.1 Education: Make AI Everyone’s Business

In the EPIC playbook, Education comes first because literacy is the seed from which every later safeguard grows. Education here does not mean teaching every nurse or records officer Python. It means making sure the people in each role understand three things: what the technology can do, where it fits into their own workflow, and why ethics matter. When those basics land, mistrust subsides, front-line staff spot dozens of micro-tasks ripe for automation, and governance rules start to feel like common sense rather than red tape (Dwivedi et al., 2021; Jobin et al., 2019).

A proven accelerator is a short, high-energy “AI-in-Action” huddle: a 15-minute live demo (speech-to-text summarizing a meeting works well), a myth-busting Q&A, and a breakout session where every participant names one pain point AI might solve. At a regional Australian hospital, these sessions resulted in 27 ideas in four weeks. Five of these became pilots, and one is now saving radiologists ninety minutes a day triaging scans, an impact line with the translational gains reported by Sendak et al. (2020). Each huddle ends with a plain-language flyer setting out the organization's AI principles, so literacy and values travel together.

1.2 Partnership: Borrow the Expertise You Don't Have

The *Partnership* pillar recognises that even Fortune-500 firms lack all the talent required for state-of-the-art AI. Cross-sector alliances, including universities, specialist start-ups, cloud vendors, and NGOs, collapse learning curves and cut time-to-value (Fountaine et al., 2019).

A lightweight three-clause Memorandum of Understanding is often enough: the organization provides a clearly framed problem and an anonymized dataset; the partner brings postgraduate talent, mentoring, and, if needed, subsidized computing; and both parties publish the findings. This is precisely how Duke Health translated its sepsis-prediction model from lab to ward: clinicians contributed workflow nuance, researchers tuned the algorithm, and together they launched a tool that flags high-risk patients hours earlier than if conducting a manual review (Sendak et al., 2020).

Public agencies short of cash can swap "data-for-research" or join multi-agency consortia that pool budgets and open-source the resulting code, a move squarely in line with the collaboration principle of the OECD AI Recommendation (2019).

1.3 Infrastructure: Lay Data Rails You Can Trust

EPIC's *Infrastructure* pillar insists that AI built on patchy data is a skyscraper on sand. Before coding begins, teams conduct a brutally honest review: *Do we have the data? Are they complete? Where might bias lurk? Are privacy controls ISO-compliant? Can the platform scale once the model is live?* (Adadi & Berrada, 2018; ISO/IEC 42001, 2023).

A fast starter is the two-day *red-amber-green data audit*. Duke Health's sepsis team did exactly this. Within 48 hours they saw that 11% of lactate results lacked time stamps (red: completeness) and Black patients were under-represented in the training cohort (red: bias). Restoring missing logs and oversampling the minority group increased accuracy, and, more importantly, prevented an equity problem before the model reached the ward.

1.4 Community: Earn Trust Before the Launch Banner Goes Up

Finally, the *Community* pillar reminds us that AI systems impact citizens, patients, and students and are judged in the court of public opinion. Early, open dialogue beats costly damage control every time (OECD, 2019). A simple template works: publish a plain-language FAQ as soon as the pilot starts by stating the problem, the data, the safeguards, and a contact email address, and then host one public webinar.

Ultimately, when every EPIC pillar is achieved, it means that staff literacy has been achieved, partners are engaged, the data are sound, and the community is on board. Then, we can move from “AI sounds interesting” to “We are structurally ready.”

TABLE 2

EPIC: FOUR PILLARS OF AI-READINESS AT A GLANCE

Pillar	Core Objective	90-Day Quick Win	Concrete Reference Example
Education	Increase AI literacy so people in every role understand capabilities, limits, and ethics.	Run "AI-in-Action" huddles: live demo + Q&A + idea harvest. Issue a plain-language AI-principles flyer.	Regional hospital huddles surfaced 27 automation ideas. One pilot now saves radiologists 90 min/day (Sendak et al., 2020).
Partnership	Plug skill gaps fast via cross-sector alliances.	Sign a three-clause MoU: problem + anonymized data ↔ postgraduate talent + cloud credits → co-publish results.	Duke Health collaborated with a university data-science institute to create a sepsis predictor that flags high-risk patients hours earlier (Sendak et al., 2020).
Infrastructure	Ensure that data, tech, and privacy rails are production-grade.	Two-day red-amber-green audit: catalogue sets, score completeness/bias/ compliance, and time-box every "red."	An audit revealed that 11% of lactate timestamps were missing. The fix increased accuracy and prevented bias (ISO/IEC 42001, 2023).
Community	Earn social licence through transparency and dialogue.	Publish a FAQ as the pilot starts. Host one public webinar. Open a feedback mailbox.	A welfare agency webinar caught wording that disadvantaged visually-impaired users. The fix pre-empted a PR crisis (OECD, 2019).

IMPLEMENT RESPONSIBLE AI GOVERNANCE

After preparation lays the railway, governance is the signaling system that enables the train to reach full speed without derailing. In the original **TOAST** paper, governance is described as the "ethical and technical spine" that binds **T**rust & Accountability, **O**ptimization & Control, **A**daptability & Innovation, and **S**ocio-**T**echnical Harmony into one operational loop (Tjondronegoro, 2025).

Here we will translate each pillar into daily practice, drawing on real cases and external standards.

2.1 Trust & Accountability: Values that Trigger Code

Both ISO/IEC 42001 (2023) and the NIST AI-RMF (2023) emphasize that principles must be explicit and testable. TOAST's Trust pillar builds on this by encouraging organizations to write the metric, threshold, and automatic action in a single sentence, thereby making safeguards codable, auditable, and understandable.

Example:
A regional bank operationalized fairness by anchoring its rules to the OECD AI Recommendation (2019): "Suspend any credit-scoring model if the approval gap between the highest- and lowest-scoring demographic groups exceeds 3 percentage points; resume only after a human risk officer signs off." This led to zero bias incidents and 70% faster model releases.

Case Study: Singapore's AI Governance Framework
Singapore's Model Register requires every government AI system to publish its purpose, metrics, alert thresholds, and automatic mitigation steps in a single registry entry (Allen et al., 2025). This ensures that values like fairness and privacy are not aspirational but instead tied to concrete KPIs and actions, aligning directly with TOAST's trust-through-code principle.

How to start immediately:
Draft a two-page charter linking each corporate value to one measurable indicator and one automatic response. Post it internally to invite feedback, because transparency begins inside the firewall.

2.2 Optimization & Control: a Lean Yet Empowered Oversight Committee

Even the best rules need stewards. TOAST's Optimization pillar calls for a five-seat oversight board with real authority and minimal bureaucracy. Its members are:

- **Model owner** (a lead data scientist): explains mechanics and limitations
- **Domain expert** (a clinician, teacher, and claims assessor): probes real-world realism
- **Privacy officer**: checks to ensure lawful data use
- **Cyber-risk chief**: reviews security controls
- **User or citizen advocate**: weighs lived-experience bias

Example:
At a Melbourne hospital, this format was used to review an AI fracture detector. Audit logs showed that recall lagged for darker-skinned patients. The board vetoed deployment, the training set was expanded, and the revised model launched with 35% faster reporting and clinician trust secured (Bennett et al., 2022).

Case Study: Thailand's Cooperative ITS
Bangkok's transport authority formed a five-stakeholder "ITS Coordination Committee" (data scientists, traffic engineers, legal advisors, privacy officers, and citizen reps) to review AI features every two weeks by using a one-page ISO-aligned agenda (Choosakun et al., 2021). The committee had veto power where legal or user needs conflicted with technical plans, which embodies TOAST's lean but decisive governance.

How to start immediately:
Form the board, schedule a 60-minute meeting every two weeks, and adopt a standard agenda: purpose → data → metrics → risk → next actions. Log decisions in a public changelog.

2.3 Adaptability & Innovation: Monitoring That Proves Instead of Just Observing

Even ISO-certified AI models can drift. TOAST treats every production model as adaptive, requiring live monitoring of performance and fairness, with retraining routed through governance checkpoints.

Example:
When Australia entered its first COVID-19 lockdown, a national supermarket chain saw its demand-forecast dashboard flash red within hours. Accuracy plunged below the charter threshold, automatically pausing the model and opening a ticket for the AI oversight board. Engineers quickly built a "shadow" version that added new mobility-curfew signals, then ran it alongside the old model. This illustrative example embodies the audit-loop method described by Raji et al. (2020).

Case Study: Hangzhou City Brain
City Brain streams real-time traffic, congestion, and incident metrics from thousands of sensors to a live dashboard. A cross-department review team meets weekly to inspect KPI breaches and schedule updates (Caprotti & Liu, 2022). This continuous telemetry and board backlog process prevents one-off audits from becoming obsolete, which is exactly the kind of adaptive loop TOAST prescribes.

How to start immediately:
Instrument one live model with drift, fairness, and latency alerts. Route alerts to the board's chat and commit to triage within 24 hours.

2.4 Socio-Technical Harmony: Transparency That Earns Permission to Operate

Governance fails if it protects code but alienates people. TOAST's ST pillar emphasizes procedural justice, which emphasizes that citizens must see how decisions are made, not just that they're accurate.

Example:
Before launching AI-based benefits triage, Canada's welfare ministry completed a federal Algorithmic Impact Assessment. A high-risk score triggered civil-society workshops and a promise to publish quarterly bias audits. The result:

40% fewer complaints, halved processing time, and reduced political opposition (Treasury Board Secretariat, 2019).

Case Study: Explainable AI in Japan
Japan's Ministry of Finance publishes case-based and counterfactual explanations of every tax-inspection AI decision (Aoki et al., 2024). They also host quarterly public forums to walk through anonymized Model Cards. Surveys show a 25% rise in perceived fairness and 30% drop in complaints, which is proof that transparency builds trust.

How to enact this:
Write a FAQ and model card for one system (Mitchell et al., 2019). Publish both online. Invite an NGO or citizen rep to the next board meeting as an observer.

2.5 Putting the TOAST Engine on the Rails

To activate all four pillars:

- Post the two-page AI charter beside the coffee machine.
- Convene the five-seat board and give it authority over one active model.
- Wire an AI model for three live metrics and alert thresholds.
- Mark the first 60-minute review on the calendar, and publicly log the outcome.
- Post the FAQ and model card for staff (as well as citizens if public-facing).

By doing this, Trust becomes enforceable, Oversight becomes lean, Adaptability becomes proactive, and Socio-technical harmony earns public permission.

TABLE 3

COMMON FAILURE MODES AND HOW TOAST NEUTRALIZES THEM

Failure pattern	TOAST pillar	Built-in remedy as exemplified by case studies
Principles theatre: Lofty values, no measurable guard-rails	**Trust & Accountability**	**Singapore's AI Governance Framework** (Allen et al., 2025): Singapore's "Model Register" requires every government AI system to publish its purpose, metrics, alert thresholds, and automatic mitigation steps in a single registry entry. This ensures that "fairness," "accuracy," and "privacy" are not just aspirational but tied to concrete KPIs and actions, which is exactly the charter approach TOAST prescribes.
One-shot audit risk: Model passes once then drifts	**Adaptability & Innovation**	**Hangzhou City Brain** (Caprotti & Liu, 2022): City Brain streams real-time traffic flow, congestion, and incident metrics from thousands of sensors to a live operations dashboard. A dedicated cross-department review team meets weekly to inspect any KPI breaches and schedule data-pipeline or model updates. This continuous telemetry and board backlog process prevents one-off audits from becoming obsolete.
Siloed ownership: Tech, legal and users never meet	**Optimisation & Control**	**Thailand Cooperative ITS** (Choosakun et al., 2021): Bangkok's transport authority established a five-stakeholder "ITS Coordination Committee" (data scientists, traffic engineers, legal advisors, privacy officers, and citizen reps). Using a one-page ISO-aligned agenda, they review new AI features every two weeks, exercising their veto power where legal or user needs conflict with technical plans.
Public backlash: Opaque "black box" systems	**Socio-Technical Harmony**	**Explainable AI in Japan** (Aoki et al., 2024): The Japanese Ministry of Finance publishes "case-based" and "counterfactual" explanations alongside every tax-inspection AI decision. They also host quarterly public forums to walk through anonymized Model Cards. Citizen surveys have shown a 25% rise in perceived fairness and a 30% drop in complaints compared to the previous year.

CONTINUOUSLY ADAPT AND INNOVATE

Many AI development teams breathe a sigh of relief once their AI models and algorithms pass validation and governance requirements, but the real test starts the moment such models meet the real, dynamically changing world. Supply chains judder, privacy clauses are amended, or new foundation models appear on GitHub. Unless the algorithms learn as quickly as their environments change, yesterday's successes quietly become tomorrow's risks.

The **Responsible AI Implementation Framework (RAIIF)** treats that reality as non-negotiable. It wraps every production system in a "learn-while-running" loop so the solution that ships today is still trustworthy, fair, and useful six months from now (Tjondronegoro et al., 2022). RAIIF keeps the loop turning through four habits: watch, improve, experiment, and motivate, each of which is mapped to the external standards that now shape the AI landscape. This section will show how to embed those habits without slowing delivery, so your AI solutions stay valuable even as everything else moves.

Learn in production or be left behind: RAIIF keeps the loop spinning through four habits, each anchored to an external benchmark.

- *Watch*: Live telemetry shows drift before users feel pain (NIST AI-RMF, 2023).
- *Improve*: Retrain → re-audit → shadow-test → log (Schwartz et al., 2023).
- *Experiment*: Trial new ideas in a governed sandbox (Dwivedi et al., 2021; Mitchell et al., 2019).
- *Motivate*: Celebrate every bug-finder (Glikson & Woolley, 2020).

Below we unpack each habit with a single field vignette and a "Monday-morning move" you may adopt.

3.1 Watch: See Drift Before Anyone Is Hurt

What to do: Connect AI solutions or data analysis pipeline's accuracy, latency and subgroup-parity to a shared dashboard. Trigger rule: if accuracy falls below a critical threshold, the system throttles itself and posts an alert to the oversight channel.

Evidence: Duke Health's sepsis predictor streams AUROC and feature coverage to a "model-health" screen (Sendak et al., 2020). Eight hours after an EHR upgrade, AUROC slipped below the charter threshold; a one-line mapping fix restored performance the same day—drift caught as a red pixel, not in missed patients.

Monday move: Turn on drift and parity alerts for one production model; direct notifications to the governance board's chat.

3.2 Improve: Retrain, Re-audit, Redeploy

What to do:

1. Retrain weekly or on alert.
2. Re-audit with the same scripts that gated launch (fairness, privacy, robustness).
3. Shadow-test the challenger vs the incumbent; promote only if the newcomer wins every KPI.
4. Log hash, data cut-off, approver, rationale.

Evidence: After Amsterdam opened a new tram line, its congestion model under-predicted bus crowding. Three days of fresh GPS feeds and a rerun of the bias suite increased accuracy seven points while preserving parity across income areas. This vignette demonstrates iteration with memory, not mayhem (a process that aligns with Raji et al. (2020)).

Monday move. Book a 30-min monthly "model-health" huddle to ensure that ops, risk, and domain lead review retrain logs and shadow tests.

3.3 Experiment: Sandbox Before Prime Time

What to do: Mask production data, isolate credentials, and require board sign-off (with a Model Card) before any pilot touches users.

Evidence: A Canadian tax office trialed a GPT-based FAQ bot on 100,000 masked queries. Ten-thousand red-team prompts produced zero hallucinations, and the board cleared a limited beta with a human clerk "on the loop." This cadence mirrors Dwivedi et al.'s (2021) call for cautious curiosity.

Monday move: Publish a two-page sandbox protocol next to the charter, and refuse any new model that bypasses it.

3.4 Motivate: Reward Every Whistle-Blower

What to do: Treat each anomaly report as design gold, and praise the finder publicly.

Evidence: A teacher flagged that an essay-grader penalized certain idioms. Refreshing the corpus cut false positives by 30%. This illustrative example for shout-out in the staff bulletin reinforced the voice-and-trust dynamic described by Glikson and Woolley (2020).

Monday move. Announce a "bug bounty" for internal staff: coffee-cards and public kudos for the next confirmed issue.

Avoid the usual development traps:

- **Set-and-forget models.** The combination of live telemetry and charter thresholds stops silent decay.
- **Retraining without re-auditing.** CI/CD pipelines embed fairness and privacy tests.
- **Shadow-IT experiments.** A sandbox policy and board sign-off keep every project on the radar.

Your first loop in one sprint:

1. Turn on drift and parity alerts for one model.
2. Hold the first 30-min model-health huddle.
3. Enforce the sandbox approach on any new idea.
4. Celebrate the next anomaly reporter.
5. Feed lessons into the next EPIC training cycle, thereby closing the loop back to **Prepare**.

With watching, improving, and safe experimenting in place, your AI doesn't just keep up. It compounds value while protecting the trust earned under TOAST governance. The next section shows how the **Prepare**, **Govern**, and **Adapt** steps lock together into a single engine of resilience.

INTERCONNECT STEPS TO ACHIEVE A POSITIVE, SELF-REINFORCING CYCLE

Viewed in isolation, the three actions might look like separate work-streams: **EPIC** increases readiness, **TOAST** secures governance, and **RAIIF** keeps AI models and algorithms robust. Their real power appears only when they lock together, turning risk management into a productivity flywheel (as shown in Figure 1).

4.1 Preparation Feeds Governance

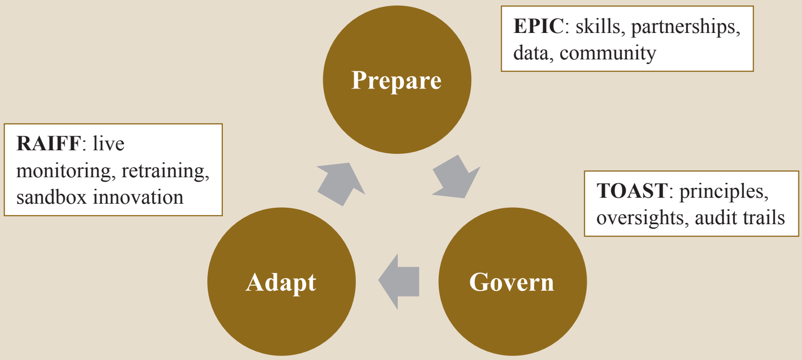
When an organization applies the EPIC lens to raise literacy, forge partnerships, tidy data, and open dialogue, its people can debate AI rules based on concrete facts rather than hunches. Clean datasets make bias tests easy to run, and a privacy-aware culture makes model cards routine, not heroic (Mitchell et al., 2019). Stakeholders invited early now see their concerns reflected in the charter. The result is a TOAST-styled oversight board equipped with accurate evidence, organizational legitimacy, and a shared vocabulary that defines prerequisites for swift, decisive risk control.

4.2 Governance Powers Adaptation

Because the board has published thresholds and real veto rights, data scientists know exactly when to retrain, what evidence to supply, and how to obtain approval. Live dashboards feed those metrics straight back to the board, satisfying its Trust & Accountability mandate without endless meetings (Raji et al., 2020). That clarity unshackles RAIIF's iteration loop: teams patch drift or bias in days, confident that the board will green light any model that clears the published bar. Safety becomes a catalyst for speed, not a drag.

FIGURE 1

The AI Risk–to–Resilience Fly-Wheel



4.3 Adaptation Renews Preparation

Every drift alert, audit finding, or sandbox experiment yields lessons, such as a feature the data pipeline must capture, a new skill the workforce needs, or a public concern that should appear in the next FAQ. Those insights roll naturally into the next EPIC cycle, with updated training modules, revised data contracts, and refreshed community briefings, all of which raise organizational readiness one notch higher. Round after round the enterprise climbs the AI-maturity curve described by Dwivedi et al. (2021): early pilots with manual checks, then automated audits at scale, and, ultimately, an AI-native culture where innovation and assurance travel in the same heartbeat..

4.4 Compound advantage

The strategy is constant (**Prepare → Govern → Adapt → repeat**) but the implementation expands by domain. A hospital uses the loop to govern diagnostic imaging, a university applies it to adaptive learning, and a city applies it to service chatbots. Like compound interest, each rotation grows the principal (skills, trust, and performance) while lowering the marginal cost of the next improvement. Organizations that internalize this rhythm become secure in the knowledge that every turn converts uncertainty into sustainable productivity and resilience. They stop asking, "*Can we handle another AI project?*" and start asking, "*How fast can we spin the flywheel again?*"

CONCLUSION: FROM SINGLE WINS TO ENDURING ADVANTAGE

Artificial intelligence dazzles most vividly in its early triumphs: a diagnostic model flags tumors invisible to the human eye, a chatbot clears a city-hall call queue, or an optimizer gifts lecturers whole afternoons. Yet first impressions rarely equal lasting value. Many pilots stall once data drift, staff rotate, or public opinion cools. The organizations that go on harvesting productivity, year after year, are those that treat AI not as a one-off project but as an evolving socio-technical capability (Dwivedi et al., 2021).

This capability rests on three mutually reinforcing moves. **EPIC** establishes the starting platform by educating people, forging partnerships, tightening infrastructure, and engaging the community. **TOAST** keeps that platform safe and swift by turning values into testable rules, giving lean oversight real authority, wiring models to live telemetry, and opening the black box to citizens (ISO, 2023; OECD, 2019). **RAIIF** lets everything breathe, observe, and learn: drift is spotted quickly, fixes are audited at speed (Schwartz et al., 2023), and new ideas mature in a sandbox before they face the world.

Played once, the sequence yields a respectable pilot; repeated, it produces a flywheel. Literacy drives better rules; better rules enable quicker iteration; iteration uncovers insights that feed the next round of learning. Like compound interest, every rotation widens the gap between enterprises that manage AI deliberately and those that leave success to luck.

Five high-leverage moves for the next 90 days:

1. **Publish an enforceable AI charter.** Translate “fair, transparent, and accountable” into numbers, thresholds, and auto-actions. Tape it beside the coffee machine and embed it in every vendor contract.
2. **Seat a five-person oversight board.** One data scientist, one domain lead, one privacy lawyer, one cyber-risk chief, and one user or citizen representative. Grant real veto power over any model that touches people or money.
3. **Switch on drift telemetry for one production model.** Set alert thresholds straight from the charter. Deliver the first “model health” email within 30 days.
4. **Shadow-retrain once.** Use fresh data to build a challenger model, run it in parallel, and promote it only if it wins fairly. Log every step, and hand auditors a watertight story.
5. **Host a community or staff showcase.** Explain what the AI does and how appeals work, invite sharp questions, and publish the Q&A. Few actions build trust faster than respectful transparency.

Complete those steps and you will have turned abstract frameworks into living muscle, with EPIC’s readiness visible on your dashboard, TOAST’s trust rules executing in code, and RAIIF’s feedback loop already paying for itself with a cleaner, smarter model.

Why endurance matters

Rules are tightening. Within a year, the EU AI Act will become law, U.S. Executive Order 14110 is already shaping federal contracts, and Australia is drafting its own AI risk-tier scheme. At the same time, the tech frontier keeps sprinting: multimodal language models this quarter, privacy-preserving federated learning the next. Organizations that have rehearsed the **Prepare → Govern → Adapt** loop will be resilient enough to meet each change with curiosity rather than panic. After all, they have:

- **Literacy** to size up new tools,
- **Governance** to launch them safely, and
- **Agility** to refresh data, skills, and controls before risk turns into re-work.

Endurance also pays a reputational dividend. Staff trust that automation augments instead of eliminating their roles (Tambe et al., 2019), customers see concrete safeguards, and regulators view the firm as a dependable innovator. Such social licence can unlock more fresh data partnerships, attract talent, and smooth regulatory pathways.

Final remarks and invitation

People sometimes call risk management a brake pedal. In practice it is the steering wheel, dashboard, and suspension that let you drive fast and stay on the road. EPIC, TOAST, and RAIIF are not ivory-tower theory. Instead, they are distilled field notes from research projects on AI at scale. Take the first loop this quarter, the second next quarter, and by year's end you may look back on a portfolio of AI models and solutions that are clever and trusted, efficient, and resilient.

REFERENCES

Adadi, A. & Berrada, M. (2018). Peeking inside the black box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Allen, J. G., Loo, J. & Campoverde, J. L. L. (2025). Governing intelligence: Singapore's evolving AI governance framework. *Cambridge Forum on AI: Law and Governance, 1*, e12. https://doi.org/10.1017/cfl.2024.12

Aoki, N., Tatsumi, T., Naruse, G. & Maeda, K. (2024). Explainable AI for government: Does the type of explanation matter to the accuracy, fairness, and trustworthiness of an algorithmic decision as perceived by those who are affected? *Government Information Quarterly, 41*(4), 101965. https://doi.org/10.1016/j.giq.2024.101965

Barocas, S., Hardt, M. & Narayanan, A. (2019). *Fairness and machine learning: Limitations and opportunities*. http://www.fairmlbook.org

Bennett, A., Hakim, A., Gomez, A., Smith, J. & Zhao, Y. (2022, November). *Mitigating skin-tone bias in automated fracture detection: A prospective clinical audit* [Conference presentation]. Radiological Society of North America 108th Annual Meeting, Chicago, IL, USA.

Caprotti, F., Liu, D. (2022). Platform urbanism and the Chinese smart city: the co-production and territorialisation of Hangzhou City Brain. *GeoJournal* 87, 1559–1573. https://doi.org/10.1007/s10708-020-10320-2

Choosakun, A., Chaiittipornwong, Y. & Yeom, C. (2021). Development of the Cooperative Intelligent Transport System in Thailand: A Prospective Approach. Infrastructures, 6(3), 36. https://doi.org/10.3390/infrastructures6030036

Davenport, T. H. & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review, 96*(1), 108–116.

Dwivedi, Y. K., Hughes, D. L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., … Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57,* 102141. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

European Union. (2024). *Artificial Intelligence Act: Proposal for a regulation laying down harmonised rules on artificial intelligence* (Council Doc. ST 16731/23). Council of the European Union. https://data.consilium.europa.eu/doc/document/ST-16731-2023-INIT/en/pdf

Fountaine, T., McCarthy, B. & Saleh, T. (2019). Building the AI-powered organization. *Harvard Business Review, 97*(4), 62–73.

Glikson, E. & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

IEEE Standards Association. (2022). *IEEE 7000-2021: IEEE standard—Model process for addressing ethical concerns during system design.* https://doi.org/10.1109/IEEESTD.2022.9703419

International Organization for Standardization. (2023). *ISO/IEC 42001:2023—Artificial intelligence management system—Requirements*. ISO.

Jobin, A., Ienca, M. & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., … Matei, A. (2019). Model cards for model reporting. In *Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency* (pp. 220–229). Association for Computing Machinery. https://doi.org/10.1145/3287560.3287596

National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework 1.0* (NIST AI 100-1). https://doi.org/10.6028/NIST.AI.100-1

Organisation for Economic Co-operation and Development. (2019). *OECD recommendation of the Council on artificial intelligence* (OECD/LEGAL/0449). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Raji, I. D., Smart, A., White, R., Mitchell, M., Gebru, T., Hutchinson, B., … Barnes, P. (2020). Closing the AI accountability gap: An end-to-end framework for internal algorithmic auditing. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 33–44). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372873

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A. & Hall, P. (2023). *Towards a standard for identifying and managing bias in artificial intelligence* (NIST Special Publication 1270). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.1270

Sendak, M. P., D'Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., … Balu, S. (2020). A path for translation of machine-learning products into healthcare delivery. *NPJ Digital Medicine, 3*, Article 106. https://doi.org/10.1038/s41746-020-00314-9

Taddeo, M. & Floridi, L. (2018). How AI can be a force for good. *Science, 361*(6404), 751–752. https://doi.org/10.1126/science.aat5991

Tambe, P., Cappelli, P. & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review, 61*(4), 15–42. https://doi.org/10.1177/0008125619867910

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25,* 44–56. https://doi.org/10.1038/s41591-018-0300-7

Treasury Board of Canada Secretariat. (2019). *Directive on automated decision making: Algorithmic impact assessment* (Version 1.2). Government of Canada. https://open.canada.ca/data/en/dataset/3c6b1fbe-0ba1-41ec-aadc-2f3009bd911b

Tjondronegoro, D. (2024). *EPIC model: Strategic pillars for AI implementation at scale* (v 1). arXiv. https://arxiv.org/abs/2410.01819

Tjondronegoro, D. (2025). *TOAST framework: Trustworthy, Optimized, adaptable, socio-technological AI integration* (v 1). arXiv. https://arxiv.org/abs/2502.00011

Tjondronegoro, D., Yuwono, E. & Richards, B. (2022). *Responsible AI implementation: A human-centred framework for accelerating the innovation process* (v 1). arXiv. https://arxiv.org/abs/2209.07076

United States Executive Office of the President. (2023). *Executive Order 14110 on the safe, secure, and trustworthy development and use of artificial intelligence. Federal Register, 88*(210), 75191–75214.

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. WHO Press. https://apps.who.int/iris/handle/10665/341996

LIST OF TABLES

LIST OF FIGURES

ABOUT THE AUTHOR

Dr. Dian Wirawan Tjondronegoro
Professor
Business Strategy and Innovation
Griffith University
Australia

WATCH THE PRODUCTIVITY TALK ON
Risk Management for Resiliency and Productivity