＜Article＞
Strategic management of national AI capability: linking AI skills, shared computing resources, and governance systems – comparative evidence and pathways for Japan

Yuusuke Harada*, Hiroaki Kajikawa*

I. Introduction

Generative AI and large general-purpose models (often called 'foundation models') have moved quickly from cutting-edge research into everyday organisational use. Studies of workplace trials and labour-market exposure suggest that effects on productivity vary by task, and that many jobs change rather than disappear. This means the main bottleneck is often not the AI tool itself, but the supporting know-how: how organisations learn, govern, and improve AI-enabled work[1-5].

Governments are therefore investing in 'AI skills'. But building capability is rarely just about training. Organisations also need places to experiment and evaluate (computing power, tools, and data) and practical governance systems (standards, audit routines, procurement rules, and safety evaluation) so that use cases can be checked and compared. When these parts are managed separately, predictable problems follow: people are trained but cannot test; compute is concentrated among a few actors; and high-level principles never turn into everyday records that managers can review[6-9].

Because the journal has many readers from the humanities and social sciences, we use the following plain-language definitions for key terms in this paper:

- **AI capability**: the practical ability to use AI in real organisations in a productive, safe, and accountable way (not only knowing about AI).
- **Generative AI**: AI that produces new content such as text, images, or code based on patterns learned from data.
- **Foundation model**: a large general-purpose model trained on broad data that can be adapted to many tasks.
- **Shared computing resources (shared compute)**: shared access to computing power (often GPUs) and common tools, models, and datasets so that many organisations can experiment and evaluate without owning their own hardware.
- **Governance systems**: the rules and routines that make AI use traceable and controllable, such as policies, standards, documentation, audits, and incident response.
- **Assurance / safety evaluation**: tests and reviews that check how a model behaves (for example robustness and security) and whether people are following the agreed rules.

* Graduate School of Humanities and Social Sciences, Hiroshima University

- **Operational evidence (practical records)**: everyday documents and logs that show what was done - such as risk assessments, access logs, evaluation reports, and incident logs.

We focus on the policy package—a linked set of policy tools that connects skills development, shared computing resources, and governance systems. We compare the European Union, the United States, and the United Kingdom—countries and regions that have begun to put duties, infrastructure, and evaluation capacity in place—and then use Japan as an implementation case to suggest practical linkage paths under coordination and budget constraints[10-13].

We ask two questions: (1) How do these leading countries and regions combine (i) skills, (ii) shared computing resources, and (iii) governance and safety checks? (2) What design lessons and practical indicators follow from the comparison, and how could Japan implement a linked system that is both convincing and feasible for day-to-day administration?

We contribute in three ways. First, we present a simple framework that links capability layers to enabling infrastructure and highlights the role of practical records in policy management. Second, we draw cross-case lessons about linked policy design, record-keeping, compute access rules, and evaluation capacity. Third, we turn these lessons into a minimal set of reusable documents and a phased roadmap for Japan, with implications for other later-starting innovation systems.

II. Conceptual framework: capability layers, infrastructures, and linkage

1. Capability layers: literacy, implementation, and frontier

AI literacy is often described as knowledge, skills, and attitudes. This is useful, but for policy and organisational management it is incomplete unless we link it to who is responsible and what counts as competent practice. In real workplaces, key questions are: Who needs to know what? Who must do what? And what kind of evidence shows that the work was done properly? Education frameworks for students and teachers break literacy into teachable elements, but they do not automatically create the routines organisations need[14-16].

We therefore distinguish three capability layers, each requiring different policy tools. The literacy layer is responsible everyday use: knowing when AI fits a task, when human judgement is needed, how to handle privacy and copyright, how to cite sources, and when to ask for review. The implementation layer is the ability to make AI work inside an organisation: buying AI tools and services (procurement), setting acceptable-use rules, handling data, keeping logs, checking suppliers, ensuring human oversight, doing internal audits, and responding to incidents. The frontier layer is advanced development and testing: training or adapting models, stress-testing them (‘red teaming’), using secure research environments, and joining leading research ecosystems. This breakdown clarifies a common confusion: broad literacy programmes are necessary, but they do not replace implementation competence or frontier capability.

2. Enabling infrastructures: training, shared computing resources, and governance

If capability has layers, it also needs supporting infrastructure. In the generative AI era, governments

increasingly act in at least three areas: training infrastructure (curricula, accreditation, assessment, and ongoing professional learning); shared computing resources and data infrastructure (shared access to GPUs, models, datasets, tools, and evaluation environments); and governance systems (standards, guidelines, audit processes, and organisations that test safety). Short courses and micro-credentials can help scale training, but they matter most when they connect to real organisational practice and assessment [7,17].

Figure 1 represents this as a three-infrastructure model aligned with the three capability layers. The model is analytically useful because it shows why capability-building often stalls: a country may invest in courses but provide little access to experimentation environments; it may provide compute but without access rules, user support, or assurance methods; or it may publish governance principles that are not translated into routines, documentation requirements, or feedback mechanisms.

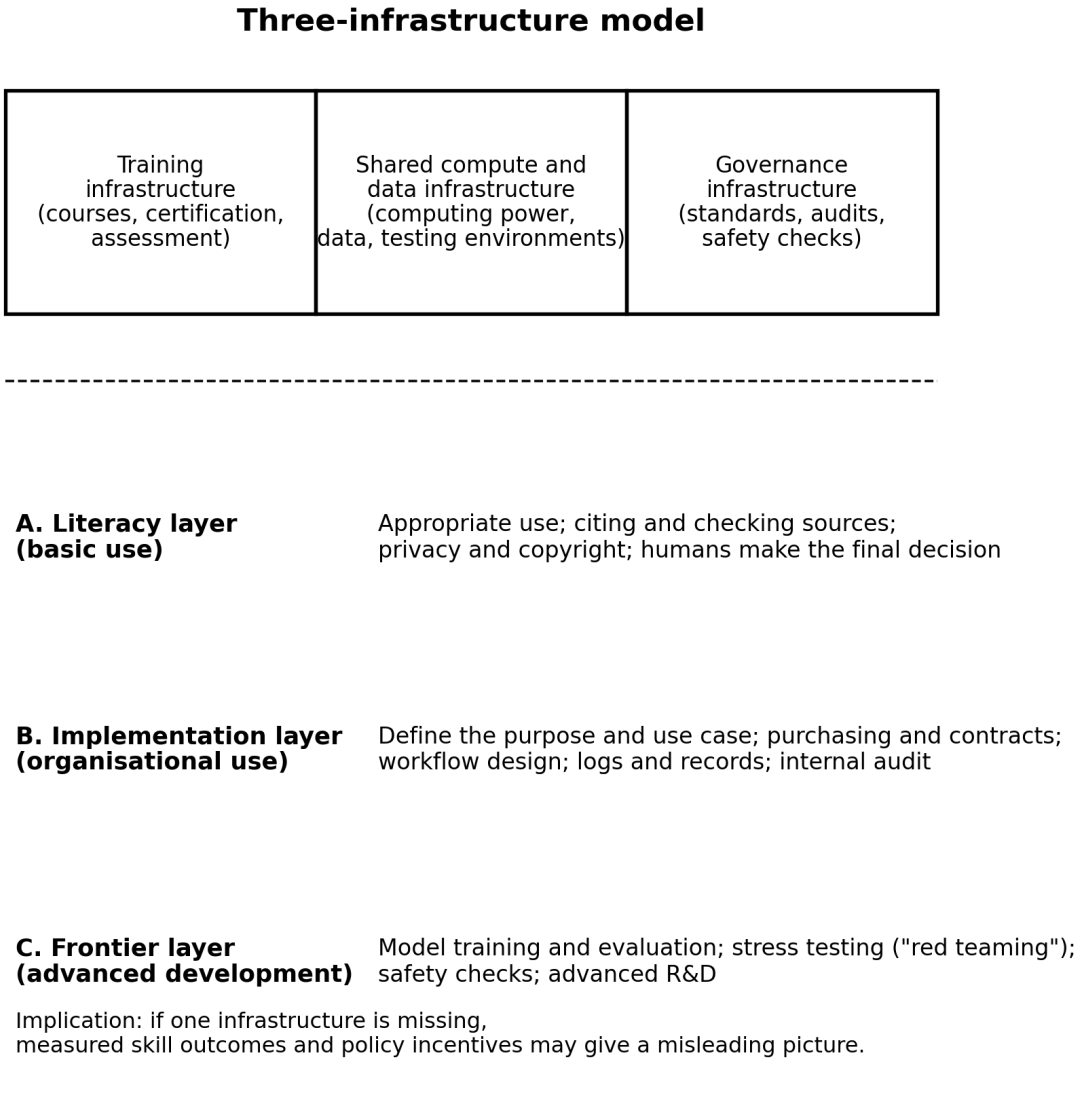

Figure 1. Capability layers (literacy, implementation, frontier) mapped to enabling infrastructures (training; computing resources and data; governance systems) (Created by the author).

Treating compute as infrastructure is especially important. In policy debates, compute is sometimes seen as a purely technical or industrial issue, separate from workforce development. Yet hands-on ability in evaluation, benchmarking, operating and maintaining models in practice (often called MLOps), and safety testing often depends on access to compute and controlled environments. Compute policy also raises trade-offs about allocation, cost, and energy impacts, especially as training and evaluation workloads grow[18-20].

Governance systems matter just as much. Risk-management and management-system frameworks connect competence requirements to procedures, documents, and regular review. They also make capability visible to managers and auditors, enabling continuous improvement rather than one-time compliance. In this sense, governance systems do not only restrict innovation; they help organisations build, trust, and scale AI capability[6,7,8,21].

3. Linkage, evidence, and feedback in policy management

A key concept in this paper is linkage: the practical links among policy tools, organisational routines, and feedback mechanisms. In a linked system, training outcomes inform procurement criteria, procurement decisions shape logging and supplier assessment, and practical records feed back into curriculum updates, standards, and resource allocation. An unlinked system may list the same tools on paper, but they do not reinforce each other.

Practical records are the glue for linkage. Here, ‘evidence’ does not mean only research studies; it also includes everyday documents such as training completion records, role definitions, risk assessments, access logs, model documentation, supplier checklists, incident reports, and corrective-action logs. These records clarify responsibility, support resource decisions (including compute allocation and procurement review), and enable learning because fixes can be traced over time. Tools such as model cards and datasheets show how standardised documentation can reduce information gaps in procurement and oversight[8,22,23].

This perspective suggests a different evaluation logic for capability programmes. Counting participants or accredited courses is administratively convenient, but it says little about whether organisations can procure, govern and improve AI systems in practice. A stronger policy management approach combines output indicators with evidence on process quality, resource use, incident handling and fixes. Training evaluation approaches can be repurposed here, but only if they connect to operational performance and governance routines rather than stopping at satisfaction or knowledge tests [24].

III. Design and methods

1. Scope, case selection and research questions

The review centred on three comparative cases and one practical implementation case. The comparative cases were the European Union, the United States and the United Kingdom. These countries and regions were selected because they have moved beyond generic AI strategies toward more operational combinations

of duties, infrastructure investment and evaluation capacity. Japan was analysed as a practical implementation case because it combines substantial soft-law and existing institutions and tools with linkage challenges typical of many innovation systems.

The review was guided by the two research questions in Section 1. To answer them, we examine: (i) who each policy package targets and what duties it creates; (ii) how implementation is expected to happen in practice; and (iii) what indicators or records make capability and compliance visible.

2. Identification, screening and source selection

Searches were conducted in academic databases and public online repositories. For the academic component, the search covered Scopus, Web of Science, Google Scholar and arXiv, with particular attention to literature on AI literacy, workforce impacts, governance, auditing, public procurement, shared computing resources (sometimes discussed as 'shared compute'), and evaluation capacity. The search terms combined domain terms (e.g., 'AI literacy', 'generative AI', 'foundation models') with governance and infrastructure terms (e.g., 'risk management', 'audit', 'compute roadmap', 'shared compute', 'AI safety institute').

In parallel, we compiled public official documents, including laws and regulations, executive orders, standards, official programme documents, infrastructure documentation and agency guidance. Screening proceeded from title/abstract to full-text checks, focusing on coverage (who it applies to), how things are expected to work in practice and the availability of easy to check indicators. The overall workflow is summarised in Figure 2[25)].

We included a source if it did at least one of the following: (a) described who is covered (for example, which actors have duties); (b) explained a practical mechanism (for example, how resources are allocated or how compliance is assessed); or (c) provided quantitative indicators relevant to capability building. The final collection included 97 sources, combining academic studies with policy and infrastructure documents. A full coded bibliography is available from the author on request.

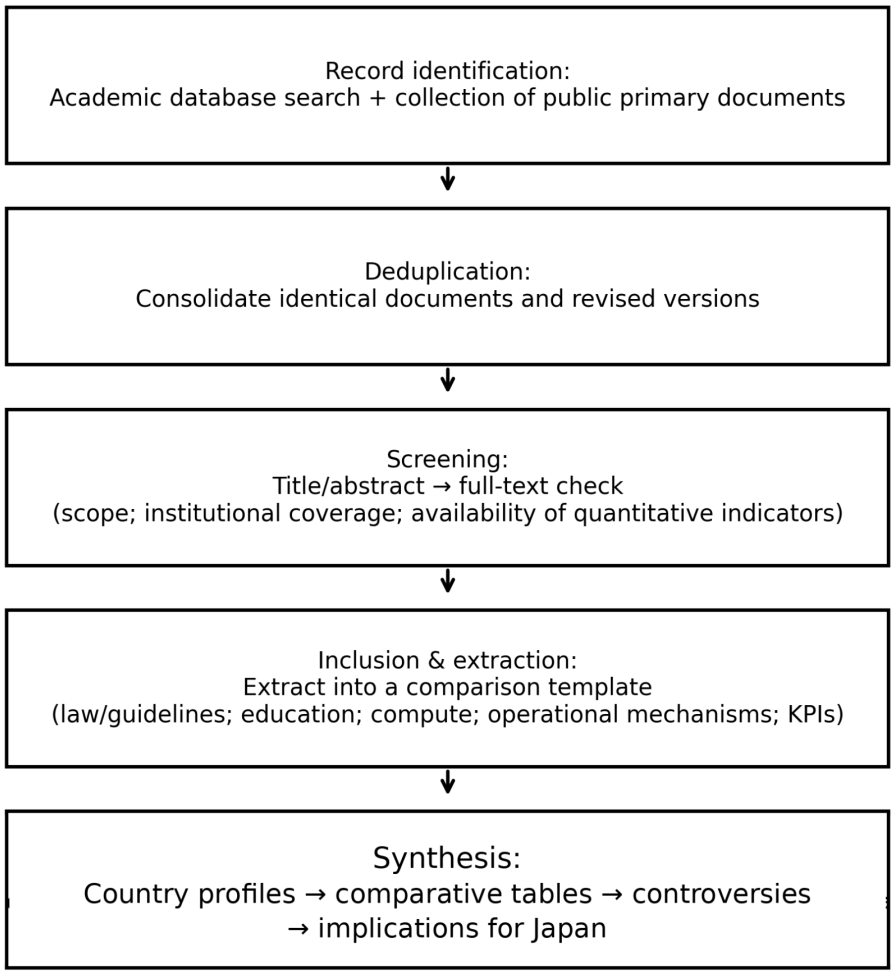

Figure 2. Broad review workflow: identification, screening, extraction and synthesis (Created by the author).

3. Coding frame and comparative strategy

Each source was coded into a common extraction template asking: who is targeted; what type of obligation or incentive is created; which enabling infrastructure is addressed (training, compute/data, governance); what how things are expected to work in practice are specified; and what evidence documents or indicators are implied. Coding did not stop at topic coverage. Each document was read for its implied theory of implementation—whether behaviour change was expected to occur through legal duty, procurement leverage, resource access, guidance, accreditation, or audit feedback.

Two comparative choices are important. First, we treat policy packages rather than individual documents as the unit of interpretation, because instruments function through complementarities and ordering. Second,

we emphasise linkage mechanisms (practical records and feedback loops) as the critical feature differentiating ‘paper strategies’ from implementable capability systems.

4. Analytical limits

As a broad review, this study does not attempt systematic effect-size synthesis. It is designed to map instrument combinations, how things are expected to work in practice and practical implementation pathways. The design therefore prioritises analytic rigour in comparison of how systems are set up over comprehensive enumeration of all national initiatives.

IV. Comparative findings I: institutionalising AI literacy, governance and assurance

1. From voluntary training to organisational duty

A first shared trend is the movement from AI literacy as a desirable educational goal to AI literacy as an organisational duty. The EU AI Act makes this shift most explicit by introducing an AI literacy duty and linking it to a risk-based governance system for providers and deployers. This matters because it changes what counts as ‘skills’: competence must be demonstrated through organisational routines, not merely through participation in training[10)].

The United States and the United Kingdom have taken different routes, relying more on executive direction, standards, and institutional capacity than on one broad, horizontal AI law. In the US case, Executive Order 14110 and related agency actions frame trustworthy AI as a cross-sector responsibility, and standards help turn this goal into practical steps. In the UK case, building safety-evaluation institutions signals a similar move toward responsibility that can be checked in practice[8,26,27)].

For policy management, the implication is that literacy initiatives need linkage to organisational evidence: training records that connect to role assignments; procurement criteria that require documented governance; and incident learning mechanisms that feed back into training design. Without such linkage, literacy remains a diffuse aspiration rather than a managed capability.

2. Standards, auditability and evaluation capacity

A second finding is that standards and assurance frameworks help turn high-level policy goals into concrete organisational practice. NIST’s AI RMF and the Generative AI Profile outline risk functions and categories that organisations can build into internal processes, while ISO/IEC 23894 and ISO/IEC 42001 provide risk-management and management-system approaches that support routines that can be checked (audited)[6,7,8,21)].

The practical importance of these frameworks lies in their documentary consequences. They encourage, and often implicitly require, documents such as role definitions, risk registers, evaluation reports and incident logs. Model documentation practices (model cards, datasheets) provide complementary documents that can travel through procurement and oversight processes, reducing information gap between developers, deployers and regulators[22,23)].

Evaluation capacity is an equally important part of this picture. Institutions such as the UK AI Safety

Institute signal that evaluation is becoming a policy capability in its own right rather than a purely technical concern, especially for generative models where claims about safety and performance can be contested and context-dependent[27].

3. Labour-market and adoption evidence

The labour-market and workplace literature reinforces the need for an integrated view. Experimental evidence suggests meaningful productivity gains in some settings, but also indicates that benefits depend on task characteristics, worker experience and organisational arrangements. Labour-market impact studies highlight exposure of specific task clusters and potential distributional effects, implying that 'skills' policy must include organisational change and not only individual training[1,2,5].

Macro-level analyses similarly emphasise uneven exposure and the likelihood that generative AI reshapes job quality and task content. These findings support a policy logic in which capability-building is designed around task transformation, institutional support and governance routines, not merely tool exposure[3,4,28,29,30].

V. Comparative findings II: shared computing resources as a capability-building instrument

1. Why compute matters for learning, experimentation and assurance

In the foundation-model era, compute is not merely an input into frontier R&D. It shapes what can be taught, tested, benchmarked and governed. Capability at the implementation and frontier layers often requires access to controlled environments for evaluation, red teaming and secure experimentation. As a result, compute access affects inclusion: which organisations can learn by doing, and which can only consume tools without the ability to evaluate claims.

Compute also creates public policy tensions around allocation, cost and sustainability. The concentration of advanced compute in a small number of firms and locations raises questions of competition, resilience and national dependence. At the same time, energy and carbon lessons make compute policy politically salient, particularly when public funds are used to subsidise access[18,19,20].

A key comparative insight is that governments increasingly treat shared computing resources as a managed service rather than as a passive asset. Access regimes, review criteria, user support, and the availability of evaluation tooling determine whether shared computing resources produce broad-based capability or only support a narrow research elite.

2. European Union: enabling ecosystems through AI Factories

The European Union's approach combines regulation with enabling infrastructure. On the regulatory side, the AI Act provides a visible governance anchor. On the enabling side, the AI Innovation Package and the AI Factories initiative aim to expand access to compute and support the ecosystem required for compliant innovation[10,11,31].

The EU case is noteworthy for policy management because it does not rely on regulation alone. Its design recognises that compliance obligations and innovation capacity have to be built together. Programmes for

shared computing resources are coupled with support functions (training, tooling, evaluation environments) that reduce the time and effort costs of experimentation for smaller actors.

3. United States: NAIRR as a managed linkage model

The US case illustrates a different route. Rather than relying on a single horizontal AI law, the United States has emphasised standards, guidance and resource linkage. The National AI Research Resource (NAIRR) Pilot frames compute access as a reviewed and supported service that links compute, data and tools for the research community[12].

NAIRR is significant as a policy-management model because it recognises that access on its own is insufficient. Users need eligibility logic, review processes, user support functions and governance rules. For countries that cannot replicate US resource scale, the transferable lesson is not the volume of compute, but the managed-linkage principle: treat compute access as an instrument connected to capability development and governance routines.

4. United Kingdom: compute roadmap and safety evaluation institutions

The UK Compute Roadmap positions compute investment as strategic infrastructure linked to innovation and security objectives rather than as a narrow research input[13].

In parallel, the UK has invested in safety evaluation institutions. The AI Safety Institute signals that assurance capacity is a public asset, supporting both adoption and the ability to contest or verify safety claims[27].

For policy design, the UK case highlights the value of aligning compute investment with evaluation institutions, and of treating evaluation outputs as reusable evidence that can travel through procurement and organisational governance.

5. Cross-case synthesis: access rules and support functions

Across cases, the design levers that matter most are not only hardware scale but governance of access: eligibility, prioritisation, pricing or subsidy rules, support functions (training and tooling), and integration with assurance processes. These levers determine whether shared computing resources expand the set of capable actors and support trustworthy spread.

Table 1 provides a compact cross-case overview of how each country or region combines skills, shared computing resources, and governance instruments.

Country/region	Skills / literacy (main tool)	Shared computing resources (main tool)	Governance / safety checks (anchor)	How it works (in practice)
European Union	AI Act makes AI literacy a duty and sets risk-based duties for organisations	AI Factories / EuroHPC expand access to computing resources and evaluation environments	AI Act compliance supported by standards and compliance checks	Combine regulation with support infrastructure so organisations can innovate and comply

United States	Guidance and standards emphasise workforce development and organisational responsibility	NAIRR Pilot provides reviewed access to computing resources, data, and tools	NIST AI RMF and executive direction on safe, trustworthy AI	Treat compute access as a managed shared service for the public interest, with user support
United Kingdom	Adoption and skills programmes linked to innovation and security policy	UK Compute Roadmap sets a national path for computing investment	AI Safety Institute and investment in safety-testing capacity	Align computing investment with evaluation bodies to support trustworthy spread
Japan	Education and business guidance (schools and organisations), but uneven uptake	ABCI 3.0 and shared HPC resources for R&D and experimentation	AI Business Guidelines, procurement guidance, privacy and security standards	Many tools exist, but weak linkage of records across education, procurement, and shared computing resources

Table 1. Comparative policy packages for AI capability building (skills, shared computing resources, and governance).

VI. Cross-case design propositions and indicators

Synthesising the comparative findings, we derive four propositions that specify where capability-building succeeds or fails as a managed set rather than as isolated programmes. Each proposition implies operational indicators that policy managers can track to assess linkage and learning.

Proposition 1 (linked sets of policies): Capability-building is produced by linked sets of policies linking literacy, resource access and assurance. Skills programmes raise baseline participation, but without experimentation environments and governance routines they do not translate into organisational capability. Conversely, compute investments without linked training and assurance risk concentration, underuse and politicised subsidy debates.

Proposition 2 (evidence linkage): Responsibility without practical records does not scale. Assigning duties is only effective when tied to reusable documents—training records, role matrices, risk registers, procurement disclosures, evaluation reports and incident logs—that enable accountability and learning across organisations. Standards-based management approaches provide a mechanism for making such documents usable across organisations[7,8].

Proposition 3 (allocation rules): Compute allocation rules are themselves skills and innovation policy. Eligibility criteria, prioritisation and support functions determine whether shared computing resources expand capability in SMEs, universities and the public sector or reinforce existing advantage. Transparent allocation criteria can also reduce conflict over subsidised access and align compute use with public-interest objectives.

Proposition 4 (assurance capacity): Assurance capacity is a strategic public asset. Domestic evaluation and red-teaming capacity reduces dependence on external providers for safety claims and supports both

adoption and frontier research. Capability policy should therefore track not only training outputs and compute capacity, but also evaluation throughput, incident learning and standards uptake[21,27].

Table 2 summarises these propositions and offers indicative operational indicators for policy managers.

Design lesson	What it means for policy	Example indicators (what to measure)
Linked policies (skills–compute–governance)	Manage capability building as a set of linked programmes; avoid skills-only or compute-only plans	Coverage across the three domains; cross-ministry steering and feedback loop
Evidence linkage	Duties and guidance work at scale only when tied to reusable documents/records and review routines	Template use rates; audit closure rates; incident reporting and follow-up tracking
Allocation rules	Rules for compute access and user support shape who can learn by doing	Diversity of users (SMEs, public sector, universities); transparency of prioritisation criteria
Public capacity for testing and evaluation	Invest in evaluation, stress-testing (‘red teaming’), and standards-based management to reduce uncertainty	How many evaluations can be done; use of AI management systems; domestic safety-testing capability

Table 2. Cross-case design lessons and example indicators.

VII. Japan as an implementation case: orchestration under coordination constraints

1. Current policy fragments and institutional assets

Japan has accumulated a substantial set of instruments relevant to capability-building, including cross-ministry business guidance, education guidance, public-sector procurement rules, and shared HPC capability. Examples include the AI Business Guidelines, school-level guidance on generative AI use, and procurement guidance for government administration, alongside the ABCI 3.0 infrastructure and related R&D support[32-36].

Japan has also attempted to put in place education capacity through programme accreditation and related public reporting. These instruments can help scale training, but their strategic value depends on whether accredited skills are linked to procurement requirements, organisational governance routines and access to experimentation environments[37].

The main weakness is not instrument absence but linkage. Education initiatives, procurement requirements and infrastructure access are governed by different institutional logics and rarely share evidence. As a result, training completion does not automatically inform procurement qualification; procurement documentation does not reliably feed into audit and incident learning; and access to shared computing resources is not systematically tied to capability development or assurance routines. This pattern increases duplication, weakens evaluability and makes accountability unclear[38-40].

2. A linkage model based on operational evidence

A practical response is to standardise a minimal practical records set—small, reusable documents that travel across instruments and reduce coordination cost. The objective is not paperwork expansion; it is interoperability: a shared evidence grammar that allows ministries, universities, infrastructure operators and suppliers to coordinate without requiring a single new AI law at the outset.

Figure 3 sketches a linkage model centred on a single entry point (for procurement, compute access, or capability programmes) connected to review, support, execution and evaluation functions. In such a model, the evidence documents generated at each step become inputs to the next step and to policy learning.

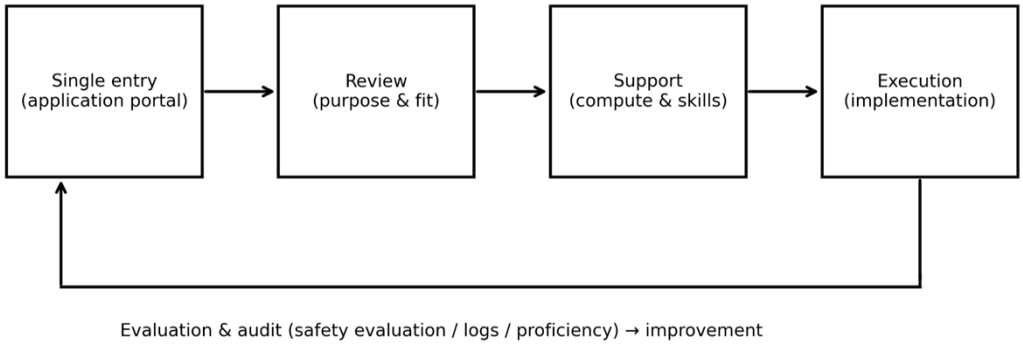

Figure 3. Linkage model for capability building: single entry point, review, support, execution, and evaluation/audit feedback (Created by the author).

Table 3 proposes a deliberately modest evidence set centred on roles, acceptable use, risk registers, procurement disclosures, access logs and incident learning. These documents can be embedded into procurement templates, compute access procedures and training evaluation, turning fragmented guidance into an implementable system.

Record / document	**What it should include (minimum)**	**Main owner**	**Why it matters / how it is reused**
Roles and responsibilities list	Named roles for supplier, user organisation, approver, and incident contact	Business owner + legal + security	Clarifies responsibility across training, procurement, audit and incident response
Acceptable-use rules	Permitted/prohibited uses, disclosure rules, exception handling	Responsible unit + HR	Turns guidance into clear boundaries and training content
Use-case list & risk assessment	Purpose, data sensitivity, risk controls, review schedule	Project owner + risk function	Makes adoption measurable and supports prioritised assurance effort
Procurement disclosure package	Model documentation, data sources, evaluation claims, security posture	Procurement + vendor	Reduces information gap; supports comparable procurement decisions
Access & usage logs (including shared	Who accessed what resources, when, for	IT / compute operator	Supports traceability, misuse detection and

computing resources)	which use case		cost allocation
Incident & corrective-action log	Incident categories, root cause, mitigations, closure evidence	Operations + audit	Turns failures into learning and supports continuous improvement

Table 3. Minimal operational records set for organisational AI capability building (illustrative).

3. Phased roadmap and governance options

Implementation should be phased. A preparation phase can align ministries and operators on common templates and minimum definitions (e.g., what counts as a 'use case', 'evaluation', or 'incident'). A pilot phase can embed templates into selected procurement exercises, education pilots and shared-compute access processes, generating evidence for refinement. Scaling can then proceed through guidance updates and procurement standardisation, supported by assurance institutions and audit routines.

A key governance choice is where to locate the 'orchestration' function. One option is a light coordination unit that curates templates, aggregates evidence signals and maintains feedback loops (e.g., updating procurement clauses and training guidance based on incident learning). Another option is to embed orchestration within existing infrastructure operators (shared computing resources, sectoral centres) who can provide hands-on support. Either way, the core requirement is to keep evidence portable across ministries and sectors.

This ordering is transferable to other innovation systems because it prioritises linkage and learning over comprehensive new legislation at the outset. It also supports strategic flexibility: as model capabilities and risks evolve, evidence templates and evaluation routines can be updated without redesigning the entire policy framework.

VIII. Discussion: strategic management of national AI capability

For technology analysis and strategic management, the core implication is that generative AI capability is a managed system-level asset. Treating skills, compute and governance as separate policy silos creates predictable common problems: skills without experimentation capacity, compute without inclusive allocation, and governance without reusable evidence. Conversely, well-designed packages can reduce uncertainty for adopters, lower time and effort costs in procurement, and support more distributed innovation across sectors.

This perspective suggests a shift in performance measurement. Completion counts, accredited programmes and GPU totals matter, but they do not reveal whether evidence travels across instruments or whether incidents generate corrective action. Operational indicators—template adoption rates, audit closure rates, incident reporting volumes, evaluation throughput, and compute allocation diversity—can provide early signals of whether capability-building is becoming scalable and trustworthy[6,21].

The analysis also highlights a strategic trade-off. Shared computing resources and assurance capacity can be framed narrowly as research infrastructure, or more broadly as an innovation-system service supporting

experimentation, evaluation and spread. The latter framing aligns better with generative AI's rapid spread into non-frontier sectors and with the need for trustworthy adoption. Japan's case illustrates that even without frontier-scale budgets, policy can improve outcomes by reducing coordination cost and making responsibility easy to check through modest evidence standardisation.

IX. Conclusion

This paper compared how the EU, US and UK assemble policy packages for AI capability building and used Japan to derive feasible implementation pathways. The main lesson is that capability depends on the co-management of people, resources and assurance: skills policy, shared computing resources and governance systems are complements. Linked sets of policies, evidence linkage, allocation rules and assurance capacity are the key design levers. For Japan and similar innovation systems, a minimal evidence set and phased linkage roadmap can improve evaluability and adoption without requiring immediate comprehensive new legislation.

References

1. Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at Work. NBER Working Paper No. 31161. https://doi.org/10.3386/w31161.
2. Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. arXiv:2303.10130.
3. International Labour Organization (ILO). (2023). Generative AI and Jobs: A global analysis of potential effects on job quantity and quality.
4. International Monetary Fund (IMF). (2024). Gen-AI: Artificial Intelligence and the Future of Work (Staff Discussion Note).
5. Noy, S. & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science, Vol. 381 No. 6654, pp. 187-192. https://doi.org/10.1126/science.adh2586.
6. ISO/IEC. (2023). ISO/IEC 23894:2023 Artificial intelligence - Guidance on risk management.
7. ISO/IEC. (2023). ISO/IEC 42001:2023 Artificial intelligence - Management system.
8. National Institute of Standards and Technology (NIST). (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1.
9. OECD. (2025). Bridging the AI skills gap.
10. European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 2024-06-13 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.
11. European Commission. (2024). AI Innovation Package (communication and accompanying documents).

12. National Science Foundation (NSF). (2024). National Artificial Intelligence Research Resource (NAIRR) Pilot (official program page).
13. UK Department for Science, Innovation and Technology (DSIT). (2025). UK Compute Roadmap. 2025-07-17.
14. European Commission Joint Research Centre. (2022). The Digital Competence Framework for Citizens (DigComp 2.2).
15. UNESCO. (2024). AI competency framework for students.
16. UNESCO. (2024). AI competency framework for teachers.
17. OECD. (2021). Micro-credentials for lifelong learning and employability.
18. Patterson, D., et al. (2021). Carbon emissions and large neural network training.
19. Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. ACL.
20. HAI (Stanford Institute for Human-Centered Artificial Intelligence). (2025). AI Index Report 2025.
21. National Institute of Standards and Technology (NIST). (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1. 2024-07-26.
22. Gebru, T., Morgenstern, J., Vecchione, B.,, et al. (2021). Datasheets for Datasets. Communications of the ACM, Vol. 64 No. 12, pp. 86-92.
23. Mitchell, M., Wu, S., Zaldivar, A.,, et al. (2019). Model Cards for Model Reporting. Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT*).
24. Kirkpatrick, J. D. & Kirkpatrick, W. K. (2016). Kirkpatrick's Four Levels of Training Evaluation. ATD Press.
25. Tricco, A. C., Lillie, E., Zarin, W.,, et al. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Annals of Internal Medicine, Vol. 169 No. 7, pp. 467-473. https://doi.org/10.7326/M18-0850.
26. Executive Office of the President of the United States. (2023). Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. 2023-10-30.
27. UK Government (Department for Science, Innovation and Technology). (2023). Introducing the AI Safety Institute. 2023-11.
28. World Economic Forum. (2025). The Future of Jobs Report 2025. 2025-01.
29. McKinsey & Company. (2024). The State of AI in 2024.
30. Microsoft and LinkedIn. (2024). Work Trend Index 2024.
31. 31) EuroHPC Joint Undertaking. (2025). EuroHPC JU selects AI Factory Antennas to broaden the AI Factories initiative. 2025-10-13.
32. Ministry of Internal Affairs and Communications (MIC), Japan; Ministry of Economy, Trade and Industry (METI), Japan. (2025). AI Business Guidelines (Version 1.1) [in Japanese]. 2025-03-28.
33. Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. (2024).

Guidelines on the Utilization of Generative AI in Primary and Secondary Education (Version 2.0) [in Japanese]. 2024-12-26.

34. Digital Agency, Government of Japan. (2025). Guidelines for Procurement and Use of Generative AI for the Evolution and Innovation of Government Administration [in Japanese]. 2025-05-27 (updated 2025-06-13).
35. ABCI. (2025). ABCI 3.0 User Guide: overview of the ABCI system.
36. National Institute of Advanced Industrial Science and Technology (AIST), Japan. (2024). Accelerating cutting-edge generative AI R&D and social implementation with ABCI 3.0 (press release; including system specifications) [in Japanese]. 2024-10-10.
37. METI, Japan. (2025). Accreditation and selection status for the FY2025 Mathematics, Data Science and AI Education Program Accreditation System (news release) [in Japanese].
38. Cabinet Office, Government of Japan. (2025). AI Strategy Council: public materials (e.g., 13th meeting) [in Japanese].
39. PPC (Personal Information Protection Commission), Japan. (2024). Perspectives on Generative AI and Personal Information Protection (collection of published materials) [in Japanese]. 2024–2025.
40. National center of Incident readiness and Strategy for Cybersecurity (NISC), Japan. (2023). Unified Standards for Information Security Measures for Government Agencies (latest version) [in Japanese]. 2023–2025.