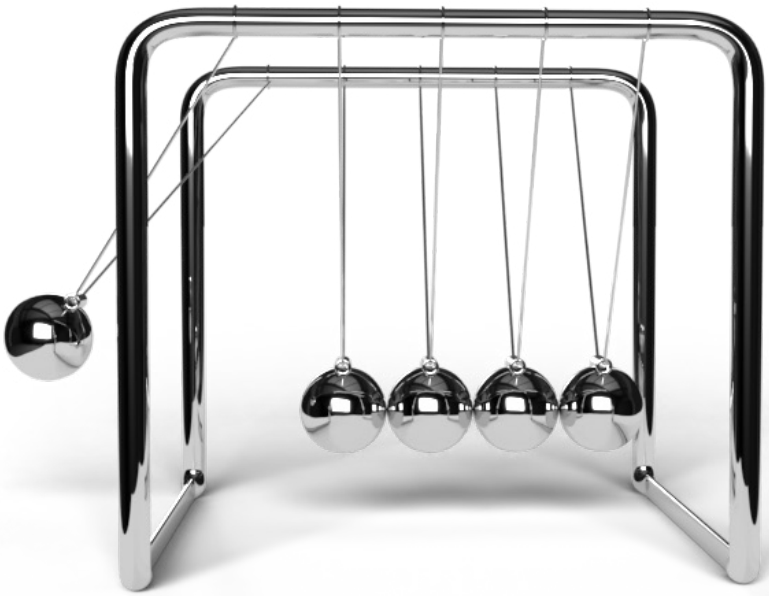
07 ASEF HIGHER EDUCATION innoLab

WHITE PAPERS

The Role of Universities in Advancing AI for Good

November 2026

ASIA-EUROPE FOUNDATION



Publisher

Asia-Europe Foundation (ASEF)

31 Heng Mui Keng Terrace

Singapore 119595

www.asef.org

Edited by

Ms Réka Tózsa

Dr Claudio Rivera

Ms Cleo Cachapero

Graphic & Layout Designer

Mr Mikhail Plata

ISBN 978-981-94-8358-7

Copyright

Asia-Europe Foundation (ASEF) 2026. All rights reserved.

Open Access

This work is distributed under the terms of the Creative Commons Attribution-NonCommercial licence, which permits any non-commercial use, distribution, and reproduction in any medium, provided the original authors and source are credited.

A Note

This publication was produced in the framework of the 7th ASEF Higher Education Innovation Laboratory. Any views and opinions expressed in the papers included in this body of work are the sole responsibility of the authors and do not reflect the views of the Asia-Europe Meeting (ASEM) or the Asia-Europe Foundation (ASEF) or any of their partner organisations.

Table of Contents

Foreword 4

Acknowledgements 5

Introduction to the ASEFInnoLab7 White Paper Collection: 8
Universities as Catalysts of the Common Good through Reflective Engagement with AI

1 Bridging the AI Divide in Higher Education through Asia–Europe Cooperation 13

2 AI for Whose Good? A Governance Framework for University Investment in AI with Underserved Communities 27

3 The Need for a Contextualised Competence Framework for GenAI in Higher Education 39

4 What policy reforms are needed to enable universities to support equitable lifelong learning and reskilling through AI? 51

5 Humanising Learning in the Age of AI: 69
Policy Conditions for Pedagogy, Assessment and Learner Formation

6 Cross-Regional Generative AI Governance in Higher Education: 81
From Reactive Policies to Coherent Institutional Frameworks

7 Why Open Education Matters More Than Ever Before: 95
Co-creating SDG-informed Open Education with Trusted AI across the Asia-Europe Axis

8 Universities as Catalysts for Aligning AI Innovation with the Sustainable Development Goals 111

9 A Framework for University Leaders to Choose Strategic AI Policy Postures: A Sandbox, An Accelerator, A Bridge and A Futures Architect / Policy Advocate 125

Foreword

I am pleased to present the ASEFInnoLab7 White Paper Collection, the key output of the 7th ASEF Higher Education Innovation Laboratory (ASEFInnoLab7), dedicated to exploring "The Role of Universities in Advancing AI for Good."

Following ASEFInnoLab's launch in 2021, it has grown into a collaborative platform connecting higher education leaders, academics, researchers, and experts from across Asia and Europe. At its heart lies a shared commitment to fostering meaningful dialogue, exchanging knowledge, and developing ideas that respond to emerging challenges at the intersection of higher education, innovation, and artificial intelligence.

This seventh edition marks an important step forward in that journey. ASEFInnoLab7 builds on previous discussions about universities' role in AI innovation ecosystems and shifts the conversation towards a broader question: how can universities help ensure that the transformative potential of AI translates into meaningful benefits for society?

The nine white papers featured in this publication reflect the collective efforts of academics, researchers and experts from diverse backgrounds. Through a 6-month long co-creation process combining virtual collaboration and in-person exchanges, we brought together multicultural, interdisciplinary teams to develop policy-oriented insights grounded in their respective contexts.

The resulting papers embody the collaborative spirit of ASEFInnoLab. Beyond contributing to academic and policy discussions, they reflect our aspiration to strengthen universities' capacity to act as responsible innovators, trusted partners, and influential voices in shaping AI's role in society.

I extend my sincere appreciation to the white paper leads, contributors, our Strategic Advisor, resource persons, institutional partners, and colleagues whose dedication, expertise, and generosity have made this publication possible.

I hope this collection will serve not only as a valuable reference for higher education leaders and policymakers, but also as a starting point for continued dialogue, cooperation, and action across Asia and Europe.

As AI continues to transform our societies, it is our hope at the Asia-Europe Foundation that these contributions inspire universities to look beyond technological advancement and take an active role in shaping a future that is more inclusive, equitable, and oriented towards the common good.

Réka TÓZSA
Director, Education Department
Asia-Europe Foundation



Acknowledgements

This collection is the result of a shared undertaking by the participants, expert leads, and partners who have contributed to the 7th ASEF Higher Education Innovation Laboratory (ASEFInnoLab7). Below, we wish to recognise the individuals and organisations whose expertise, dedication, and support have been essential in bringing this publication to fruition.

At the heart of this collection are the nine white paper teams, whose work reflects six months of sustained collaboration, thoughtful exchange, and collective effort. Together, they have produced papers offering valuable insights into how universities can play a leading role in advancing artificial intelligence (AI) for the benefit of society. We warmly appreciate the time, energy, and expertise each member has devoted to the entirety of the programme alongside on-going professional and personal commitments. Their willingness to engage across disciplines, institutions, and regions has been fundamental to shaping the ideas presented in this volume.

We extend particular recognition to the white paper leads, whose leadership and dedication have guided their respective teams throughout the development process. Beyond ensuring the quality and guiding the direction of each paper, they have played a central role in driving the collaborative work that anchors the success of ASEFInnoLab7. We are grateful to the following experts for taking on this responsibility and working closely with ASEF throughout the programme:

1. Prof SHEN Yi and Mr ZHANG Shuyan, Fudan University, China
2. Dr Yang BONG, Birmingham Newman University; Dr Neelima SAILAJA and Dr Minjia CHEN, University of Nottingham, United Kingdom
3. Prof Daniel BURGOS, Universidad Internacional de La Rioja, Spain
4. Prof Lampros STERGIOULAS, The Hague University of Applied Sciences, Netherlands
5. Prof Zainal SANUSI and Prof Mira KARTIWI, International Islamic University Malaysia
6. Dr Shalinka JAYATILLEKE, La Trobe University, Australia
7. Dr João PITA COSTA, Mr Davor ORLIČ, and Dr Anja POLAJNAR, International Research Centre for Artificial Intelligence (IRCAI) under the auspices of UNESCO, Slovenia
8. Dr Sharina YUNUS, Universiti Teknologi Brunei, Brunei Darussalam
9. Prof Tejpavan GANDHOK and Prof Spriha BHANDARI, OP Jindal Global University, India

Our appreciation also goes to ASEFInnoLab7's Strategic Advisor, Dr Claudio RIVERA (RTU Riga Business School, Latvia), whose guidance and expertise have been integral to the programme's implementation. His contributions across the different activities of this edition, particularly in guiding the development and finalisation of the white papers, has helped ensure the coherence, quality, and relevance of this collection.

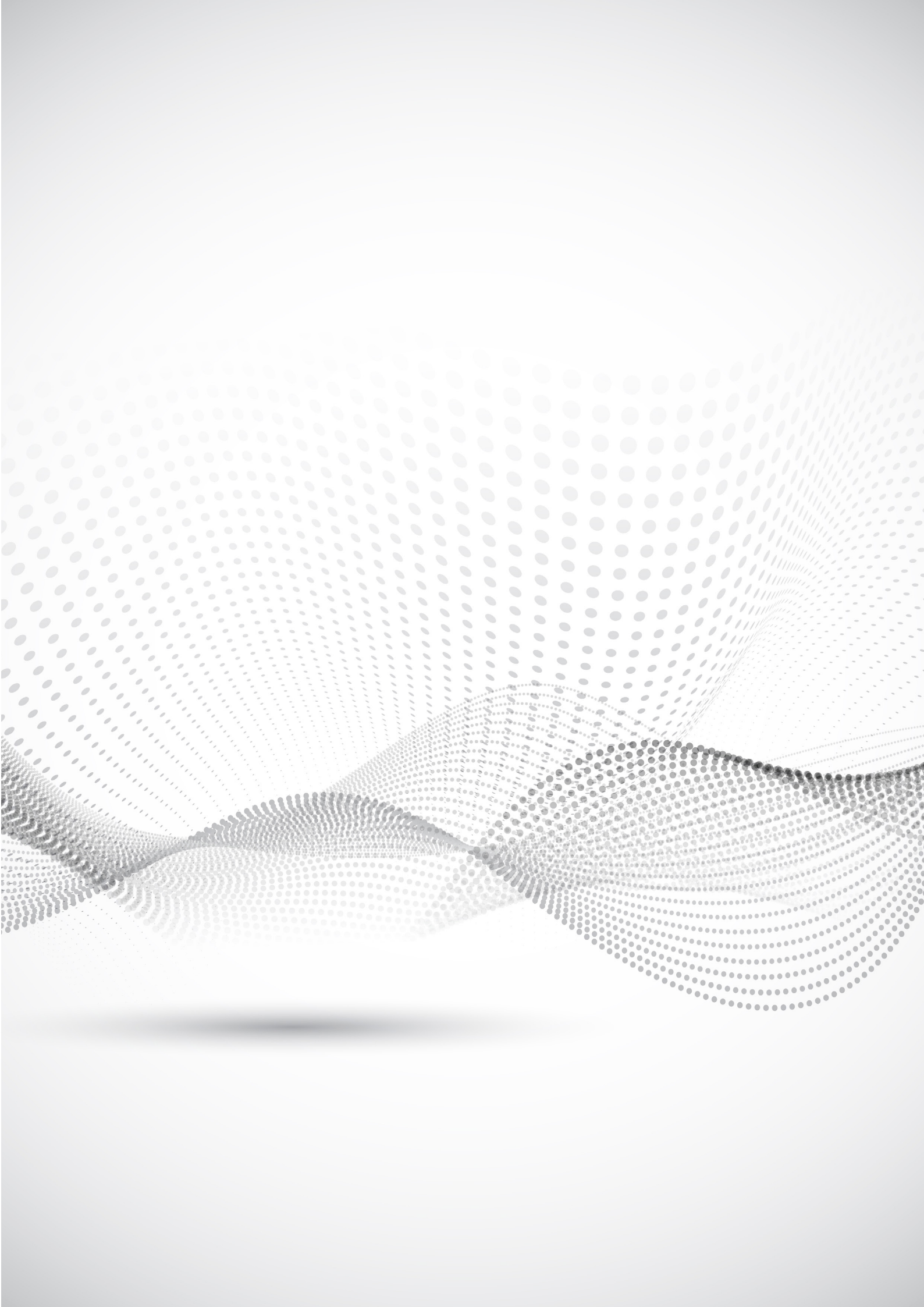
We are especially grateful to our implementing partners this edition, Fudan University in China and University of Nottingham in the United Kingdom, including its Horizon Centre of Doctoral Training and the Nottingham University Confucius Institute. Their intellectual and financial contributions as well as exceptional hospitality have been instrumental to the programme. Both institutions played a pivotal role by hosting an in-person event, contributing to its key activities, and leading a white paper team. Their partnership exemplifies the kind of institutional collaboration that ASEFInnoLab seeks to foster between Asia and Europe.

We likewise acknowledge the contributions of the rest of our partners: the International Research Centre for Artificial Intelligence (IRCAI), International Association of Universities (IAU), UN Internet Governance Forum (IGF), German Academic Exchange Service (DAAD), ASEM Education Secretariat (AES), and the Asia-Europe for AI (AE4AI) Network. Through their engagement, networks, and expertise, these organisations have enriched the programme's discussions, reinforced its cross-regional character, and helped bring its work to wider international audiences.

We express our sincere appreciation to the core team at ASEF's Education Department, especially Ms Reka TOZSA, Education Director; Ms Cleo CACHAPERO, Lead for Higher Education Projects; and the rest of the team for their contributions throughout the process. We also extend special thanks to Mr Mikhail PLATA, whose work on the design and layout has given this collection its final visual form.

Finally, we gratefully acknowledge the Ministry of Foreign Affairs, China for its generous funding support, which has made the implementation of ASEFInnoLab7 possible. This support has enabled a meaningful platform for knowledge exchange and collaboration between Asia and Europe, culminating in a collection that underscores the pivotal role of universities in ensuring that advances in AI translate into meaningful and lasting benefits for society.





Introduction to the ASEFInnoLab7 white paper collection

Universities as Catalysts of the Common Good through Reflective Engagement with AI

A shared inquiry with a practical purpose

The nine white papers in this collection are the result of the ASEFInnoLab7 co-creation process. They bring perspectives across Asia and Europe into a shared inquiry about how universities can use the development of artificial intelligence (AI) to serve the common good. Our purpose has been to provide strategic insights and actionable pathways for global, national and institutional leaders: a view of the key themes, the dilemmas they expose, and the alternatives through which leaders can respond. The collection invites readers to connect technological possibilities with educational purposes, institutional capabilities, and the people whose lives those choices will affect.

Co-creation is central to that purpose. The papers grew through collective discussion, the development of team arguments, and subsequent reflection and review. Participants had to translate experiences and competing interpretations into propositions that others could question, adapt and use. Working across cultures made differences in institutional resources, public expectations and ethical assumptions visible. These differences are a source of strategic knowledge. They help explain why an approach that appears effective in one setting may require different infrastructure, authority or relationships in another.

The collection connects three directions: AI4GlobalGood, AI4PublicInterest and AI4InstitutionalExcellence. Global good directs attention to sustainability and shared social challenges. Public interest asks whose understanding, participation and trust must be strengthened with a focus on national and regional scope of governance. Institutional excellence concerns the governance, capabilities and practices that allow universities to contribute responsibly. These directions depend on one another. An institution cannot sustain a public mission without the capacity to act provided by national authorities; technical capacity of national authorities earns public legitimacy through the purposes granted by international bodies and local communities.

Several questions hold the inquiry together. What role should universities earn as the production and circulation of knowledge change? When does AI support learning, and when does it replace the intellectual work through which learning occurs? Which ethical principles require shared protection, and where does context demand different implementation? What would support responsible institutional readiness? Each question leads towards a practical test: what concrete decisions can leaders make so that AI strengthens human capability and public value?

We approach these questions with an understanding of AI that extends beyond “public domain narratives”. AI interacts with changes in work, industry, demography and the organisation of knowledge. Universities encounter that transformation in teaching, research, administration and their relationships with communities and employers. The resulting decisions cannot be delegated entirely to technology suppliers or reduced to the acquisition of software.

Equity and the authority to define public benefit

The first paper, "Bridging the AI Divide in Higher Education through Asia–Europe Cooperation," reframes the divide as a challenge of institutional capability. Access to a tool is only one condition of meaningful participation. Universities also need infrastructure, skills, the ability to interrogate outputs, and sufficient influence over how systems are deployed. Otherwise, adoption can increase dependence while leaving the institution with little power to shape its educational and research priorities. Cooperation across regions becomes valuable when it strengthens the capacity of partners to act and decide.

This understanding of the divide matters for leaders at every level. Global cooperation can support shared knowledge, infrastructure and a wider voice in governance. National policy can address the conditions that prevent institutions or communities from participating. University leaders can examine who receives access, training and support within their own institution. Measures of adoption should therefore be complemented by attention to capability, inclusion and agency.

The second paper, "AI for Whose Good? A Framework for University Investment in AI with Underserved Communities," extends that inquiry to the definition of benefit itself. A university may describe a project as socially useful while giving the affected community little influence over the problem, the data or the criteria for success. The paper's Receiver–Doer–Judge distinction provides a way to examine who receives an intervention, who produces it and who decides whether it has created value. Community engagement belongs throughout the research process, including the decisions that establish its direction.

This has consequences for the use of resources and authority. Consultation after a system has been designed may reveal problems, but it offers limited opportunity to change the assumptions that shaped it. Engagement earlier in the process allows lived experience and local knowledge to influence priorities, design and evaluation. Institutions should consider how participation is supported, how responsibilities are shared, and how communities can question or reject proposed uses.

The seventh paper, "Co-creating SDG-informed Open Education with Trusted AI Across the Asia-Europe Axis," brings these concerns into open education. Its connection between inclusion, trusted AI and the Sustainable Development Goals places educational access within a wider international responsibility. Open materials can support participation across institutions and regions, but access also depends on language, accessibility, support and the ability to judge the material. The strategic issue is how to sustain knowledge infrastructure that people can use meaningfully and that institutions can maintain responsibly.

Together, these papers suggest that common good must be made visible in the design of cooperation. Leaders can ask whether a partnership expands the capability of less-resourced participants, whether underserved communities influence the definition of success, and whether educational resources remain usable beyond the initial project. These questions have practical implications about budgets, governance and relationships.

Learning and capability across a changing higher education system

The third paper, "The Need for a Contextualised Competence Framework for GenAI by Target Groups in Higher Education," addresses the different responsibilities within a university. Students, academics, administrators and leaders need a shared foundation in AI literacy, but they face different decisions and risks. A student must understand the limits of assistance in an assessed task; a teacher must determine what evidence demonstrates learning; an administrator must protect information; a leader must judge institutional dependence and accountability. Training should reflect these responsibilities and the context in which they are exercised.

The fifth paper, "Humanising Learning in the Age of AI: Policy Conditions for Pedagogy, Assessment and Learner Formation," gives this capability agenda an educational

foundation. It proposes AI as an amplifier rather than a substitute, formation before performance, governance proportionate to stakes, and a human-centred foundation that can be discussed across cultures.

Education produces human capability through activity, feedback, revision and reflection. AI may make some of those activities more accessible and offer useful support. It may also remove the effort through which a learner develops understanding. The task for teachers is to distinguish avoidable friction from intellectually necessary engagement. Assessment should make the learner's contribution and improvement visible, with opportunities to explain reasoning, evaluate evidence and demonstrate relevant independent capability.

The fourth paper, "What policy reforms are needed to enable universities to support equitable lifelong learning and reskilling through AI?" extends that responsibility beyond the conventional degree. Its roadmap considers universities as participants in flexible lifelong learning ecosystems. Changing tasks and careers increase the need for routes that people can enter at different stages of life. The public value of these routes depends on whether they are accessible, recognised, and connected to meaningful learning opportunities, including for people whose time, finances or prior qualifications restrict participation.

The reflections accompanying this collection identify the impact of AI on reducing the capacity of people to influence their "job reality" and on enhancing their talent. These are useful lenses for leadership of higher education institutions. External technological decisions can weaken people's sense of agency over their work, while assisted performance can obscure the capability that remains when the assistance is unavailable or unreliable. Universities should help learners and professionals understand changes at the level of tasks, retain responsibility for judgment, and develop the capacity to question technological recommendations.

A practical response combines curriculum renewal, support for educators, and dialogue with employers about real tasks. It also requires foresight about capabilities society will need beyond today's vacancies. Universities can connect academic and vocational routes, test flexible credentials, and design learning around challenges that require action. The outcome to seek is a person who can use assistance well, recognise its limits, and continue learning as circumstances change.

Institutional coherence and strategic alternatives

The sixth paper, "Cross-Regional Generative AI Governance in Higher Education: From Reactive Policies to Coherent Institutional Frameworks," examines the gap between having policies and possessing a functioning governance system. Separate rules for teaching, assessment, research and administration can create uncertainty when they rest on different assumptions or assign unclear responsibilities. A coherent approach links those domains, gives people understandable guidance, and creates a route for reviewing experience. The paper describes a pathway that offers institutions a way to discuss the present position of the institutions and possible next steps.

Readiness consequently involves more than publication of a policy. Leaders need to know whether staff and students understand the guidance, whether the institution can support its implementation and whether difficult cases have an accountable route for resolution.

The eighth paper, "Universities as Catalysts for Aligning AI Innovation with the Sustainable Development Goals," connects governance to the direction of innovation. Its case-based approach examines how universities can translate broad commitments into choices about research, partnerships and strategy. Referring to a development goal does not in itself demonstrate a contribution. Institutions need to examine the mechanism through which an activity is expected to create benefit, the people involved and the possible costs.

The ninth paper, "A Framework for University Leaders to Choose Strategic AI Policy Postures," supplies a vocabulary for such choices: Bridge, Sandbox, Accelerator and Futures Architect/Policy Advocate. A Bridge connects actors and knowledge that would otherwise remain separated. A Sandbox supports bounded experimentation and learning about risks. An Accelerator helps promising work progress towards implementation. A

Futures Architect or Policy Advocate contributes foresight and helps shape the wider conditions of responsible innovation.

These roles should inform decisions rather than become labels of prestige. A university with strong local relationships may contribute most through connecting communities and researchers. Another may have the infrastructure and expertise to support rigorous experimentation. Institutions may combine roles, but a credible choice needs resources, accountable leadership, and a way to evaluate its contribution.

Pathways for global national and institutional leaders

The collection offers a global view through dialogue across regions. Its recommendations are best read through three connected questions: What is the distinctive strategic insight of each paper and the critical tension it offers? Which actionable pathways are available to the reader's institution, given the actors, resources and constraints? Which collaboration models in the reader's context make those pathways feasible and accountable? This reading allows leaders to move from general aspirations to decisions without assuming that every setting requires the same sequence or institutional model.

For global leaders and international networks, a central task is to widen participation in the development and governance of AI. Cooperation can support open knowledge infrastructure, shared learning, and capability development across unequal settings. It should also create opportunities for institutions and communities with limited resources to influence priorities.

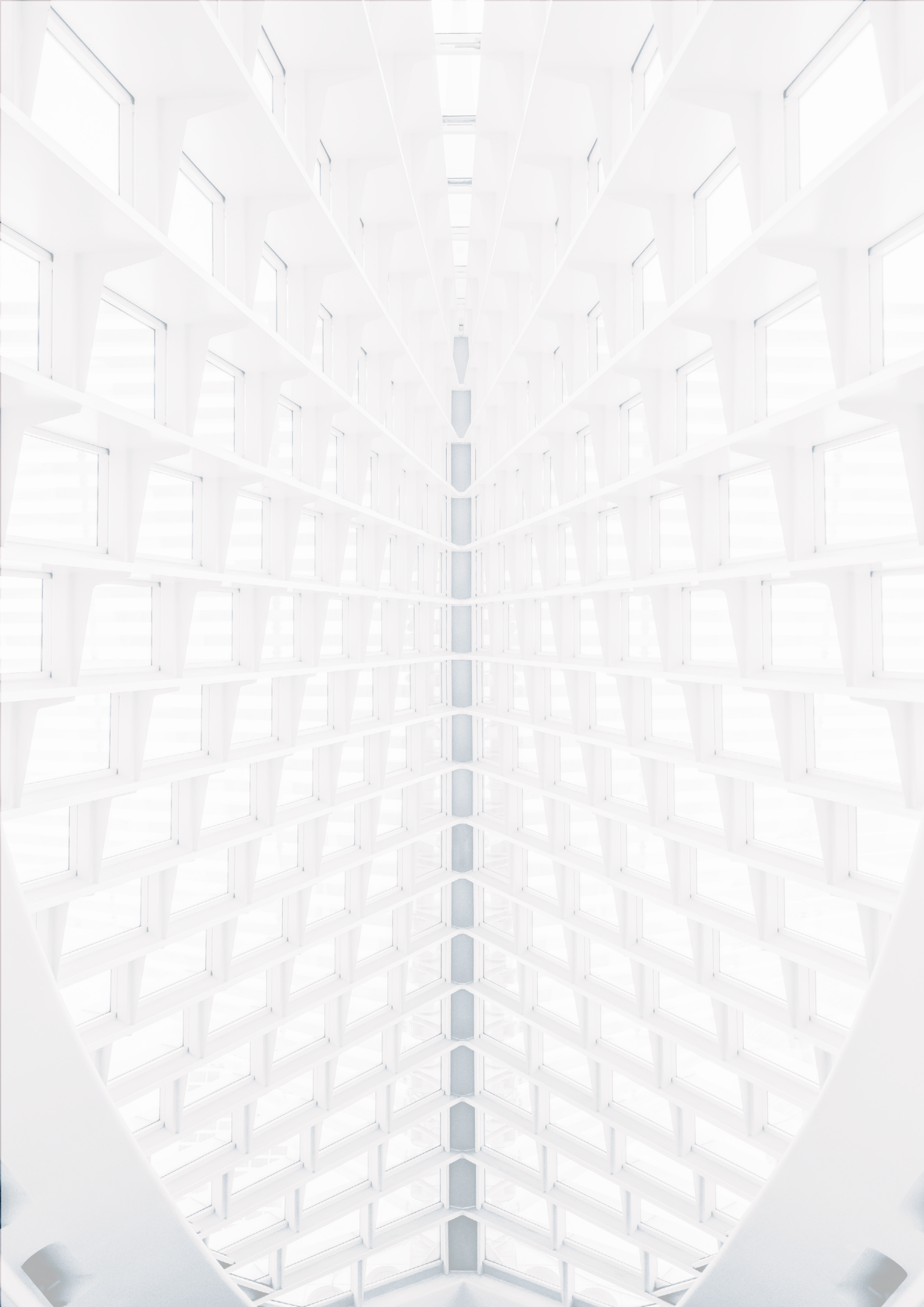
For national leaders, the papers invite a coordinated view of higher education, innovation and human capital development. Policy can help connect lifelong learning with recognition and access, support educator development, and address the conditions that sustain institutional capacity. It can also create space for responsible experimentation while maintaining clear accountability. Universities need room to adapt implementation to their missions, alongside public expectations that protect inclusion and educational quality.

For institutional leaders, an initial pathway is to examine actual practice. During the first 90 days, for example, a university can audit significant AI uses, risks, assessment practices and gaps between policy and experience. By six months, it can support teacher pilots and discussions with staff and employers about changing tasks. After a year, it can evaluate learning, inclusion and implementation difficulties, and use that evidence to revise its guidance.

The nine papers ask universities to act as catalysts rather than passive recipients of technological change. Their contribution is directional: to connect AI capability with public value through learning, innovation, governance and inclusion. This requires humility about the existing challenges and confidence in the responsibilities institutions can exercise.

We hope that this collection will inspire initiatives of "co-creation" between the readers' institutions and communities. This is one of our aspirations: that this collection becomes a resource for that continuing work, with the hope that its strategic insights will support informed co-created choices and its proposed pathways will strengthen willingness to pursue the common good through our universities.





Bridging the AI Divide in Higher Education through Asia–Europe Cooperation

AUTHORS
Shen Yi
Zhang Shuyan
Robert Kerwin Billones
Lafifa Jamal
Rachel Tan
Ramneek Athwal
Yang Xiuzhen

This paper helps policymakers, legislators, and higher education administrators close the structural AI capability gap in higher education across Asia and Europe by redirecting their attention toward equal and inclusive institutional cooperation.

Executive summary

Higher education institutions (HEIs) across Asia and Europe are caught in a striking paradox: artificial intelligence (AI) tool adoption exceeds 90% in both regions, yet the ability to govern, create, and equitably benefit from AI remains profoundly unequal. The AI divide has fundamentally shifted, moving beyond questions of who has access to AI towards questions of who has the institutional capacity to shape it. Unless confronted, this asymmetry risks creating a lasting global hierarchy, separating AI users from its governors. This paper identifies four interlocking dimensions of the divide:

- **A governance gap**, as evidenced by the fact that only 19% of HEIs have formal AI policies;
- **A literacy and talent gap**, where tool adoption outpacing skill readiness, driving talent migration from under-resourced to better-resourced institutions across Asia and Europe;
- **An infrastructure gap**, which is defined not by connectivity but by asymmetric capacity within shared computational environments like the European High Performance Computing Joint Undertaking (EuroHPC JU) and China's National Supercomputing Internet; and
- **A social and cultural gap**, which cover issues such as measurable performance deficits for Bengali-language users, a 13.93% cultural bias rate in leading AI models against Chinese contexts, women comprising just 22% of the AI workforce, and the persistence of data colonialism in training datasets.

The paper diagnoses four structural failure modes: implementation asymmetry, supply-side bias, self-reinforcing resource concentration, and a competitiveness framing that systematically marginalises equity. It argues for a strategic reorientation grounded in four principles: from policy supply to institutional autonomy; from infrastructure provision to capability equalisation; from adoption to institutional good use; and from competitiveness to equitable and reciprocal participation.

To operationalise these principles, the paper proposes five cross-regional mechanisms: an Asia–Europe higher education AI capability compact; a shared AI compute and data commons; equity-weighted eligibility criteria for research funding; a multilingual AI literacy and corpus fund; and a cross-regional AI divide monitoring observatory. Together, these mechanisms constitute a cooperative framework capable of replacing the current asymmetric model with one that ensures both regions develop the genuine institutional capacity to shape and govern AI on their own terms. By strengthening equitable institutional capability, reciprocal participation, and inclusive governance, the proposed framework positions Asia–Europe cooperation as a practical pathway for advancing AI for the common good.

1 Problem framing

AI tools have spread widely. What's Next?

More than 90% of higher education institutions (HEIs) across Asia and Europe already use artificial intelligence (AI) tools. We have won the access battle. But while governments celebrate rising adoption rates, a deeper and more consequential divide is widening in plain sight: the capacity to govern, create, and democratically benefit from AI remains profoundly unequal. We are measuring the wrong metric, and the cost of that mistake is a structural transfer of AI power from the Global South to the Global North, from regional universities to elite institutions, and from public interest to private capture.

1.1 What is the AI divide in higher education?

The *artificial intelligence divide* (AI divide) extends and deepens the concept of the *digital divide*. While the early digital divide primarily concerned inequalities of physical access, internet connectivity, and device availability,[1,2] the AI divide points to a more complex, multi-layered inequality of capabilities. It encompasses a full spectrum of disparities: from access to AI tools and infrastructure, to the cultivation of AI skills and literacy, the diversity of AI data and language resources, participation in AI innovation ecosystems, and the right to shape AI governance rules.

Even as access to AI tools becomes increasingly democratised with cloud services and open-source models dramatically lowering cost, the ability to translate these technologies into economic value, scientific innovation, and social impact remains highly uneven. This unevenness represents a central policy challenge of our time.

In the higher education context, the complexity of the AI divide is further highlighted. Higher education plays a triple role in AI knowledge production, talent formation, and governance legitimacy. It also serves as an institutional bridge between the development of AI technology and its social consequences.[3] When HEIs are unable to effectively manage AI, the democratic deficit in AI governance will expand. According to UNESCO's 2025 global survey of HEIs across 90 countries, 90% of respondents reported using AI tools in their professional work,[4] with Asia (93%) and Europe (91%) leading adoption rates. The AI divide in higher education today is therefore not primarily about access but about the structural conditions that determine who governs AI, who benefits from it, and who is left behind.

The working definition.

AI divide in higher education refers to "the structural inequality in the ability of HEIs, industry, government, and society to create, adopt, govern, and equitably benefit from AI."



1.2 Why pursue Asia–Europe cooperation? Why now?

Asia and Europe combine substantial AI activity, with considerable variation in institutional capability. This creates both a shared challenge and an opportunity for reciprocal cooperation. Both regions bring diverse higher education systems, resources, and areas of expertise. Some institutions have advanced AI infrastructure and research capacity, while others are still building foundational capabilities. This creates opportunities to share infrastructure, expertise, educational practices, and contextual knowledge. Existing Asia–Europe university networks, research partnerships, and policy platforms provide a foundation for such cooperation. Timing is also important. AI standards, datasets, and governance approaches are developing rapidly. Institutions need opportunities not only to adopt emerging AI systems but also to contribute to their development and governance. Asia–Europe cooperation can provide a practical platform for strengthening institutional AI capability through shared resources, reciprocal learning, and support that recognises different starting points.

1.3 Positioning and distinctive gap

Between 2022 and 2025, the number of national AI strategies more than doubled, from about 30 to around 70, across Asia and Europe.[5] These four years represent the most intensive wave of AI-related higher education policymaking on record.[6] However, in existing statistics and research, the implementation and translation of relevant policies by HEIs are often overlooked. This paper fills a critical gap by shifting the analytical lens to the institutional level, focusing on the universities, colleges, and polytechnics where most students receive their education and where the implementation and effects of AI divide interventions are ultimately realised. It argues that the dominant competitiveness framing of AI policy can contribute to the AI divide, and proposes a reorientation towards equity, capability equalisation, and cross-regional reciprocity. This paper positions institutional capability as an important link between AI policy, investment, and equitable participation.

This paper develops an integrated framework for understanding and addressing the AI divide in higher education across Asia and Europe. It first identifies four dimensions of the divide — governance maturity, literacy and talent, infrastructure, and social and cultural inclusion — and then diagnoses four structural failure modes that limit the effectiveness of existing responses. Building on this analysis, the framework connects these challenges to key policy trade-offs, strategic principles, cross-regional mechanisms, and five pilot projects, thus providing a structured pathway from problem diagnosis to practical institutional action.



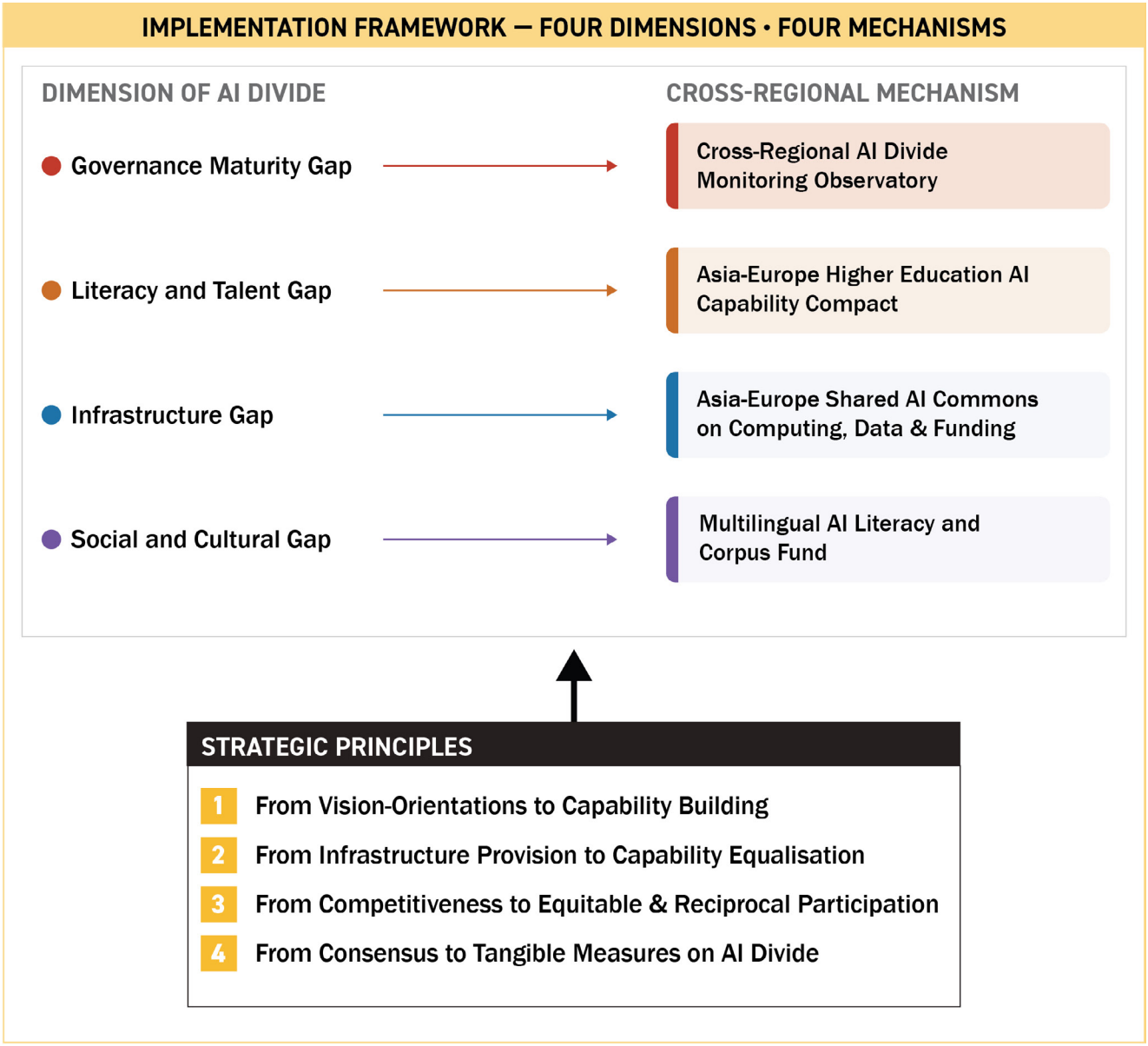


Figure 1. *A framework on bridging the AI divide in higher education through Asia–Europe cooperation: Structural failure modes, policy trade-offs, implementation framework, and pilot projects.*

2 Key insights

2.1 Four dimensions of the AI divide in higher education

2.1.1 The governance maturity gap

Although nearly 70 countries have adopted national AI strategies and policies,[7] HEIs in Asia and Europe have yet to translate national policies into feasible institutional practices. UNESCO's 2025 survey of 400 institutions across 90 countries found that only 19% had a formal AI policy, 42% were still developing one, and 39% had neither.[8]

A deeper governance failure concerns representation: as multinational technology companies gain increasing influence over global AI governance, HEIs, particularly those from the Global South, remain significantly underrepresented in the forums where AI norms, usage rules, and cross-regulatory frameworks are developed, reinforcing existing asymmetries in global governance.

2.1.2 the literacy and talent gap

The pace of AI adoption across industries significantly outpaces institutions' capacity to develop genuine AI capability. A 2026 UK survey found that 94% of undergraduates used generative AI (GenAI) for assessed work, while only 48% felt their teaching staff were helping them develop AI skills for future careers—a stark gap between tool adoption and institutionally supported

capability formation.[9] A Japanese study shows a similar pattern: 84.7% of first-year medical students had used GenAI, but only 49.2% had formal learning.[10]

High-skilled and AI talent mobility further complicates capability development. Outward mobility from countries including Bangladesh, the Philippines, and Malaysia, alongside strong demand in higher-resourced destinations, may make it difficult for countries to retain experienced AI talent.[11,12]

2.1.3 The infrastructure gap

Infrastructure inequality has evolved from disparities in institutions' internet connectivity towards differences in institutions' internal capacity to access, operate, and integrate advanced shared computing environments. In Europe, EuroHPC has 19 AI Factories providing access to supercomputing resources for all European scientists.[13] AI capability depends not only on skills but also on access to computational infrastructure. EuroHPC's AI Factories illustrate this institutional approach, providing researchers, industry, and public-sector users with shared access to advanced supercomputing resources. In China, the National Supercomputing Internet connected over 30 centres across 14 provinces and surpassed one million users by January 2026.[14] AI capability depends not only on skills but also on access to digital infrastructure and resources, reflecting the multidimensional nature of the digital divide.[15]

The infrastructure divide is thus shifting from physical access to national computing, to institutional capacity to convert shared infrastructure into research and educational value. This transition risks creating a new tier of inequality, one defined not by connectivity but by the ability to translate computing into impact.

2.1.4 The social and cultural gap

Three sub-dimensions define this gap. First, linguistic and cultural inequality. AI systems optimised for English-language data systematically disadvantage non-English speakers. Research on Bengali-language AI performance reveals consistent gaps: smaller models produce substantially higher token counts per Bengali instance, degrading accuracy and increasing computational overhead for over 170 million speakers.[16] The Proceedings of the 40th AAAI Conference on Artificial Intelligence introduced the GeWu benchmark which found a 13.93% cultural bias rate against Chinese contexts across every tested model, demonstrating that Chinese-language capability alone does not eliminate culturally specific bias.[17]

Second, gender and socioeconomic stratification. Women account for only 22% of global AI professionals. In Bangladesh, Malaysia, and the Philippines, socioeconomic inequalities continue to shape students' access to digital devices, reliable connectivity, and other resources needed for meaningful participation in technology-enabled learning.[18]

Third, data sovereignty. Who controls the data used to train AI systems, and who can audit or contest those datasets? HEIs across Asia have, in many cases, produced research and pedagogical datasets incorporated into commercial AI models without attribution or benefit-sharing. This is a form of data colonialism[19] that extracts value from populations with limited power to define the terms of that extraction.

2.2 Four structural failure modes

Mapping national and regional practices against the four dimensions reveals why the proliferation of initiatives has not closed the divide:

- **Implementation asymmetry.** National guidelines are cheap to produce; institutionalisation requires sustained investment in personnel, training, and organisational routines.[20] Malaysia illustrates this gap: its Digitalisation of Higher Education Action Plan 2025–2030 sets national priorities and phased initiatives, but implementation depends on institutional capacity.[21]
- **Self-reinforcing resource concentration.** Networked infrastructure such as EuroHPC Antennas and China's National Supercomputing Internet does not automatically equalise institutional capacity.[22] Shared access does not guarantee the skills, support, and organisational capacity needed for meaningful use. Universities still differ in technical staff, compliance capacity, project preparation, and ability to integrate advanced computing into research and teaching.

- **Supply-side bias.** Governments build computing centres, establish AI programmes, and draft curriculum frameworks, but provision does not automatically translate into capability. The UK provides a striking illustration: 94% of undergraduates reported using generative AI to help with assessed work, yet only 48% felt that teaching staff were helping them develop AI skills for their future careers.[23]
- **Competitiveness framing.** When AI policy is designed to win industrial races, universities become talent pipelines rather than sites of democratic capability formation. This explains both the neglect of the language and cultural divide (as there are no immediate competitive returns in addressing it) and the absence of equity-oriented monitoring. The result is a policy environment that prioritises speed and scale over inclusion and reciprocity.

3 Policy trade-offs

The core tension.

Should AI policy in higher education be framed primarily as a tool of national competitiveness, or as a public interest obligation? This is not a rhetorical distinction. It determines who receives funding, which institutions are prioritised, what gets measured, and ultimately who shapes AI governance. Both directions carry real consequences.[24,25]

3.1 Competitiveness first framing

AI as a national competitive asset.

Concentrate investment in frontier institutions and high-potential sectors. Measure success by rankings, patents, and economic output.

Consequences if this direction is chosen:

- Resources may flow disproportionately to research-intensive and already well-resourced universities, widening intra-national and inter-regional gaps.
- Cross-border AI talent mobility is increasing. Graduates from lower-resourced economies increasingly pursue opportunities in higher-resourced economies, underscoring the need for policies that convert mobility into knowledge exchange, talent retention, circulation, return pathways, and stronger domestic AI capacity.[26]
- Language and cultural minorities risk systematic marginalisation. Languages such as Bengali remain comparatively under-resourced in NLP, while culturally specific datasets and benchmarks receive less attention than resources serving dominant languages and markets.[27,28]
- The Global South risks increasing dependence on AI systems whose underlying models, infrastructures and data architectures are developed elsewhere, raising concerns about asymmetrical control over data and knowledge.[29]
- Short-term national gains in rankings and GDP are achievable but produce fragile, externally dependent AI ecosystems.
- AI governance reflects only the values and priorities of dominant economies and elite institutions.

3.2 Equity and public interest framing

AI as a public interest obligation.

Distribute capability across institutions and communities. Measure success by equity of access, institutional capacity, and democratic representation in AI governance.

Consequences if this direction is chosen:

- Equity-weighted funding reaches teaching-focused, regional, and resource-constrained institutions.
- Talent retention may strengthen as local institutions develop genuine AI capability, reducing the incentive to emigrate.
- Multilingual and culturally diverse AI systems can broaden representation and improve the contextual relevance of AI across diverse communities.

- Data sovereignty can be strengthened when institutions have meaningful capacity to contribute to, govern, and audit the datasets and data infrastructures that shape AI.
- AI governance can incorporate a broader range of institutional, linguistic, and cultural perspectives through more inclusive participation.[30]
- Potentially slower short-term competitive gains, while prioritising broader institutional capability, resilience, and legitimacy.

3.3 The compounding risk of inaction

Maintaining the status quo through the selective governance pattern identified in this paper produces the worst of both approaches. Without deliberate equity-oriented mechanisms, existing mechanisms (such as the EuroHPC JU, national AI funds, and curriculum reforms) may continue to produce uneven institutional benefits, particularly where access is not accompanied by the skills, organisational capacity and support required for meaningful use. The window to shape AI governance norms is narrowing: standards are being established and training corpora developed now, while institutions that lack capacity risk inheriting systems they had little role in shaping and limited ability to contest.

Strategic implication.

The cost of the AI divide is not only unequal access today, but reduced agency over how AI systems and governance norms are shaped tomorrow. AI for the common good must therefore address both access and agency.

4 Recommendations

The following four principles define the strategic reorientation required and anchor five cross-regional mechanisms proposed in Section 6.

4.1 From vision-orientations to capability building

AI policymakers and institutional regulators need to simultaneously shift their mindset, focusing from top-down macro policymaking to the specific implementation processes of institutions. The diffusion of governance measures at the national level is due precisely to their low production costs. Institutionalisation requires continuous investment in personnel, training, and organisational practices.[31] Asia–Europe cooperation should enhance institutional governance, leadership, and personnel capacity building, and form a mechanism for mutual learning and shared development.

4.2 From infrastructure provision to capability equalisation

AI infrastructure cooperation should pursue capability equalisation by strengthening the capacity of institutions to access, operate, integrate, and benefit from AI resources. Shared infrastructure must be accompanied by technical onboarding, training, data management support, cybersecurity guidance, and institutional integration support.[32] Attention should be directed to teaching-focused, regional, and resource-constrained institutions, which educate the majority of students in every country examined.

4.3 From competitiveness to equitable and reciprocal participation

AI strategies should position equity, language, culture, and participation alongside technological development and competitiveness, rather than subordinating them to these objectives. This means creating conditions in which institutions across different levels of capacity can participate, contribute, learn, and benefit. Capacity differences should inform the distribution of support and resources; participating institutions should simultaneously have opportunities to contribute knowledge, expertise, educational practices, data, and other resources to the wider cooperation ecosystem.

4.4 From consensus to tangible measures on AI divide

At present, Asian and European countries show broad recognition of the AI divide but frame its priorities differently. The Philippines emphasises AI talent, workforce preparation, and economic development; Malaysia combines inclusive AI adoption with infrastructure, talent, governance, and national capability development; while Bangladesh places particular emphasis on linguistic and cultural representation in AI systems.[33,34,35]

5 Implementation pathways

The following five mechanisms translate the four principles into coordinated cross-regional action. They are designed to be complementary; each addresses a specific dimension of the AI divide and sequenced to build institutional momentum toward capability equalisation.

5.1 Cross-regional AI divide monitoring observatory

Governments, higher education networks, and international partners should establish a cross-regional AI divide monitoring observatory to obtain comparable evidence on the policy priorities of institutions in Asia and Europe. The observatory collects data at the institutional level from four aspects: institutional governance (AI policies, governance structures, and ethical safeguards), talent literacy (AI literacy, teaching methods supported by AI, assessment redesign), infrastructure (algorithms, data, computing power centres, energy, network connections), and social culture (language, gender, cultural diversity). The observation station will shift its monitoring focus from national adoption indicators to institutional policy demands, thereby enabling policy, financial, and cross-regional learning and cooperation based on mutual understanding.

Proposed commitment.

Pilot a common institutional AI capability framework across a small group of Asia–Europe partner HEIs and openly publish baseline data and refine indicators before wider implementation.

5.2 Asia–Europe higher education AI capability compact

To enhance the initiative of institutions in AI governance, a shared mechanism should be established among universities in Asia and Europe to monitor the implementation progress of AI governance in each specific institution. The transformation of AI should be evaluated through measurable institutional capabilities, including governance, employee expertise, teaching integration, student support, and responsible implementation, rather than through the adoption of agents such as tool procurement or project initiation. ASEF, participating governments, higher education networks, and relevant partners should establish an Asia–Europe higher education AI capability compact as a shared framework for the development of institutional AI capabilities. The compact would establish common reference points across six domains: institutional AI governance, staff and leadership capability, AI-integrated curriculum and assessment, student AI literacy and support, responsible AI practices, and institutional implementation capacity.

Proposed commitment.

Establish institutional AI implementation plans and report progress against common capability milestones.

5.3 Asia–Europe shared AI commons on computing, data, and funding

Higher education partners in Asia and Europe should establish shared AI computing and data-sharing infrastructure, providing fair access to AI investment, computing resources, datasets, models, and technical support. Public resources, based on existing national and regional programs, serve as cross-regional access and support portals rather than a repetitive system. Existing and future higher education funding schemes related to AI should incorporate weighted eligibility criteria to account for differences in institutions' start-up capabilities. Priority should be given to institutions that focus on teaching, are regional, and have limited resources. The core objective is capability balance: enabling institutions with different initial capabilities to access, operate, and integrate shared AI infrastructure and investments, and benefit from them.

Proposed commitment.

Establish a pilot Asia–Europe AI commons providing shared access to AI compute, data, models, and technical support, with equity criteria prioritising resource-constrained HEIs and transparent reporting of funding distribution.

5.4 Multilingual AI literacy and corpus fund

Rather than acting independently across universities, it is more important to establish an inclusive framework so that local cultural resources are equally represented

in the AI development process (Recommendation 4.3). Asia–Europe cooperation should establish a multilingual AI literacy and corpus fund to strengthen AI capability across different linguistic and cultural contexts. The fund supports collaborative development of non-dominant-language training corpora, AI evaluation benchmarks, translated and locally adapted AI literacy curricula, multilingual teaching and learning resources, and culturally responsive AI resources.

The linguistic and cultural dimensions of the AI divide are among the most neglected, yet they have significant implications for equity. Bangladesh has directly addressed this through commitment to a Bengali-based LLM. The fund addresses this structural gap at scale, treating corpus construction and culturally grounded evaluation benchmarks as questions of educational justice, not just technical infrastructure.

Proposed commitment.

Launch the fund with initial contributions from at least two Asia–Europe partner governments and prioritise non-English language corpora and open-licensed AI literacy curricula.

5.5 Pilot projects: From principle to practice

Five institutional pilots operationalise the mechanisms above and generate the evidence base for future cooperation:

- Joint Asia–Europe AI literacy benchmarking pilot — co-design a shared AI literacy measurement instrument across ten institutions (five Asian, five European), providing sufficient institutional diversity for a meaningful cross-regional baseline of staff capability, student AI readiness, and institutional support provision.
- Co-designing and open sharing of AI literacy resources — collaborate with industry to develop at least one adaptable AI literacy module per national partnership, released under an open license to enable adoption by smaller and regional institutions.
- Asia–Europe faculty and institutional leadership exchange — engage systematic professional development for academic staff, institutional leaders, and professional services personnel, with particular attention to teaching-focused and resource-constrained institutions. Participants jointly produce practical outputs: assessment approaches, governance processes, and AI literacy programmes.
- Collaborative Asia–Europe AI data and corpus pilot — contribute to a shared repository of AI-related datasets, with priority for non-English language datasets, underrepresented regional contexts, and culturally grounded evaluation resources. Governance arrangements cover attribution, benefit-sharing, access, and audit rights.
- Participatory institutional AI governance pilot — establish structured mechanisms through which students, academic staff, and professional services staff contribute to decisions on AI use, policy, assessment practices, and institutional procurement. This extends equitable participation from access to AI resources to participation in AI governance.



A final note for policymakers

The AI divide examined in this paper is not inevitable. It is the cumulative result of policy choices about what to measure, who to fund, and what framing to use. Every time AI policy is evaluated by adoption rates rather than capability equity, the divide widens. Every time infrastructure investment bypasses regional and teaching-focused institutions, the divide deepens. Every time a governance process excludes HEIs from the Global South, the divide becomes more entrenched.

The mechanisms proposed here are not aspirational. They build on existing infrastructure, institutions and research. What they require is a deliberate reorientation of purpose: from capability accumulation at the frontier toward capability equalisation across the system.

The bottom line.

Universities across Asia and Europe are not merely sites for the adoption and implementation of AI; they are also active contributors to its development and direction. Every governance framework produced, dataset or corpus developed, benchmark designed, educational practice introduced, and policy model tested within a university can help shape the kind of AI ecosystem the region will have. The ambition of this paper is to enable more universities across different languages, national contexts, institutional missions, and levels of capacity to take their place in shaping that future. Closing the AI divide begins with recognising that the diversity of Asia and Europe's HEIs is not a coordination challenge to overcome, but one of the region's most underutilised assets. In this sense, closing the AI divide is integral to advancing AI for the common good across Asia and Europe.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council [Grant Number EP/S023305/1] and the "AI Governance and Safety: Understanding the Impacts of AI and Enabling Partnerships" [Grant Number FC-2425-5-154] by the Department of Science, Innovation and Technology's International Science Partnerships Fund (ISPF) via the Royal Academy of Engineering under the Frontiers Champions - Tranche 5 scheme.

Authors

Prof SHEN Yi

Fudan University, China

Mr ZHANG Shuyan

Fudan University, China

Prof Robert Kerwin BILLONES

De La Salle University, Philippines

Prof Dr Lafifa JAMAL

University of Dhaka, Bangladesh

Dr Rachel TAN

INTI International University, Malaysia

Mrs Ramneek ATHWAL

University of Nottingham, United Kingdom

Miss YANG Xiuzhen

Fudan University, China



References

1. Van Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. *Poetics*, *34*(4–5), pp. 221–235.

2. Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4).

3. Weuts, R. et al. (2025). *White paper for universities on navigating artificial intelligence innovation ecosystems in the area of AI governance*. Asia-Europe Foundation. https://asef.org/publications/asefinnolab-white-paper-governance/

4. UNESCO. (2025, 2 September). *UNESCO survey: Two-thirds of higher education institutions have or are developing guidance on AI use*. https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use

5. OECD.AI. (2025). OECD.AI: OECD Artificial Intelligence Policy Observatory. https://oecd.ai/

6. UNESCO IESALC. (2025). *Shaping the future: Higher education's impact on national AI policies*. https://www.iesalc.unesco.org/en/articles/shaping-future-higher-educations-impact-national-ai-policies

7. OECD.AI. (2025). OECD.AI: OECD Artificial Intelligence Policy Observatory. https://oecd.ai/

8. UNESCO. (2025, 2 September). *UNESCO survey: Two-thirds of higher education institutions have or are developing guidance on AI use*. https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use

9. Stephenson, R. and Armstrong, C. (2026). *Student generative AI survey 2026 (HEPI Report 199)*. Higher Education Policy Institute. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/

10. Tajima, H. et al. (2025). Perceptions and intentions to use generative artificial intelligence among first-year medical students in Japan: Cross-sectional survey study. *JMIR Medical Education*, *11*, Article e77552.

11. World Bank. (2025). *Digital progress and trends report 2025: Strengthening AI foundations*. https://openknowledge.worldbank.org/server/api/core/bitstreams/f2509a0f-7153-4f32-b180-bc11e90c4940/content

12. Haesler, C. et al. (2026). The global race for AI talent is far from over. *Boston Consulting Group*.

13. European Commission. (2026). *Shaping Europe's digital future – AI Factories*. https://digital-strategy.ec.europa.eu/en/policies/ai-factories

14. Supercomputing Internet. (2026, 4 January). *National Supercomputing Internet user base surpasses one million*. https://www.scnet.cn/home/news/123619

15. Lythreatis, S. et al. (2022). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, *175*, Article 121359.

16. Mahfuz, T. et al. (2025). Too late to train, too early to use? A study on necessity and viability of low-resource Bengali LLMs. *Proceedings of the 31st International Conference on Computational Linguistics*. pp. 1183–1200. Association for Computational Linguistics.

17. Lin, Y. et al. (2026). GeWu: A culturally-grounded Chinese benchmark for multi-stage social bias evaluation in large language models. *Proceedings of the AAAI Conference on Artificial Intelligence*, *40*(38), pp. 32033–32041.

18. UNESCO and EQUALS Skills Coalition. (2019). *I'd blush if I could: Closing gender divides in digital skills through education*. https://unesdoc.unesco.org/ark:/48223/pf0000367416?

19. Couldry, N. and Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press. https://doi.org/10.1515/9781503609754

20. Benayoune, A. et al. (2026). Artificial intelligence policy challenges and institutional readiness in Omani higher education. *Discover Education*, *5*, Article 86.

21. Ministry of Higher Education Malaysia. (2025). *Digitalisation of higher education action plan 2025–2030*. https://www.mohe.gov.my/en/download/publications-journals-and-reports/pelan-tindakan-pendigitalan-pendidikan-tinggi-2025-2030

22. Epstein, G. S. (2026). The distribution of scientific power in the age of AI. *Social Sciences & Humanities Open*, *14*, Article 103374.

23. Stephenson, R. and Armstrong, C. (2026). *Student generative AI survey 2026 (HEPI Report 199)*. Higher Education Policy Institute. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/

24. UNESCO. (2022). *Recommendation on the ethics of artificial intelligence*.

25. OECD. (2026). *OECD digital education outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. https://doi.org/10.1787/062a7394-en

26. World Bank. (2025). *Digital progress and trends report 2025: Strengthening AI foundations*. https://openknowledge.worldbank.org/server/api/core/bitstreams/f2509a0f-7153-4f32-b180-bc11e90c4940/content

27. Joshi, P. et al. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 6282–6293. Association for Computational Linguistics.

28. Blodgett, S. L. et al. (2020). Language (technology) is power: A critical survey of "bias" in NLP. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 5454–5476. Association for Computational Linguistics.

29. Couldry, N. and Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press. https://doi.org/10.1515/9781503609754

30. UNESCO. (2022). *Recommendation on the ethics of artificial intelligence*.

References

31. Benayoune, A. et al. (2026). Artificial intelligence policy challenges and institutional readiness in Omani higher education. *Discover Education*, *5*, Article 86.

32. Lythreatis, S. et al. (2022). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, *175*, Article 121359.

33. Department of Trade and Industry (DTI) Philippines. (2024). *National AI Strategy Roadmap 2.0: The strategy framework*. Center for AI Research. https://naisr.cair.ph/strategy-framework/

34. AI Malaysia Berhad. (2026). *National AI Action Plan 2026–2030*. https://strapi.naio.okie.my/uploads/AI_Nation_2030_National_AI_Action_Plan_2026_2030_EN_V_180826_ddf0221683.pdf

35. Government of Bangladesh. (2026). *National Artificial Intelligence Policy 2026–2030 (Draft V2.0)*. AI Policy for Bangladesh. https://aipolicy.gov.bd/docs/national-ai-policy-bangladesh-2026-2030-draft-v2.0.pdf





AI for Whose Good?

A Governance Framework for University Investment in AI with Underserved Communities

AUTHORS

Yang Bong
Neelima Sailaja
Minjia Chen
Devisakti Annamali
Jin Tada
Leïla Cellérier
Muhammad Zahid Iqbal
Vicente Pitogo

Executive summary

Artificial intelligence (AI) is rapidly reshaping everyday life, but its benefits remain unevenly spread. Left to prevailing market and research incentives, AI is poised to deepen inequalities that extend beyond the digital divide of technology accessibility to a more consequential AI divide – in who shapes it, whose data and languages it represents, and whose definition of progress it encodes.

Universities are unusually well-placed to confront this divide and have strong reasons to do so now.[1,2] Their advantage lies in three core institutional missions: teaching, research, and community engagement. Through talent development, knowledge creation, and societal partnership, these missions give universities a pathway for co-creating AI with communities, translating innovation into public value, and sustaining the reciprocal partnerships through which both improve.

While no other actor combines all three at comparable scale, many universities have so far treated AI mainly as a question of teaching integrity, research capacity or operational efficiency, and few have made it part of their public mission. The reasons to act now are also the university's own. AI is already disrupting universities' roles as providers of education and research, and institutions that fail to adapt their mandates risk losing the public trust and support on which their long-term sustainability depends.

This white paper argues that universities should make explicit whose good their AI work is meant to serve, prioritising underserved communities that are most exposed to AI's harms and least represented in its design. It offers a framework for universities and communities to determine together what outcomes matter and how AI initiatives should be governed, resourced and evaluated. At its core, "good" should be defined not for these communities but with them, through the capabilities they themselves have reason to value. The paper speaks to universities at any stage, from those yet to direct AI towards public good to those already building AI for identifiable communities.

Drawing on development ethics, participatory design, and the capability approach, the paper proposes a 'University AI for Community Good Framework' to support universities considering or undertaking AI initiatives for public and community benefit. The framework helps university leaders align institutional strategy with social impact and inclusion, make AI research and development more responsive to those most affected, and guard against 'techno-solutionism', which is the assumption that deploying technology is sufficient to produce benefit.

At its centre is a three-part test that treats communities as receivers, doers and judges at once. An AI initiative serves a community's good only when its members gain freedoms they value, when it runs on terms they helped set, and when they decide whether it is worth pursuing and whether it has succeeded. The framework rests on four recommendations: let communities define "good"; reform incentives for community-engaged work; treat communities as knowledge co-creators; and sustain partnerships beyond the grant while connecting across regions.

1 Problem framing

AI is already reaching the communities universities serve, through commercial products, public services and AI platforms. Governments, companies and institutions have been investing rapidly in AI, with an increasing focus on social good, while many universities have yet to decide what role to play in shaping it. A central question remains: who decides whether these investments are "good" and whether they do any good? An AI initiative may be technically successful, ethically reviewed, and well-funded, yet still fail to serve the communities it intends to benefit if those communities have little influence over which problems are addressed, how solutions are designed, or what outcomes count as success.

Since the late 2010s, "AI for Good" has grown into a recognised field, propelled by the United Nations' flagship campaign and the Sustainable Development Goals, corporate initiatives and a growing body of scholarship.[3] The aspiration is much needed, but its framing of "good" has three limitations. First, "good" remains broad and underspecified. Research in the field presupposes a definition of "good" that is seldom stated,[4,5] and a generic and supply-driven notion of "good" licenses techno-solutionism, as illustrated through the One Laptop per Child programme.

Second, good is often defined by the researchers, companies, governments, NGOs or funders behind an initiative, and rarely by the communities they claim to serve, while the language of social good is increasingly attached to corporate and governmental initiatives and their priorities.[3] This biases effort toward incremental, fundable problems over efforts that require more structural reforms,[4] and can lend the language of social good to projects that entrench the very disadvantages they purport to address.[5,6]

Third, when whose good is not specified in AI endeavours, the label "good" can itself do harm. Projects that define their beneficiaries and their needs from the outside risk producing systems that, as scholars[5] described, make their proponents "feel good" while extending harm and consuming substantial resources. The result is wider inequality in AI innovation, and beyond it, epistemic exclusion where communities are absent both from participation and from the account of whose needs count in the "AI for Good" discourse.

The One Laptop per Child programme began at MIT's Media Lab as a centrally designed answer to children's education across developing countries and was adopted on a large scale. In one of its largest deployments costing roughly USD 180 million,[7] a randomised evaluation in Peru found no effect on children's test scores in mathematics or language, despite a sharp rise in their access to computers.[8] The programme continues to serve as a reminder that technology initiatives however well-intentioned, can fail to deliver the intended outcomes when the problem and solutions were defined without the affected communities.[9]

Recent work shows how to put these concerns into practice. Costanza-Chock[10] argues that inclusive intent is insufficient: design justice – a movement that treats the design process itself, and not only what it produces, as a site of justice – holds that the practices, norms and ownership structures of design must themselves be redistributed toward those bearing the burden of its harms. London and Heidari[11] apply Sen's capability approach, which judges progress by what people are actually free to be and do, marking a shift from utility-maximising evaluation toward an agency-driven, diversity-sensitive standard, and name the distinct ways a system can instead harm a user: paternalism, coercion, deception, exploitation, and domination. Bondi et al.'s[4] PACT framework – a Participatory Approach to enable Capabilities in communities – draws the participatory conclusion: "social good" cannot be specified in advance by researchers or funders, but must be worked out by the communities a system will affect, who identify which capabilities matter most to their flourishing. Cenci and Cawthorne[12] sharpen the procedural point by asking less how to design a technology than who decides what it is for.

The unresolved question at the heart of the field is therefore not whether AI can generate "good", but whose good it advances, whether AI is the right response at all, and who gets to decide.[3] The field's researcher- and funder-led framings of "good" need reframing if it is to produce outcomes that reach people on the ground. Universities are where much of that framing happens:

they define the problems, train the researchers, and hold the data partnerships, so they carry both the obligation and the opportunity to change it.

2 Key insights

Three connected insights inform the recommendations that follow. First, AI for Good remains shaped predominantly by institutional definitions of what constitutes "good". Second, the capability approach and participatory design offer a stronger test for defining and evaluating good. Third, emerging institutional practices show how universities can redistribute authority towards underserved communities.

Definition: **Underserved Communities**

This paper treats "underserved" not as a fixed property of a group, geographical or sociological, but as a relation between a particular AI system and the people it affects. A community is underserved by a system when it bears that system's effects, when its knowledge, languages and priorities were absent from its definition and design, and when it has no standing to evaluate or redirect it. The designation is context-dependent, is confirmed by the community rather than assigned by the university, and assumes no homogeneity. It reaches beyond the digital divide of unequal access to technology: a connected community can still have no say.

2.1 The challenges of AI for Good

Three mechanisms explain why the benefits of AI for Good do not reach the communities most exposed to its harms, and these apply to any university entering the field. The first is power imbalance: what constitute "good" in AI projects are often defined by its funders or researchers rather than by affected communities.[5] The second is distributive inequality: initiatives optimised for well-connected, data-rich environments tend to leave out those at the margins.[6] The third and least examined is epistemic exclusion, where mainstream approaches can treat community knowledge, languages and data as resources that can be extracted, while leaving institutional assumptions about intelligence and progress unchallenged.[13]

In response, a participatory turn has sought to reposition affected communities from beneficiaries to co-designers.[4,14] Speculative co-design, for example, asks communities what they would build if existing technological assumptions and constraints were lifted. The resulting ideas are treated as evidence of what communities value, rather than simply as specifications for a predetermined product.[6] Participation, however, does not automatically redistribute authority. Communities may contribute ideas, data and labour after researchers and funders have already determined the problem, governance arrangements and criteria of success. The central question is therefore not only who participates, but who has authority to define and evaluate the good being pursued.

2.2 From participation to evaluative authority

Determining "whose good" first requires an account of what "good" is. The Receiver-Doer-Judge concept[15] provides one. Building from the capability approach,[16] it shifts evaluation from the resources deployed and the outputs produced to the substantive freedoms people have reason to value across the three dimensions that a person holds at once: a receiver of resources and support, a doer who acts to pursue their goals, and a judge who evaluates what is worth pursuing. Applying this to AI initiatives, this changes the guiding question from 'what problems can AI solve?' to 'whose freedoms is this system expanding, on whose terms, and at whose expense?'.

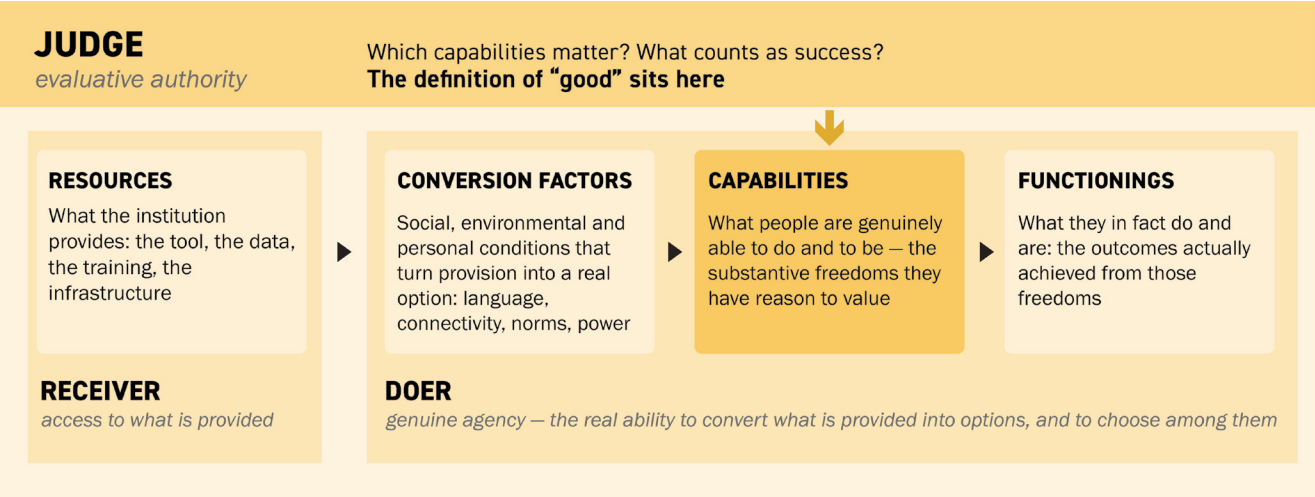


Figure 1. *The capability chain from resources to functionings, with the Receiver, Doer and Judge roles (after Bonvin and Laruffa[15])*

Applying these roles to communities, and not only to individuals, matters because the exclusion at issue is structural rather than individual: it resides in how problems are defined, whose data is used and on what terms, and which outcomes are permitted to count as success – conditions no individual can renegotiate alone. It matters, too, because collective capabilities are not simply the combined capabilities of individual members. They arise through collective action, institutions and social relationships, giving communities capacities that individuals could not exercise alone.[17] The judge role this paper calls for is collective in precisely that sense: not an aggregation of individual opinions, but a capacity exercised through structures.

Read against the three roles, AI for Good positions communities unevenly. The receiver role is over-emphasised as communities appear as a category of beneficiaries of systems designed elsewhere, for problems named elsewhere. The doer role is granted in part, as members contribute data, local knowledge and testing. The judge role is largely withheld, where the purpose of a system and the capabilities it targets remain those that the institution judged worthwhile. A university can therefore report extensive community participation in a project whose success the community had little say in judging.

This gives universities a three-part test of "good", one part for each role (see Section 4 and Figure 2). An AI initiative serves the good of a community when, as receivers, its members gain freedoms they have reason to value, not only access to a tool; when, as doers, the initiative runs on terms they helped set; and when, as judges, they decide whether it is worth pursuing, including whether AI is the right response at all, and whether it has succeeded. Because the judge role is the one most often withheld, it carries the most weight in what follows.

2.3 Emerging institutional practices

A growing number of universities are now conducting influential research on responsible and human-centred AI. Stanford Institute for Human-Centered Artificial Intelligence (HAI) produces an annual Index that tracks AI's societal impacts globally[18]. Oxford's Internet Institute produces governance frameworks for public-sector AI deployment[19], and MIT's work on responsible AI addresses algorithmic accountability and risk[20]. Yet guidance on how universities should conduct their own AI work with underserved communities remains underdeveloped. Research on community partnerships finds that effective partnerships share five features: respect for lived experience; engagement before key decisions are made; shared authority over goals and methods; transparency about data, processes and benefits; and long-term investment in trust.[21,22] To illustrate, three initiatives below show how these principles can operate in AI:

University of Nottingham and communities in Kerala, India. A multi-year collaboration positioned subsistence farmers as dataset co-creators rather than data sources. Using familiar practices and WhatsApp, farmers photographed and annotated fields across changing monsoon conditions to support a Visual Place Recognition system. The collaboration combined older participants' ecological knowledge with younger participants' digital fluency, while a locally rooted cooperative offered a potential structure for governing the resulting system.[23]

Auckland University of Technology, New Zealand. Shedlock et al.[24] show how Indigenous knowledge can shape the design process rather than merely supply data. Māori pōwhiri – the protocol for welcoming visitors – was used to structure the development of a large language model. The obligations and sequence of the encounter informed how the system was built, demonstrating how Indigenous knowledge and language can guide AI development from the outset and support digital sovereignty.

The Abundant Intelligences Programme. This six-year consortium brings together thirteen universities and eight community organisations. Its research Pods position Indigenous knowledge-holders, language-keepers and community organisations alongside scientists, engineers and humanists as co-investigators rather than informants. Communities therefore help determine what is built and which forms of knowledge justify building it.[13]

These initiatives differ in scale and context, spanning a European university working in South Asia, a university in Oceania and an international Indigenous-led consortium. However, they share three characteristics: communities contribute justificatory knowledge, which shapes why and how a system is built; their participation begins before the technology is defined; and partnerships are designed to extend beyond a single project cycle. Their structures, including a cooperative, a Māori protocol, and a standing international network, also provide a basis for retaining knowledge and governance after individual researchers or grants withdraw.

3 Policy trade-offs

A capability-centred approach is a reorganisation of value, power and resources, and it confronts universities with three interconnected trade-offs: research outputs versus research impact, institutional control versus community agency, and institutional investment versus long-term sustainability. The three form a chain, starting from what is measured, to who holds authority, to what must change to sustain the work.

3.1 Research outputs vs research impact

The first trade-off concerns what counts as success. Rapid deployment can produce publications, prototypes, partnerships and visible short-term impact. Indeed, current academic and funding structures often favour projects with short timelines and measurable outputs.[4] A capability-centred approach places greater emphasis on community agency, shared governance and long-term sustainability and distributes costs, benefits and responsibilities differently.

The deployment-first model can accelerate experimentation, yet it may position communities mainly as data providers, users or beneficiaries. Such arrangements can reinforce dominant worldviews, widen capability gaps, and concentrate value within well-resourced institutions. Evaluation is where this concentration is hardest to see and hardest to undo. While success is recorded only as publications, patents and prototypes, a project can satisfy every measure applied to it and still leave the community that hosted it no better able to act – and the institution holds no instrument that would tell it so.

3.2 Institutional control vs community agency

The second trade-off concerns legitimacy. Universities carry responsibility for the legitimacy they bring to AI initiatives. Their involvement can support sustained engagement with communities, validate technologies and strengthen public confidence. Where governance

remains concentrated among researchers and technology partners, this legitimacy may support extractive practices associated with data colonialism, the appropriation of communities' data and knowledge on terms set elsewhere and for benefits that flow elsewhere.[25]

Communities may contribute knowledge and data while receiving limited authority, compensation or long-term benefit, and the reputational risk rests with the university whose name is on the work. Capability-centred and participatory approaches redistribute this authority. Communities can shape research priorities, govern data and evaluate how AI affects their lives. They can also develop durable capabilities in AI literacy, data stewardship, knowledge curation and collective decision-making. Universities gain research grounded in context, stronger public legitimacy and relationships that extend beyond individual projects.

3.3 Institutional investment vs long-term sustainability

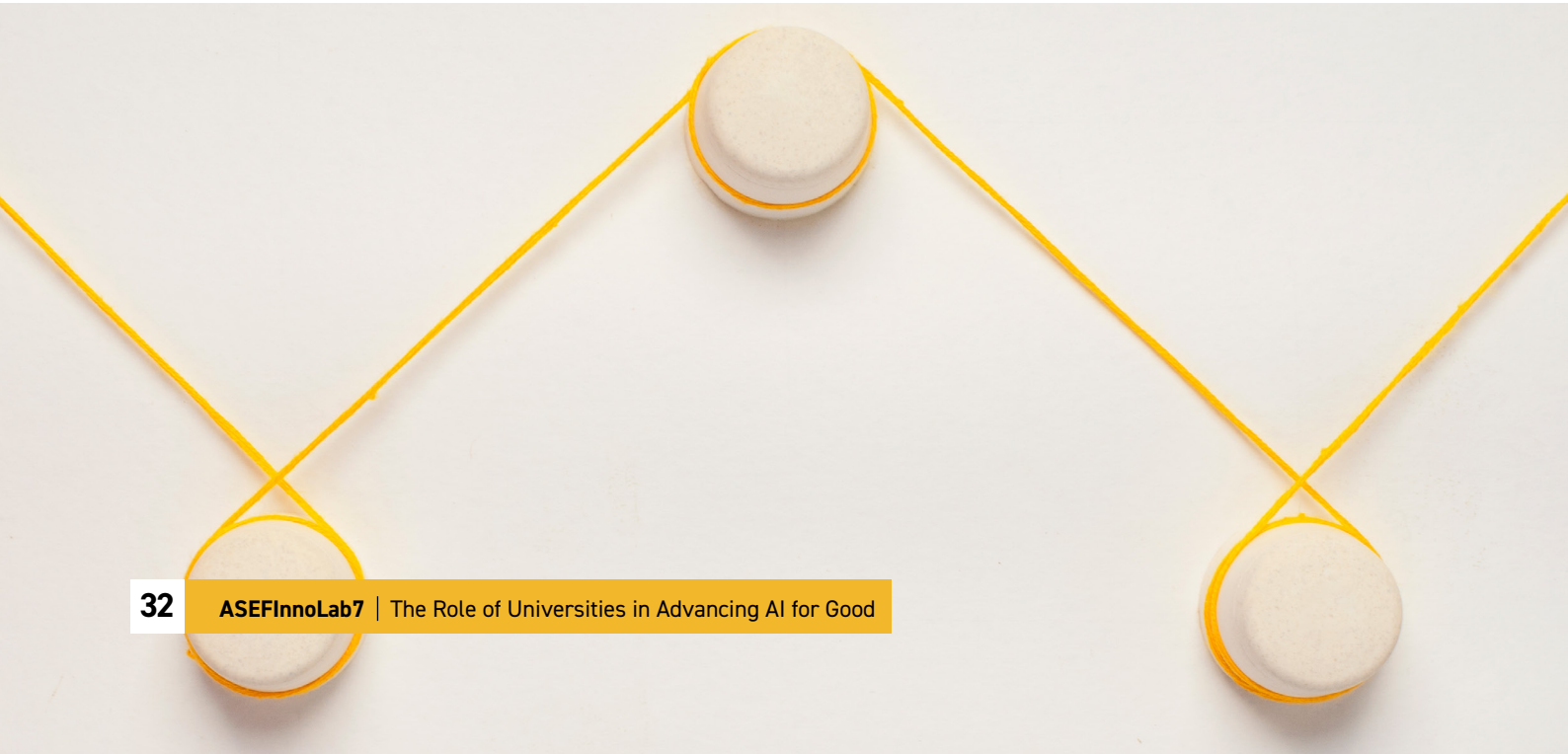
These gains require substantial institutional investment. Co-design, participatory evaluation and sustained partnership require additional time, funding and specialist expertise. Universities may need to support longer project cycles, fewer concurrent initiatives and a wider range of research outputs. Researchers may need to share control over project priorities, data access and intellectual property. Funders may need evaluation frameworks that recognise community capacity, trust and institutional change alongside publications, patents and technical outputs.

Sustainability creates a further trade-off. Short pilot projects can demonstrate immediate value, while long-term initiatives require continuing investment in governance, infrastructure and local expertise. Multi-year funding, open and interoperable technologies, community data agreements, and local maintenance capacity can strengthen continuity after external partners leave.

4 Recommendations

The paper argues that universities should define public good with communities and invest in AI to serve it. Implementing this means moving beyond isolated projects to a strategy that runs through every level of the institution. The recommendations apply to universities at any stage of the AI for Good spectrum, including those starting out, or those that are already active.

This section sets out the 'University AI for Community Good Framework' (Figure 2) that turns the three roles in Section 2.2 into institutional practice, guided by the four recommendations laid out below. Each names its audience, including university leadership at strategic level, academics and researchers at management level, community partnerships at operational level, and funders, local governments, and enterprises whose decisions shape each level.



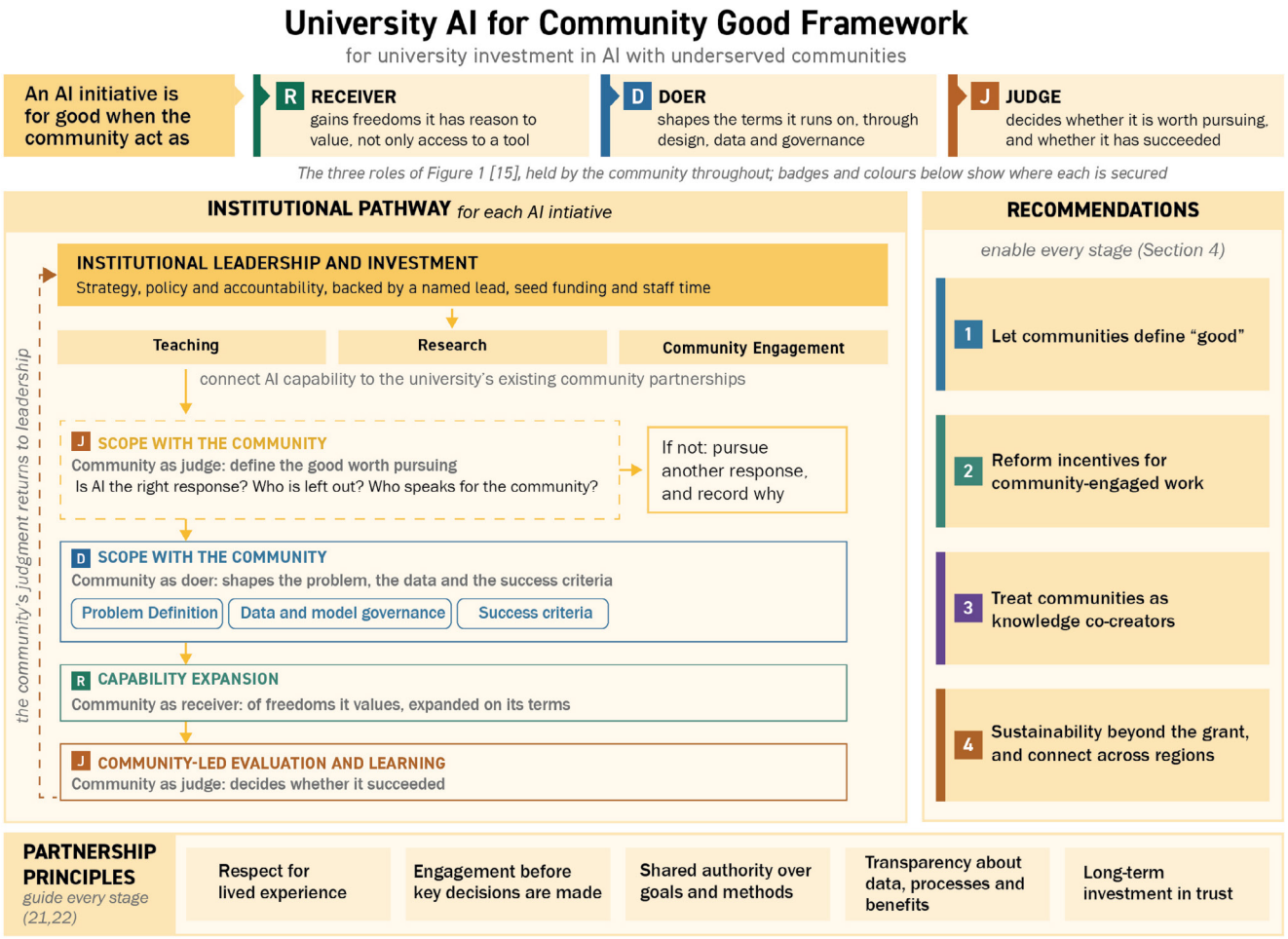


Figure 2. *The University AI for Community Good Framework. The top row sets the test of "good" by applying the Receiver, Doer and Judge roles of Figure 1 to the community. Leadership commits strategy and resources, and the three missions connect AI capability to existing community partnerships. Each initiative is scoped with the community, including whether AI is the right response, then moves through shared design and governance, capability expansion and community-led evaluation, which feeds back to leadership.*

4.1 Let communities define "good"

As laid out in Sections 1 and 2, "good" cannot be determined in advance by researchers or funders but must be defined by the affected communities. In practice, this means moving communities beyond the *receiver* role alone and into those of *doer and judge*. Communities should not merely be allowed to participate in AI projects. Instead, they should hold substantive authority in defining what "good" means for them, the outcomes pursued, the acceptable uses of their knowledge and data, and how benefits are evaluated.

Implementation Pathways

Audience: University leadership and governance bodies; community organisations; funders.

- Leadership should commit the institution to community-defined good through policy, governance, and accountability arrangements, and back that commitment with dedicated resources: a named lead, seed funding, and time for staff to build community partnerships.
- Each project should state who defines the problem, who owns the data, who governs the system and who will judge success, and should treat a decision not to use AI as a legitimate outcome. Proposals should be tested against the five partnership principles (Figure 2).
- Before starting work, identify who the proposed system excludes: those who will experience its effects, those whose knowledge and language are omitted, and those with no standing to evaluate it (see the definition in Section 2). Who speaks for a community should also be established, working through structures already accountable to their own members, such as cooperatives or elected local bodies.

- Universities should establish data and AI governance agreements that give communities a meaningful say over how their data is collected and used, and who benefits. The agreements should also set out an exit route for both the research team and the community if the partnership breaks down, including what happens to the data and the system.

4.2 Reform incentives for community-engaged work

Participatory, capability-centred work takes more time and resources than deployment-first research, and current evaluation criteria of publications, grants and speed to deployment work against it. Unless the surrounding incentives change, a single team's commitment to community-engaged work would be reabsorbed by the pressures around it. Reform must therefore be institutional, not left to individual goodwill, and must reach the national assessment frameworks and ministries that set many universities' evaluation criteria.

Implementation Pathways

Audience: University leadership; research councils and funders; national assessment bodies and ministries.

- Universities should recognise and reward community-engaged work in promotion criteria, workload credit, and dedicated awards, giving it standing alongside conventional outputs.
- Research councils, funders (including industry and ministries of higher education) and the national bodies that set assessment criteria should support long-term partnerships in which communities hold governance roles from the beginning, rather than short cycles that end before trust is built.
- Ministries of higher education could create a better funding environment for institutional reform that encourages universities to define excellence beyond publication and citation metrics. Competitive and matching-fund schemes could incentivise sustained university–industry–community partnerships, and projects that demonstrate measurable public value.
- Universities should report measures of community capability, partnership quality, institutional learning, and public value created through AI initiatives alongside traditional indicators such as publications, patents and prototypes.

4.3 Treat communities as knowledge co-creators

A capability approach asks not what problem AI can solve for a community, but which capabilities the community values and how technology might expand them (Section 2.2). This has two implications for what universities teach. First, communities should be approached not through a deficit lens, which sees only what they lack, but as holders of rich local knowledge and problem-solving capacity that AI should build on rather than replace. Second, if knowledge is to flow in both directions, communities must be recognised as teachers, not only learners.

Implementation Pathways

Audience: Academic staff and programme leads; community partners.

- Universities should embed capability-based AI literacy, which ties AI skills to what learners and their communities want to be able to do, into existing programmes, not standalone access initiatives.
- Curricula should be co-designed with underserved communities, so that local knowledge systems inform what is taught alongside technical skills.
- Where regulations allow, researchers should bring community partners into teaching as recognised and credited co-instructors, training students in the participatory and speculative co-design methods illustrated in Section 2. The aim is to move communities from consumers of AI to co-creators of it.

4.4 Sustainability beyond the grant

Sustainability should be a condition of ethical AI projects and not an afterthought (Section 2.3). Community partnerships and the systems they produce need to outlast the grant that started them. Projects that collapse when a grant ends

damage public trust and deepen the very disillusionment that participatory work seeks to overcome.

Implementation Pathways

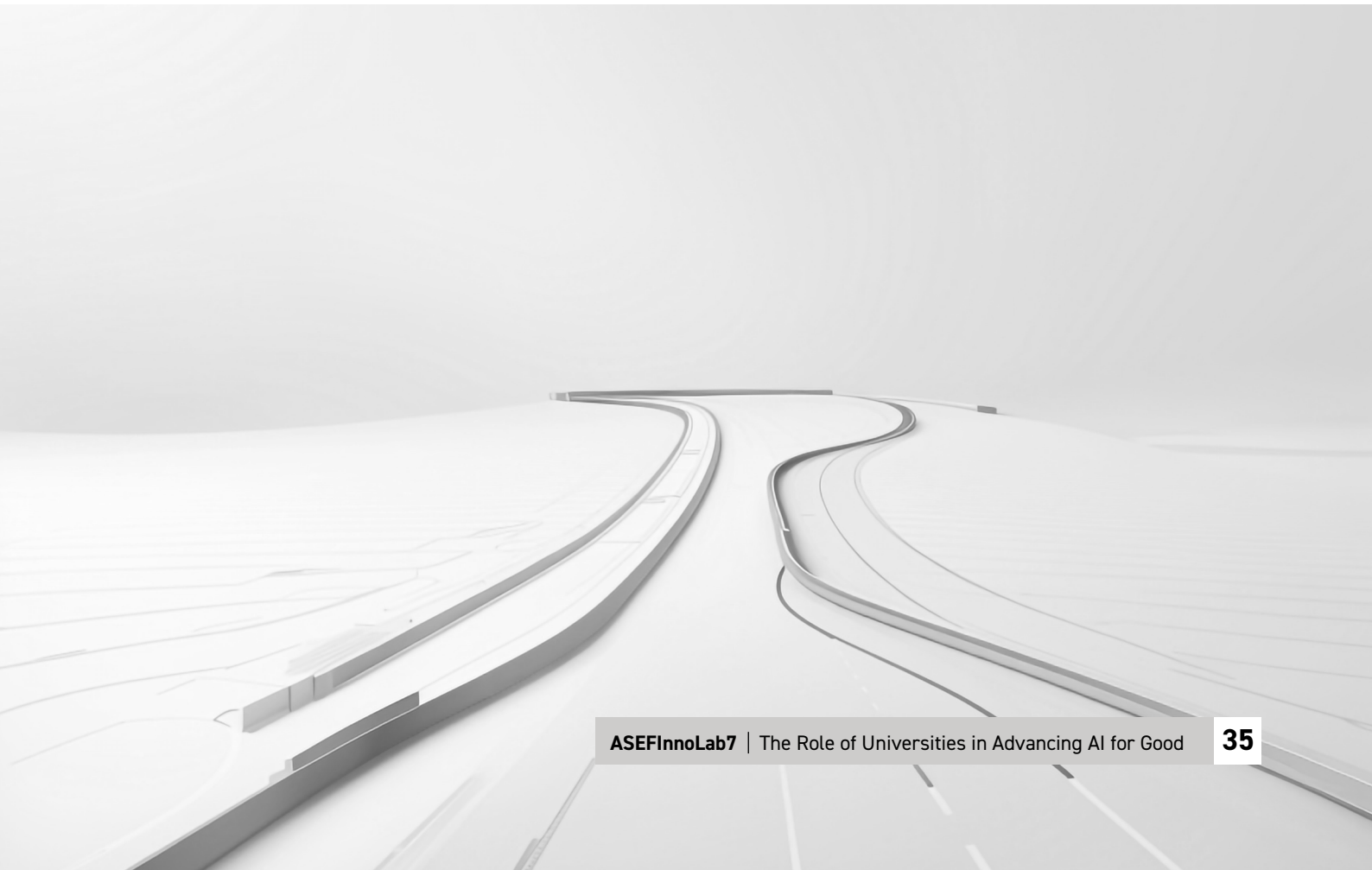
Audience: University leadership; local governments; enterprises; philanthropic and regional funders.

- Universities should bridge funding between grants and a named partnership owner who stays when projects end.
- Universities should encourage pooled regional or institutional mechanisms, in which local governments fund maintenance and connect systems to public services, enterprises supply infrastructure and procure on community-agreed terms, and philanthropic partners bridge grant cycles, so that the community members and knowledge-holders whose contributions make these systems work stay engaged.
- Universities across regions should build a shared knowledge hub that exchanges methods and governance frameworks on sustainable AI projects, and documents how local practice adapts across contexts, avoiding the extractive research dynamics the literature warns against (Section 2.1).

5 Conclusion

AI has the potential to either widen or narrow existing societal inequalities, and universities are in the position to influence which way it goes. That influence is exercised project by project, through choices about who frames the problem, whose data a system learns from, and who is present when success is defined. Left to prevailing incentives, those choices define good through what secures funding, gets published or advances institutional priorities, which makes inaction a consequential choice in itself.

Universities can change that default, and few institutions are better placed to do so. Through teaching, research and community engagement, they already hold the talent, knowledge and partnerships that community-led AI needs. The University AI for Community Good Framework asks leaders to use them deliberately: to define good with communities first, to invest in AI to serve the good so identified, and to hold each initiative to a three-part test: that it expands freedoms the community values, on terms the community helped set, and that the community judges it a success. In return, universities sustain their public mandate for societal impact and advancing public good, a good that the public itself defines.



Use of generative AI

The authors used Claude (Anthropic)[26] as an editorial aid in improving clarity and concision, tightening structure, and checking consistency across sections.

The tool was not used to generate the paper's argument, to select or summarise the literature, or to produce findings. All sources were chosen and verified by the authors, and all final wording was written or approved by the authors. The authors take full responsibility for the content of this paper.

Authors

Dr Yang BONG
Birmingham Newman University, United Kingdom

Dr Neelima SAILAJA
University of Nottingham, United Kingdom

Dr Minjia CHEN
University of Nottingham, United Kingdom

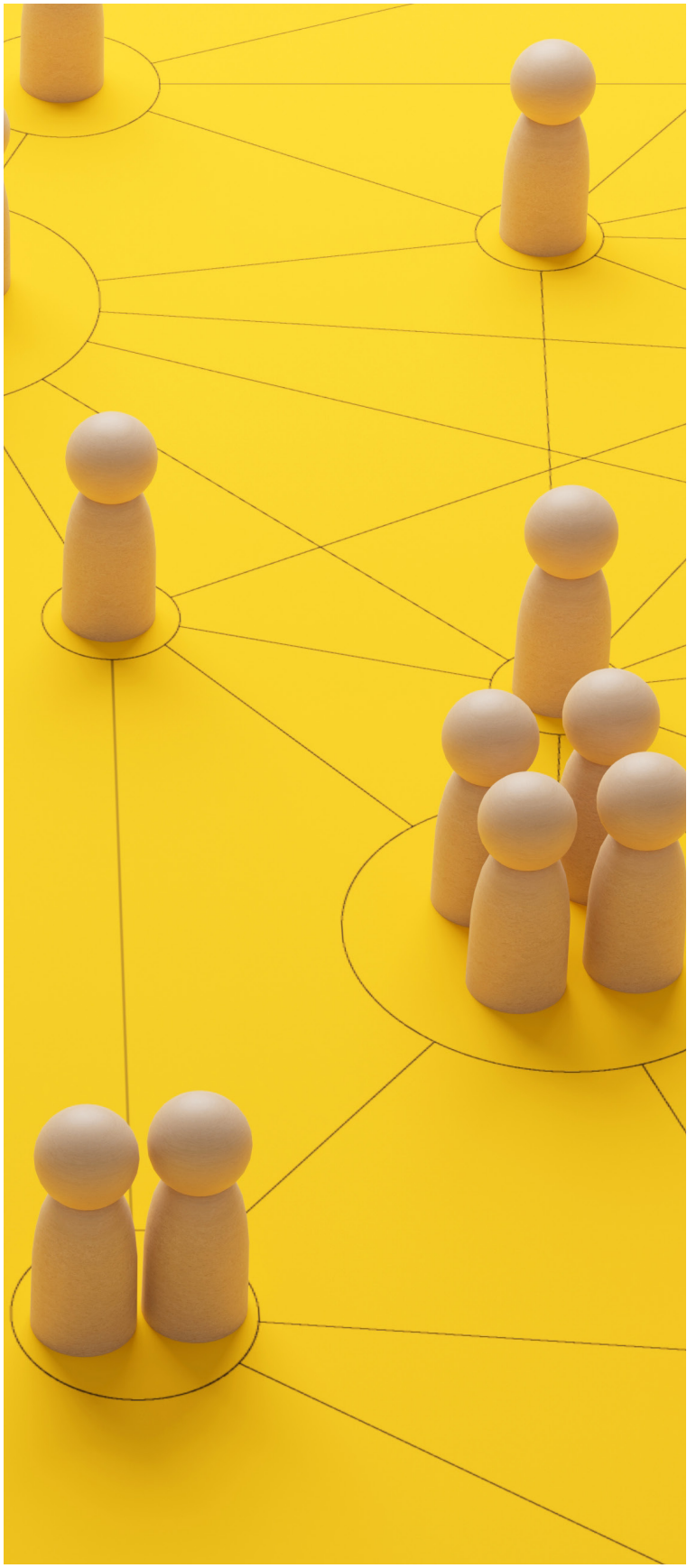
Dr Devisakti ANNAMALI
Universiti Sains Malaysia, Malaysia

Mr Jin TADA
National Graduate Institute for Policy Studies, Japan

Ms Leïla CELLÉRIER
University of Geneva, Switzerland

Dr Muhammad Zahid IQBAL
Teesside University, United Kingdom

Dr Vicente PITOGO
Caraga State University, Philippines



References

1. Bong, Y., Yunus, S., Weerawarna, R., Senkerik, R., Tan, R., Masood, T., Wawn, T., & Camilleri, V. (2025). *White paper for universities on navigating artificial intelligence innovation ecosystems in AI for sustainable development*. ASEF Higher Education Innovation Laboratory. https://doi.org/10.25878/QWXR-RP84

2. Wawn, T., Masood, T., Driha, O. M., Astuti, Y., & Bong, Y. (2024). *Positioning for success: How universities in developing countries can thrive in the AI era by focusing on sustainable development*. Asia-Europe Foundation. https://nottingham-repository.worktribe.com/output/56144058

3. Cowls, J. (2021). 'AI for social good': Whose good and who's good? Introduction to the special issue on artificial intelligence for social good. *Philosophy & Technology, 34*(Suppl. 1), 1–5. https://doi.org/10.1007/s13347-021-00466-3

4. Bondi, E., Xu, L., Acosta-Navas, D., & Killian, J. A. (2021). Envisioning communities: A participatory approach towards AI for social good. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 425–436). Association for Computing Machinery. https://doi.org/10.1145/3461702.3462612

5. Moore, J. (2019). AI for not bad. *Frontiers in Big Data, 2*, Article 32. https://doi.org/10.3389/fdata.2019.00032

6. Oluwatuyi, R., Pillay, V., Mazwi, J., Castro, A., Farao, J., Srivatsa, S., Bagalkot, N., & Densmore, M. (2026). Collectively reimagining artificial intelligence with marginalized communities. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems* (pp. 1–21). Association for Computing Machinery. https://doi.org/10.1145/3772318.3791421

7. Özler, B. (2012, June 14). One Laptop per Child is not improving reading or math. But, are we learning enough from these evaluations? *Development Impact*. World Bank Blogs. https://blogs.worldbank.org/en/impactevaluations/one-laptop-per-child-is-not-improving-reading-or-math-but-are-we-learning-enough-from-these-evaluati

8. Cristia, J., Ibarrarán, P., Cueto, S., Santiago, A., & Severín, E. (2017). Technology and child development: Evidence from the One Laptop per Child program. *American Economic Journal: Applied Economics, 9*(3), 295–320. https://doi.org/10.1257/app.20150385

9. Ames, M. G. (2019). *The charisma machine: The life, death, and legacy of One Laptop per Child*. MIT Press.

10. Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press.

11. London, A. J., & Heidari, H. (2024). Beneficent intelligence: A capability approach to modeling benefit, assistance, and associated moral failures through AI systems. *Minds and Machines, 34*(4), Article 41. https://doi.org/10.1007/s11023-024-09696-8

12. Cenci, A., & Cawthorne, D. (2020). Refining value sensitive design: A (capability-based) procedural ethics approach to technological design for well-being. *Science and Engineering Ethics, 26*(5), 2629–2662. https://doi.org/10.1007/s11948-020-00223-3

13. Lewis, J. E., Whaanga, H., & Yolgörmez, C. (2025). Abundant intelligences: Placing AI within Indigenous knowledge frameworks. *AI & Society, 40*(4), 2141–2157. https://doi.org/10.1007/s00146-024-02099-4

14. Saxena, D., Kahn, Z., Moon, E. S.-Y., Chambers, L. M., Jackson, C., Lee, M. K., Eslami, M., Guha, S., Erete, S., Irani, L., Mulligan, D., & Zimmerman, J. (2025). Emerging practices in participatory AI design in public sector innovation. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1–7). Association for Computing Machinery. https://doi.org/10.1145/3706599.3706727

15. Bonvin, J.-M., & Laruffa, F. (2018). Human beings as receivers, doers and judges: The anthropological foundations of sustainable public action in the capability approach. *Community, Work & Family, 21*(5), 502–518. https://doi.org/10.1080/13668803.2018.1526777

16. Sen, A. (1999). *Development as freedom*. Oxford University Press.

17. Ibrahim, S. S. (2006). From individual to collective capabilities: The capability approach as a conceptual framework for self-help. *Journal of Human Development, 7*(3), 397–416. https://doi.org/10.1080/14649880600815982

18. Stanford Institute for Human-Centered Artificial Intelligence. (2026). *The 2026 AI index report*. Stanford University. https://hai.stanford.edu/ai-index/2026-ai-index-report

19. Neudert, L.-M., & Howard, P. N. (2020). *Four principles for integrating AI & good governance* (Working Paper 2020.1). Oxford Commission on AI & Good Governance, Oxford Internet Institute, University of Oxford. https://www.oii.ox.ac.uk/news-events/reports/four-principles-for-integrating-ai-good-governance/

20. Slattery, P., Saeri, A. K., Grundy, E. A. C., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., & Thompson, N. (2026). The AI risk repository: A meta-review, database, and taxonomy of risks from artificial intelligence. *Patterns, 7*(5), Article 101517. https://doi.org/10.1016/j.patter.2026.101517

21. Greene, D., Levine, D., & Nahar, K. (2025). Raising the bar in community partnerships: Principles of co-design and sustainability. In B. S. Garg (Ed.), *Community health models: An attempt to address emerging challenges and technology in diverse settings*. IntechOpen. https://doi.org/10.5772/intechopen.1013843

22. Lansing, A. E., Romero, N. J., Siantz, E., Silva, V., Center, K., Casteel, D., & Gilmer, T. (2023). Building trust: Leadership reflections on community empowerment and engagement in a large urban initiative. *BMC Public Health, 23*, Article 1252. https://doi.org/10.1186/s12889-023-15860-z

23. Responsible AI UK. (n.d.). *Off-grid AI deployments: Re-gearing AI model development through community co-design*. Retrieved May 18, 2026, from https://rai.ac.uk/new_projects/off-grid-ai-deployments-re-gearing-ai-model-development-through-community-co-design/

24. Shedlock, K., Chung, A. Q. H., Corbett, J., & Rawson, Z. (2025). Indigenous knowledges as justificatory knowledge in design science research: An expository case. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 56*(4), 36–55. https://doi.org/10.1145/3774330.3774334

25. Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology, 33*(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

26. Anthropic. (2026). *Claude* [Large language model]. https://claude.ai

The Need for a Contextualised Competence Framework for GenAI in Higher Education

AUTHORS
Daniel Burgos
Radwan Eskhita
Favour Borokini
Xaysavath Soukchaleun
Fernaz Narin Nur
Rafis Abazov

This white paper considers whether GenAI serves higher education and the common good through equitable access to quality education, enhanced human capacity, reliable knowledge, participation, and socially responsible innovations, though the competence development based on a contextualised framework implemented by target groups.

Executive summary

Generative artificial intelligence (GenAI) is no longer in its experimental phases. Technology has evolved into an institutional resource that can support teaching and learning in colleges and universities. GenAI is already impacting teaching, learning, assessment, research and administration. However, institutional governance, professional development, and infrastructure are still unevenly developed. This paper argues that GenAI can contribute to the common good only through a purposeful, clustered and contextualised adoption process.

Using insights gained from existing research literature and practical institutional experience, this white paper discusses GenAI on three interrelated levels: the university as an institution, the target groups (or clusters) of stakeholders utilising GenAI, and the competencies and processes required for its responsible integration. There are many clusters at the university: students, academic faculty, academic management, support staff, parents, executive boards, content providers, and others. All of them share a common need for critical GenAI literacy. Nonetheless, they require different role-specific capabilities, safeguards and forms of support. This paper concentrates on the first four clusters.

This white paper examines evidence and practice relating to the value added by GenAI for universities and its target groups, stakeholders, and society at large. It further explores issues related to institutional integration, role-specific uses, the need for a competence framework, and some limitations and issues involved. The white paper concludes with specific recommendations for each targeted cluster, focusing on how a GenAI competence-based framework can add value to universities and contribute to the common good.



1 Introduction

Generative artificial intelligence (GenAI) is increasingly being interwoven into all aspects of society, including higher education. In higher education, GenAI continues to impact how university students learn, how educators teach and assess students, and how academicians conduct and communicate research. Furthermore, GenAI influences how institutions conduct academic and administrative matters. As technology advances and calls for constant updating of skills, society must take into account the problem of rapid progress in innovation from an ethical standpoint. In such a case, GenAI cannot be regarded as an independent technology but rather as a technology reliant on the context.

This white paper considers whether the transformation serves the common good through equitable access to quality education, enhanced human capacity, reliable knowledge, participation, and socially responsible innovations. Such benefits could arise on different levels, ranging from improved individual productivity to increased efficiency and effectiveness of institutions and beyond. However, these benefits do not come automatically.

Indeed, universities require an overarching framework that communicates to different clusters of users – such as students, faculty, institutional managers, and support staff – what responsible and effective implementation of GenAI is. Such a model should outline the required competencies for each of the aforementioned target groups, assessment of progress, as well as requirements on issues related to privacy, integrity and access to technologies. Responsible implementation of GenAI promises universities a host of benefits related to improved teaching and learning, enhanced trust, increased participation, and strengthened contribution to society.

2 Desk review on how GenAI is integrated in the university system

A key question is whether the transformation serves the common good through equitable access to quality education, enhanced human capacity, reliable knowledge, participation, and socially responsible innovation.

This section reviews the evidence on GenAI adoption in higher education institutions (HEIs), drawing on peer-reviewed studies, institutional surveys, and policy documents published between 2022 and 2026. Consistent with the audience of this white paper, the focus is institutional practice; national and international regulation is discussed only where it shapes the room to serve institutions.

HEIs are facing a fast and largely unexpected disruption. Since the public launch of large language models (LLMs) that were widely available from late 2022, GenAI research in higher education has increased exponentially – a bibliometric study of more than 4,000 papers reveals this period as a turning point in the field.[1] The speed of development has compressed the time normally available for institutions to test technologies, build consensus, and establish governance. Even further, to allow individuals to adapt.

Institutional responses have largely been uncoordinated, with each institution adopting its own approach to GenAI. A survey by UNESCO found that fewer than 10% of schools and universities had developed formal policies or governance frameworks for GenAI.[2] In addition, new data reveals that while there is a changing landscape, the picture is uneven, with only one third of students in the United Kingdom having formal training in how to use tools such as ChatGPT, despite 92% of students using them.[3] This gap raises questions about how effective existing institutional support actually is.

This disparity can be seen even among well-resourced institutions. A study of 14 Big Ten universities in the United States revealed that there was no common governance model: institutions had mixtures of multi-unit governance (which divides governance responsibility amongst teaching, research, and information technology units); role-specific governance (whereby students, faculty and staff received different governance); and academic discipline specific governance (whereby different disciplines had different governance, which could occur within the same institution).[4]

GenAI has been most impactful in three fundamental areas, namely teaching, research and administration, and GenAI tools are impacting assessment practices,

norms of research collaboration, and administrative workflows. Because decisions in one domain affect the others, these applications require coordinated institutional checks and balances rather than isolated local responses.[5] Unfortunately, institutional guidelines are often too broad to be applicable to GenAI, with research revealing narrow perspectives on GenAI held by universities. They view it mainly as a challenge to academic integrity, with little focus on the collaborative and technology-driven processes of knowledge creation.[6] These are two distinct problems rather than a contradiction: existing institutional guidelines are too vague to inform day-to-day practice, while the underlying institutional view of GenAI remains narrowly focused on integrity risks. This suggests the need from institutions to develop broader GenAI conceptual frameworks that balance risk management with innovation in knowledge creation.

Policy divergence is a recurring theme in literature when discussing generative language models. After the analysis of policy documents from 40 universities across six global regions, Jin et al.[7] found governance approaches ranging from prohibition to active encouragement of GenAI use. National-level regulation diverges just as much. As an illustration, China's regulation focuses on the service providers but does not mention the education context, a divergence that lies in the different focus of their respective guidance. Meanwhile, Japan is geared toward higher education administrators, and the USA focuses on K-12 settings (primary and secondary education, from kindergarten to the final year of secondary school) and includes a wider array of AI technologies.[3]

There have been efforts by some large organisations to give form to this chaotic environment. UNESCO published important guidelines concerning the responsible use of AI in research. The institution emphasised the need for transparency, accountability and disclosure.[2] The European Union has also issued guidelines on the responsible use of AI. The directives of the European Commission (2022)8 are, importantly, addressed to primary and secondary educators and not to higher education ones – a sign of the difference in coverage that the section highlights. Despite these guidelines, policy fragmentation persists. Such is the case of national and institutional regulations, which are still in place and sometimes conflicting and even counterproductive.[9]

All these studies focus on two points: first, institutional policy is not keeping up with the speed of GenAI adoption; second, where policy exists, it is not consistent or uniform between institutions or jurisdictions. They also differ from national and international regulations in what they consider to be the primary issue. While national and international guidance emphasises transparency, human-centredness and rights protection, institutional studies emphasise academic integrity and operational risk.[6,7]

3 Use of GenAI in higher education by clusters

The challenge is not only to create rules for the use of GenAI, but to develop an institutional culture that embraces GenAI and integrates it into the mission of education and the common good, that can be effectively implemented by clusters.

The application of GenAI at universities varies due to the dissimilar nature of roles, risk exposure, and anticipated outcomes by the stakeholders. Students, faculty members, university administrators, and support personnel use the tools for different uses with varying forms of accountability. The grouping of such stakeholders as clusters allows universities to design more relevant policies, training programmes, and support systems while maintaining a common institutional framework.

- **Students use GenAI largely.** For instance, in the 2023-2024 academic year, it was established that more than 60% of undergraduate learners in China relied on GenAI in some capacity in their academic work.[10] According to Ravšelj et al.,[11] learners use these AI tools for writing, brainstorming, summarizing learning materials, programming, and language support. However, issues may arise where students take information from AI technology without confirming its accuracy and using it to the detriment of critical thinking and problem-solving,[12] which might increase shortcomings.[3,13]

- **Academicians use GenAI for both educational work and scientific research.** Nowadays, many teachers use AI technology for lesson planning, testing, and producing learning content.14 Some use it for literature research, coding, and scholarly writing. On the other hand, faculty members are afraid of the inaccuracy of AI algorithms, bias in information obtained from AI, and that students will rely on AI instead of thinking independently.[14] Training programmes are known to increase the acceptance and confidence of staff in the application of GenAI.[15]

- **Academic managers, including department heads, deans, and other senior leaders, use GenAI in planning, decision-making, and institutional management.** AI tools may help to analyse institutional data, monitor students' progress, ensure quality, implement strategic plans, and develop policies to make data-driven decisions.[16] Nevertheless, university leaders still face difficulties because governing frameworks and policies related to GenAI are in their early stages of development in HEIs.[17] There is a need for collaboration rather than centralisation,[4] which requires that institutional and academic managers have both technological and governance skills.

- **Support staff use AI tools to write emails, make appointments, arrange files, answer questions from students, and conduct information searches.**[14] While these are very useful functions for making the processes more efficient, they also imply the need for employees to acquire new digital skills. The support staff consists of administrative officers, librarians, student services staff, and information technology (IT) specialists.

Table 1 summarises how GenAI applications, benefits, risks and institutional support requirements differ across the four clusters, which are interdependent. Management and procurement decisions shape access to tools; academic guidance influences student practice; and support staff provide the infrastructure, security and information literacy required for responsible use. A cluster-based approach should therefore combine a common institutional framework with differentiated responsibilities and support.

Cluster	Main applications of GenAI	Potential contribution to the common good	Key risks	Institutional support required
Students	Brainstorming, writing support, summarisation, programming and language assistance	More equitable access to learning support, improved participation and personalised learning	Inaccurate information, overreliance, reduced critical thinking, privacy breaches and academic misconduct	AI-literacy education, disclosure rules, verification skills and equitable access
Academic staff	Lesson planning, learning-material development, formative feedback, assessment design, coding and research support	More inclusive teaching, timely feedback and increased time for student engagement	Bias, hallucination, confidentiality concerns, unreliable assessment and erosion of trust	Professional development, disciplinary guidance, protected experimentation time and peer learning
Academic managers	Policy drafting, report summarisation, scenario development, communication and institutional planning	More responsive governance, informed planning and accountable resource allocation	Automation bias, opaque decision-making, unfair outcomes, vendor dependency and data misuse	Governance framework, risk assessment, procurement standards, human oversight and appeal mechanisms
Support staff	Routine communication, student enquiries, information organisation, research support and service delivery	More accessible services, faster responses and reduced administrative barriers	Confidentiality breaches, inaccurate advice, cybersecurity threats and inappropriate automation	Role-specific training, data-handling rules, workflow redesign and technical safeguards

Table 1. *Cluster-based uses, benefits and risks of GenAI in higher education*

What is notable across all four clusters is that the adoption of GenAI took place much faster than the universities were prepared for. Not all institutions have established proper policies, training programmes, and ethical frameworks.[3,5] This lack of institutional readiness should be addressed by a competence framework to ensure that GenAI is valuable to higher education.

4 Required core and specific competencies for using GenAI by clusters

What does AI competence mean? It is being aware of the use of these AI technologies, having knowledge about their capacities and constraints, making appropriate requests, understanding the output, verifying claims based on available evidence, and knowing when human knowledge prevails.

The competencies required for responsible GenAI use can be organised under the broader concept of critical AI literacy. This concept combines technical understanding with judgment about evidence, ethics, social consequences, and appropriate human oversight. A shared core is necessary across the institution, but competence must also be adapted to the tasks and responsibilities of each stakeholder cluster.

AI literacy refers to "a set of competencies that enables individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool online, at home, and in the workplace"[18] and is a subset of the broader digital literacy competence.[19] In the context of higher education, AI competence means being aware of the use of these AI technologies, having knowledge about their capacities and constraints, making appropriate requests, understanding the output, verifying claims based on available evidence, and knowing when human knowledge prevails.[20]

Further, critical AI literacy, unlike ordinary AI or digital literacy, emphasises critical inquiry in the use and engagement of AI that involves not only the acquisition of skills such as prompt engineering or software design, but also includes the ability to understand the wider societal and ethical implications of AI.[21,22]

For students, there is a need to understand the best GenAI versions at their disposal. Students also need to be aware of the importance of academic rigour and integrity, as well as the cognitive deskilling risks of excessive reliance on AI.

Among faculty members, AI is already being used to help create and enhance assessments. In an indirect manner, the prevalence and reliance on AI use among students have also led faculty to reevaluate how knowledge is being assessed in higher education settings, including the possibility of returning to in-person examinations.[23] Statements like the Russell Group principles on the use of GenAI tools in education[24] also highlight the role universities and staff might play in the appropriate and effective use of AI in the learning experience.

However, the deployment of AI in higher education has also been connected to a deterioration of trust between academics and students. The impact of this erosion of trust is also visible in other parts of the research process, such as publication and peer review, leading to new rules on the unchecked use of AI in research papers.[25]

The significance of human-centredness is equally applicable to academic managers and support staff who need to be aware of the ways AI use contributes to the erosion of trust in the teacher-student relationship. This might happen due to misuse by students, and because of loss of agency by faculty, which includes proctoring, potential bias, and cheating.[26,27] They need to be aware and willing to engage with concerns around reputational harm potentially triggered by the integration of AI into higher education. These competencies cannot be built in isolation – they require the institutional processes and investment discussed in the following section.

Table 2 provides a structured, cluster-based synthesis of the core and specific competencies required across HEIs to foster responsible GenAI adoption anchored in the common good. Grounded in the broader concept of critical AI literacy, this summary table outlines how technical capabilities, ethical judgment, and output verification skills must be tailored across four key institutional stakeholder groups: students, academic staff, academic managers, and support staff.

Academic Cluster	Main GenAI Applications	Required Core and Specific Competencies
Shared institutional Core	General inquiry, text/code generation, information search	*Critical AI Literacy:* Understanding GenAI capabilities and constraints, verifying outputs against evidence, protecting data privacy, adhering to disclosure rules, and knowing when human judgment must prevail
Students	Brainstorming, writing support, summarisation, programming, and language assistance	*Independent Learning & Output Verification:* Discerning AI limitations, fact-checking hallucinations, applying critical thinking to prevent cognitive deskilling, preserving academic integrity, and transparently acknowledging AI use
Academic staff	Lesson planning, learning material development, formative feedback, assessment redesign, coding, and research support	*Pedagogical & Evaluative Judgment:* Integrating GenAI to enhance teaching and research without compromising academic rigor, designing process-oriented assessments, identifying algorithmic bias, and maintaining peer-review/research integrity
Academic management	Policy drafting, report summarisation, scenario development, institutional planning, and strategic decision-making	*Strategic AI Governance & Human Oversight:* Evaluating vendor risks (data retention, lock-in), conducting formal risk assessments for high-stakes decisions, mitigating automation bias, and fostering open dialogue on AI concerns
Support staff	Administrative communications, student enquiries, information organisation, research support, and service delivery	*Data Security & Workflow Efficiency:* Developing role-specific digital skills, handling confidential data securely, implementing cybersecurity protocols, and identifying appropriate vs. inappropriate automation boundaries

Table 2. *Core and specific competencies by clusters*

5 A case study: processes and instruments to properly integrate GenAI in higher education

Six factors to show GenAI adoption in higher education are based on technology, governance, institutional readiness, and continuous evaluation

This section is focused on six factors that are applicable for AI competence development: leadership and governance, policy, training, funding for infrastructure, partnerships and evaluation. For this purpose, one case study of higher education in Laos will be used, which is based on national policy and the experience of Souphanouvong University in Laos. The reason for selecting this country is justified by the fact that it represents a low-resource country, where the success of initiatives relies on the efforts of policymakers, financial support, and preparedness of institutions, rather than the peculiarities of technology.

- **First, effective leadership and governance are necessary for the successful implementation of GenAI in universities.** Higher education leadership should have a clear vision and strategy guiding the implementation of GenAI technologies and ensuring that the integration corresponds to the university's mission, goals, and long-term thinking. GenAI tools should not be implemented separately from the processes of teaching, research, administrative work, and community work. Consequently, the successful development of GenAI in universities will be dependent on good leadership.

- **Second, there is a need for the development of clear policies and regulations guiding the appropriate use of GenAI in education, learning, research, and administration.** Such policies should aim at promoting transparency but also maintaining academic freedom, data privacy, intellectual property, and research integrity. These interventions should also guide the ethical use of AI across all members of the university community. As stated by UNESCO,[2] the need for a governance framework cannot be overstated because of its value in mitigating risks and strengthening collaboration among stakeholders.

- **Third, universities need to ensure that ongoing training is available for teachers, non-teaching employees, and students, so that they can make appropriate use of GenAI.** Teachers need to develop practical skills required to effectively

use GenAI in various contexts, such as teaching, administrative activities, and research activities, while students need to understand how to verify the accuracy of the information created by GenAI rather than merely applying it blindly.[15,28]

- **Fourth, universities require sufficient funding to build digital infrastructure to support emerging AI technologies like GenAI.** Some common examples of resources that HEIs can invest in include internet services, cloud storage, cybersecurity systems, and digital learning platforms, among others. Giannakos et al.[29] revealed that one of the major determinants of AI implementation in universities is the presence of a sound digital infrastructure.
- **Fifth, universities need to foster stronger collaboration with governments, businesses, international organisations, and other universities to promote GenAI integration.** This collaboration provides the much-needed opportunity to exchange best practices and expertise.[30]
- **Sixth, and last, universities should monitor the impact of GenAI on the educational quality, performance, and employability of their graduates.** Constant assessment allows them to overcome emerging implementation challenges associated with GenAI and ensure that these initiatives correspond to the strategic goals of universities and the country at large.[17,31]

Collectively, these six factors show that GenAI adoption in higher education depends not only on the technology itself but on governance, institutional readiness, and continuous evaluation. Although this case study is grounded in Lao PDR's experience – where limited funding, competencies, and policy support remain the central constraints – the lessons learned have practical value for HEIs adopting GenAI elsewhere.

6 Limitations of the implementation of GenAI in higher education

A key limitation related to the deployment of GenAI in higher education is its current unreliability as a study tool.[29] In its current state of development, GenAI tools do not reliably produce accurate content and cannot serve as a competent stand-in for a tutor. Unfortunately, while these inaccurate or fabricated content, also known as hallucinations, may be easily spotted by knowledgeable experts who use AI, students and other novice users may not be capable of detecting them. In this regard, GenAI should therefore support, rather than replace, qualified teaching, disciplinary expertise, and independent verification. [2,31]

The integration of GenAI tools in higher education raises ethical issues because it exacerbates socio-economic inequalities related to the digital divide and academic precarity. As further stated by García-López and Trujillo-Liñán,[32] the application of GenAI in the academic setting may redefine the dynamic between the student and teacher, thus making one question what learning entails at present. The authors also imply that disparities in the use of GenAI tools could exacerbate the digital divide and result in more inequalities in educational achievements. Moreover, the use of GenAI could also lead to cognitive dependency in learners and prevent them from developing critical skills.[33] Finally, GenAI presents epistemic threats to scholarly research due to its reliance on training data, which can skew research findings.[34]

In terms of legal limitations, the effective use of GenAI tools in higher education is constrained by evolving or inconsistent regulatory frameworks governing their use.[23] These legal concerns include intellectual property, data protection, copyright and academic integrity, highlighting the need for transparency regarding the data used to train GenAI models and clear guidance on acceptable and prohibited uses by students, faculty and professional staff.[23,35]

The use of GenAI by higher education stakeholders also raises significant concerns about data privacy and the protection of sensitive information entered into AI systems.[36] This is particularly sensitive issue in research settings due to the value of originality in academic research and the risk of harm to students and, therefore, liability to educational institutions. The limitation further extends to research subjects and their data, as collected, stored and analysed by researchers within the university and their right

to data protection and privacy, thereby necessitating clear policies and guidelines as to how data is stored. For less well-resourced institutions, ensuring the necessary legal and data-protection safeguards may be even more difficult.

7 Recommendations for the effective use of GenAI for the common good in higher education

Focus by cluster. An AI competence-based framework should focus on the most impactful clusters, such as students, academic staff, academic management, and support staff.

As presented before, an AI competence-based framework is required in higher education. This competence framework should describe the basic knowledge and responsibilities of every stakeholder within the institution. Focus should be placed on defining how effectively and responsibly use AI works, evaluating output, protecting information, acknowledging use, and identifying situations where human judgment should take precedence. Furthermore, the framework should provide details regarding the responsible use of GenAI by students, academics, managers, and other employees in their daily operations. Further, each university and cluster should adapt the competence framework tailoring it to their missions and needs and to align with its values, disciplines and philosophy.

While it is important to consider all target groups in higher education when developing recommendations, the most attention should lean on the most impactful clusters, such as students, academic staff, academic management, and support staff. Each cluster has its own specific needs, requirements, expectations and approaches when integrating AI into teaching, research and administrative practices.

For students, the competence framework should be designed in a way that emphasises responsible and independent use of GenAI while studying. In the initial stages, the students should know when to use GenAI tools, their benefits and limitations, and how to use critical thinking to evaluate information generated with their help. As their skills develop, they should learn transparency rules, which require acknowledging the use of AI.[37] Responsible use can prevent overreliance on GenAI.

For academic staff, the framework needs to assist educators to use GenAI in a way that enhances the teaching, assessing and researching processes without jeopardising critical thinking and academic judgment. It is important for academic staff to understand the areas in which GenAI can help with teaching and learning, have clear guidelines concerning the use of these tools, and create assessments that will test not only the end product but also the whole learning process. Furthermore, it is also crucial for universities to provide sufficient time, support, and opportunities for academic staff to learn from their colleagues across various disciplines.

With regards to academic management and support staff, the competence framework should go beyond merely knowing how to use the GenAI tools. It should involve the knowledge of how to govern GenAI responsibly by safeguarding privacy and streamlining cybersecurity, accessibility, record management, and procurement. In cases where GenAI is used for high-stakes decision-making, formal risk assessments should be undertaken, and qualified personnel must review the recommendations suggested by the tools before any decisions are made. In addition, institutions should carry out a thorough evaluation of vendors by considering issues like data retention, transparency of the AI tools, integration with other systems, vendor lock-in risks, and alternative options to GenAI.

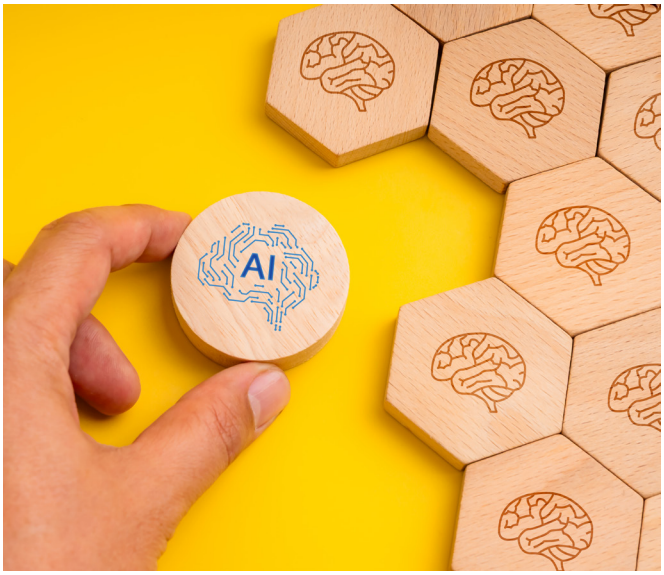
8 Conclusion

GenAI's effect in higher education is already being felt. However, a competence framework, properly contextualised by institution, and effectively designed by target groups or clusters, is required to ensure effective and responsible use.

This paper analyses disruptions in terms of institutional governance, stakeholder groups, competencies, implementation processes, and limitations, all of which support one conclusion: GenAI's education value depends on the conditions surrounding its use. In the absence of the right context, the use of GenAI could lead to the propagation of disinformation, inequality, dependency, and distrust. On the contrary, when used correctly, it could improve responsiveness, productivity and efficiency. And, to this extent, some guidance is needed in the form of a competence-based framework by clusters.

The bottom line.

A competence framework, properly contextualised by institution, and effectively designed by clusters, is required to ensure effective and responsible use of AI in higher education for the common good.



Acknowledgements

The work of Ms Favour Borokini on this paper is partially supported by the Engineering and Physical Sciences Research Council [Grant Number EP/S023305/1].

Authors

Prof Daniel BURGOS
Universidad Internacional de La Rioja, Spain

Dr Radwan ESKHITA
Lucerne University of Applied Sciences and Arts, Switzerland

Ms Favour BOROKINI
University of Nottingham, United Kingdom

Mr Xaysavath SOUKCHALEUN
Souphanouvong University, Lao PDR

Dr Fernaz NARIN NUR
Daffodil International University, Bangladesh

Dr Rafis ABAZOV
Institute for Green and Sustainable Development, Kazakhstan

References

1. Frumin, I., Vorochkov, A., Kiryushina, M., Platonova, D., & Terentiev, E. (2026). Mapping the generative AI research in higher education: 2022–2024 insights. *Higher Education Quarterly, 80*(1). https://doi.org/10.1111/hequ.70075

2. UNESCO (2023). Guidance for generative AI in education and research. *UNESCO*. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

3. .Jiang, Y., Xie, L., & Cao, X. (2025). Exploring the effectiveness of institutional policies and regulations for generative AI usage in higher education. *Higher Education Quarterly, 79*(4), Article e70054. https://doi.org/10.1111/hequ.70054

4. Wu, C., Zhang, H., & Carroll, J. M. (2024). AI governance in higher education: Case studies of guidance at Big Ten universities. *Future Internet, 16*(10), Article 354. https://doi.org/10.3390/fi16100354

5. Bacalso, F. D., Villamor, F. E., Paclipan, C. P., Orquia, C. J. B., Rendon, I. P., Albiso, A. T., & Himang, C. M. (2026). Institutional approaches to generative AI management in higher education: A systematic review. *Frontiers in Education, 11,* Article 1814426. https://doi.org/10.3389/feduc.2026.1814426

6. Nikolic, S., Sandison, C., Haque, R., Daniel, S., Grundy, S., Belkina, M., Lyden, S., Hassan, M., & Neal, P. (2024). ChatGPT, Copilot, Gemini, SciSpace, and Wolfram versus higher education assessments: An updated multi-institutional study of the academic integrity impacts of generative artificial intelligence (GenAI) on assessment, teaching, and learning in engineering. *Australasian Journal of Engineering Education, 29*(2), 126–153. https://doi.org/10.1080/22054952.2024.2372154

7. Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence, 8*, Article 100348. https://doi.org/10.1016/j.caeai.2024.100348

8. European Commission. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators.* Publications Office of the European Union. https://education.ec.europa.eu/focus-topics/digital-education/actions/plan/ethical-guidelines-for-educators-on-using-artificial-intelligence

9. Azevedo, L., Robles, P., Best, E., & Mallinson, D. J. (2025). Institutional policies on artificial intelligence in higher education: Frameworks and best practices for faculty. *New Directions for Adult and Continuing Education, 2025*(188), 70–78. https://doi.org/10.1002/ace.70013

10. Guo, F., Zhang, L., Shi, T., & Coates, H. (2025). Whether and when could generative AI improve college student learning engagement? *Behavioral Sciences, 15*(8), Article 1011. https://doi.org/10.3390/bs15081011

11. Ravšelj, D., Keržič, D., Tomaževič, N., Umek, L., Brezovar, N., Iahad, N. A., et al. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE, 20*(2), Article e0315011. https://doi.org/10.1371/journal.pone.0349691

12. Acosta-Enriquez, B. G., Farroñan, E. V. R., Zapata, L. I. V., Garcia, F. S. M., Rabanal-León, H. C., Angaspilco, J. E. M., & Bocanegra, J. C. S. (2024). Acceptance of artificial intelligence in university contexts: A conceptual analysis based on UTAUT2 theory. *Heliyon, 10*(19), Article e38315. https://doi.org/10.1016/j.heliyon.2024.e38315

13. Almassaad, A., Alajlan, H., & Alebaikan, R. (2024). Student perceptions of generative artificial intelligence: Investigating utilization, benefits, and challenges in higher education. *Systems, 12*(10), 385. https://doi.org/10.3390/systems12100385

14. Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Hassan, G. M., et al. (2025). Implementing generative AI (GenAI) in higher education: A systematic review of case studies. *Computers and Education: Artificial Intelligence, 8*, Article 100407. https://doi.org/10.1016/j.caeai.2025.100407

15. Møgelvang, A., Cipriani, E., & Grassini, S. (2025). Generative AI in action: Acceptance and use among higher education staff pre- and post-training. *Technology, Knowledge and Learning,* Advance online publication. https://doi.org/10.1007/s10758-025-09915-w

16. Selvaratnam, R., & Venaruzzo, L. (2024). Human-centered approach to the governance of AI in higher education: Principles of international practice. *Journal of Ethics in Higher Education, 5*, 79–102. https://doi.org/10.26034/fr.jehe.2024.6864

17. Dotan, R., Parker, L. S., & Radzilowicz, J. G. (2024). Responsible adoption of generative AI in higher education: Developing a "points to consider" approach based on faculty perspectives. *AI and Ethics.* Advance online publication. https://doi.org/10.1145/3630106.3659023

18. Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). https://doi.org/10.1145/3313831.3376727

19. Chiu, T. K. F., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). What are artificial intelligence literacy and competence? A comprehensive framework to support them. *Computers and Education Open, 6*, Article 100171. https://doi.org/10.1016/j.caeo.2024.100171

20. Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *International Journal of Information Management, 71*, Article 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

21. Lewison, M., Flint, A. S., & Van Sluys, K. (2002). Taking on critical literacy: The journey of newcomers and novices. *Language Arts, 79*(5), 382–392. https://doi.org/10.58680/la2002255

22. Veldhuis, A., Lo, P. Y., Kenny, S., & Antle, A. N. (2024). Critical artificial intelligence literacy: A scoping review and framework synthesis. *International Journal of Child-Computer Interaction, 43*, Article 100708. https://doi.org/10.1016/j.ijcci.2024.100708

23. Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgment for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education, 49*(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321

24. Russell Group. (2025). *Russell Group principles on the use of generative AI tools in education*. https://www.russellgroup.ac.uk/sites/default/files/2025-01/Russell%20Group%20principles%20on%20generative%20AI%20in%20education.pdf

25. MIT Sloan Management Review Middle East. (2026, May 18). *arXiv to ban authors for a year over unchecked AI use in research papers.* https://www.mitsloanme.com/arxiv-to-ban-authors-for-a-year-over-unchecked-ai-use-in-research-papers/

26. Coghlan, S., Miller, T., & Paterson, J. (2021). Good proctor or "big brother"? Ethics of online exam supervision technologies. *Philosophy & Technology, 34*, 1581–1606. https://doi.org/10.1007/s13347-021-00476-1

27. Baker, R. S., & Hawn, A. (2021). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*(4), 1052–1092. https://doi.org/10.1007/s40593-021-00285-9

28. Ministry of Education and Sports. (2025). *Digital Strategy for Education Lao PDR 2025–2035*. https://moes.edu.la/images/digital%20strategy/Digital%20Strategy%20for%20Education%20Lao%20PDR_2025-2035_en-compressed.pdf

29. Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative AI in education. *Behaviour & Information Technology, 44*(11), 2518–2544. https://doi.org/10.1080/0144929x.2024.2394886

30. Asia-Europe Foundation. (2025). *ASEFInnoLab white paper for universities on navigating AI innovation ecosystems in the area of AI in education.* https://asef.org/publications/asefinnolab-white-paper-education/

31. OECD (2026). OECD digital education outlook 2026: Exploring effective uses of generative AI in education. *OECD Publishing.* https://doi.org/10.1787/062a7394-en

32. García-López, I. M., & Trujillo-Liñán, L. (2025). Ethical and regulatory challenges of generative AI in education: A systematic review. *Frontiers in Education, 10*. https://doi.org/10.3389/feduc.2025.1565938

33. Crompton, H., & Burke, D. (2024). The educational affordances and challenges of ChatGPT: State of the field. *TechTrends, 68*(2), 380–392. https://doi.org/10.1007/s11528-024-00939-0

34. University of Reading. (n.d.). *Generative AI and research guidance.* https://www.reading.ac.uk/research-innovation-hub/managing-your-research/ai-use-in-research-and-innovation

35. Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

36. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

37. Demirel, E. T. (2024). The use and perceptions towards AI tools for academic writing among university students. *Innovations in Language Teaching Journal, 1*(1), 1–20. https://doi.org/10.53463/innovltej.20240328



What policy reforms are needed to enable universities to support equitable lifelong learning and reskilling through AI?

AUTHORS
Lampros Stergioulas
Cinta Gallent-Torres
Munir Moosa Sewani
Anshumaan Goel
Katerina Makri
David Timis

The paper is intended for national and regional policymakers, university leaders, governing bodies, quality assurance agencies, international organisations, higher education networks, and social partners across Asia and Europe. Its scope extends beyond the use of AI in individual courses to consider institutional missions, funding and initiatives, staff capacity, data governance, and universities' contribution to lifelong learning, workforce development, and the public good.

Executive summary

Artificial intelligence (AI) is transforming both higher education and labour markets, intensifying the need for lifelong learning and reskilling while also creating new risks of exclusion. While universities are expected to offer flexible, lifelong learning, existing programmes remain hard to access for adult, non-traditional and disadvantaged learners. In parallel, unequal access to digital infrastructure and AI tools fuels a growing AI divide, while algorithmic bias, opaque decision-making, and weak institutional capacity risk worsening current inequalities.

This white paper argues that AI-enabled lifelong learning is a governance challenge, not just a pedagogical one. Based on global research and policy frameworks, it shows how universities can integrate AI innovation with equitable access, responsible governance and public service.

The analysis identifies three central challenges:

1. Lifelong learning remains out of reach for many non-traditional, displaced and marginalised groups;
2. Bias, data privacy, opacity and limited stakeholder participation continue to pose risks; and
3. Conventional degrees, funding models, and credentialing systems fail to support flexible, portable reskilling.

In response, the white paper makes six recommendations organised around six complementary policy functions:

1. Secure sustainable funding, digital infrastructure, credential recognition, and quality assurance;
2. Expand university offerings beyond traditional degrees to support lifelong education;

3. Establish participatory, transparent, and risk-sensitive AI policies;
4. Connect universities with employers, communities, public authorities, and global networks;
5. Guarantee inclusion for marginalised learners and measure equity outcomes; and
6. Execute reforms through a sequenced, monitored, and accountable framework.

These reforms position universities as key actors in building equitable lifelong learning systems. The central policy message is clear: AI's value in higher education should be judged not by adoption speed, but by its ability to broaden access, safeguard rights, drive social mobility, and support learning across evolving careers.

1 Introduction

Higher education is entering an era in which learning can no longer be limited to early adulthood. Demographic change, digital transformation and technological disruption require individuals to update their knowledge and skills throughout their lives.[1] At the same time, lifelong learning should not be reduced to employability alone but rather recognised for its broader role in supporting participation, agency, and social inclusion across the life course.[2]

In this context, universities are expected to provide flexible and inclusive learning opportunities for adults at different stages of their personal and professional trajectories. Existing lifelong learning provision often reaches employees, organisations and university staff more consistently than groups such as migrants, people in remote or deprived areas, early school-leavers, and persons with disabilities.[3] Artificial intelligence may help address some of these barriers through adaptive support, accessible formats and delivery at scale, but it may also reproduce or exacerbate inequalities through unequal access, limited AI literacy, algorithmic bias, opaque decision-making, and weak institutional capacity.[4,5]

AI should therefore be treated not only as a pedagogical innovation but as a policy and institutional governance issue. Whether it contributes to more equitable lifelong learning depends not only on whether AI is adopted, but on the institutional, financial and regulatory conditions under which it is developed and used.

Research on AI in higher education has concentrated primarily on pedagogical and learner-facing applications, including profiling, prediction, personalised learning, assessment and academic integrity.[6,7] At the same time, AI is also a governance issue in education because decisions about its adoption affect who can access new learning opportunities, how data are collected and used, who bears the risks, and who is accountable for decisions.[2] Although more literature on responsible AI governance is emerging, it remains fragmented.[8]

1.1 Strategic function of the white paper: Purpose, scope and target audience

Technological innovation alone will not deliver inclusive outcomes. Universities also require appropriate funding, flexible pathways and credential systems, recognition of prior and non-formal learning, quality assurance, and responsible institutional governance. AI can support equitable lifelong learning only when technological innovation aligns with accountable, inclusive and well-resourced institutional governance. The policy question is, therefore, not only how universities can use AI for lifelong learning, but what conditions are needed to ensure that AI expands access, protects rights and produces meaningful learning opportunities for those who have traditionally been underserved.

The primary data analysed here comprise peer-reviewed journal articles, institutional reports, higher education policy documents, and previous white papers. This study prioritises publications addressing higher education policy, AI-enabled lifelong learning, and equitable reskilling in Europe, Asia and beyond. The findings are synthesised into actionable policy reform recommendations to help universities scale equitable AI-driven lifelong learning and reskilling initiatives.

2 Conceptual framework

2.1. AI and the transformation of higher education

Lifelong learning and reskilling are critical across Asia and Europe due to evolving labour demands and access barriers. The European Commission's Union of Skills initiatives expect higher education institutions (HEIs) to play an active role in upskilling, reskilling, and lifelong learning.[9] Similarly, the Association of Southeast Asian Nations (ASEAN) Qualifications Reference Framework (AQRF) focuses on validating non-formal learning via flexible modules and micro-credentials.[10] The Digital Competence Framework for Citizens,[11] the Digital Education Action Plan in Europe,[12] and the ASEAN Work Plan on Education[13] emphasise use of AI for lifelong learning.

AI may strengthen this transition by supporting personalised learning and reducing inequalities through accessible, inclusive support for underrepresented groups.[14] At the same time, its benefits cannot be assumed to be equitable. For example, a lack of social contact in the learning experience or teachers' overreliance on AI for instruction may fail to engage students in personalised learning, thereby exacerbating educational inequities. Unequal access to AI tools and AI literacy negatively affect students' academic performance and undermine their career prospects.[15,16,17,18]

Evidence also suggests that institutional readiness remains uneven. For example, a global United Nations Educational, Scientific and Cultural Organization (UNESCO) survey of higher education stakeholders across 90 countries found substantial uncertainty about the effective pedagogical and research uses of AI, alongside limited understanding of its technological, ethical, and social implications.[19] Key findings from a UNESCO report highlight both the expansion of higher education and the persistent inequalities that accompany this growth. Globally, the tertiary education enrolment ratio more than doubled, from 19% in 2000 to 44% in 2024, while the graduation ratio increased at less than half that rate, from 17% to 27%.[20] This gap shows the limitations of using enrolment alone as an indicator of educational progress, as increased access to higher education does not necessarily translate into successful completion. Moreover, this expansion has not benefited all population groups equally, with disadvantaged groups continuing to face structural barriers to participation and completion. For lifelong learning, simply broadening access is insufficient unless accompanied by targeted support to guarantee equitable completion.

These findings reinforce the policy need to combine innovation with equity, inclusion, capacity building, and responsible governance. For university leaders, the question is not simply whether AI can improve teaching or institutional efficiency, but whether it is designed, implemented and governed in ways that serve the public interest.

2.2. Lifelong learning and reskilling in the AI era

Lifelong learning refers to the continuous development of knowledge and skills throughout the life course in response to changing personal, social, economic and technological circumstances, while reskilling focuses more specifically on acquiring new competencies in response to changing occupational and labour market demands.[21] Both concepts are interrelated and interdependent. Reskilling requires access to and participation in lifelong learning opportunities.

The shift toward a knowledge-based, rapidly evolving global economy is pushing traditional, front-loaded education models into obsolescence. A robust framework for lifelong learning and reskilling is no longer optional; it is an economic and social imperative. AI and other emerging technologies are transforming occupations and workplace practices, while a rapidly ageing population is increasingly expected to remain economically and socially active for more years.[22,23,24]

Regional policy frameworks actively reflect this agenda. ASEAN initiatives advance flexible learning, credential recognition, and labour mobility[13] while Europe integrates lifelong learning into its Pillar of Social Rights[25] and relies on agencies like the European Centre for the Development of Vocational Training (Cedefop) for lifelong career guidance.

Beyond driving economic growth, lifelong learning fosters community integration, reduces isolation,[26] and strengthens democratic values through critical thinking

and empathy. Globally, this aligns with SDG Target 4.4, which emphasises equipping youth and adults with technical, vocational and digital skills for decent work.[27]

Looking ahead, lifelong learning systems will need to be flexible, personalised, and interconnected. Based on this idea, Cedefop[28] proposes five policy scenarios for lifelong and lifewide learning in 2040 to help EU and national policymakers respond to future choices. The study identifies key trends, including digitalisation, modularisation and quality assurance, that have influenced the transparency and transferability of learning outcomes over the past two decades, while highlighting ongoing challenges such as system fragmentation and unequal access.

2.3. Universities as drivers of equitable lifelong learning

Higher education is widely recognised as a public good and an integral part of the right to education. However, in many contexts, public resources have historically been limited, thereby restricting access to higher education for broader segments of the population.

According to a report by the UNESCO Institute of Lifelong Learning,[3] 66% of HEIs align their lifelong learning policies with the SDGs, 61% prioritise gender equality, and 74% view community engagement as a primary driver. Consequently, the report urges national governments to make lifelong learning a core university mission. By designing inclusive pathways, universities can break socioeconomic barriers, promote social mobility, and drive civic engagement alongside workforce development.

3 Equity and inclusion challenges

3.1. The digital divide and unequal access to AI technologies

The structural integration of AI promises significant opportunities for personalised learning, yet it simultaneously threatens to widen the historic digital divide into a structural AI divide. In higher education, the divide has evolved into a multi-dimensional disparity in access to advanced proprietary AI models, computing power, data infrastructure and AI literacy. Universities risk creating more unequal learning conditions, where well-funded institutions and affluent learners leverage advanced, subscription-based AI tools, while disadvantaged lifelong learners are relegated to outdated or less effective open tools. UNESCO's policy work has emphasised that without deliberate public intervention and robust infrastructure safeguards, the rapid digitalisation of education via generative AI will reinforce existing disparities rather than alleviate them, creating an uneven landscape for upskilling and career development.[29]

To prevent this polarisation, universities must treat access to quality educational AI tools as an essential element of digital inclusion, ensuring structural barriers do not prevent lower-income workers from accessing digital competences required by today's labour market.

3.2. Underrepresented and vulnerable groups

When designing AI-driven adult education systems, HEIs must account for the distinct barriers faced by underrepresented and vulnerable learner groups. This includes older displaced workers, rural populations, individuals with disabilities, and migrant groups whose learning pathways are often non-linear. The digital gender divide also manifests within this context. A recent study indicates significant differences in confidence levels, training opportunities, and AI skills, with only 28% of female higher education students reporting proficient AI skills compared to 73% of their male peers.[30] For older or other vulnerable adults seeking to reskill, commercial AI learning platforms can present accessibility challenges if they do not adequately address universal design, linguistic diversity or different levels of digital literacy.

If universities deploy automated adaptive-learning tools without considering the social, educational and regional contexts of non-traditional students, they risk excluding the groups that may benefit the most from accessible lifelong learning opportunities.

3.3. Barriers to equitable participation in lifelong learning

While AI-enabled learning solutions offer opportunities to expand adult reskilling, significant systemic barriers can still restrict marginalised workers from participating equitably in lifelong learning. Beyond basic access to technology, adult learners face time constraints,

financial pressures, and institutional rigidities that limit their ability to engage with AI-supported educational platforms. Low- and middle-wage workers, who are often most vulnerable to AI-driven job displacement, balance full-time employment, caregiving obligations, and unpredictable work schedules, leaving little capacity for traditional or even self-paced online coursework. Furthermore, institutional entry barriers, such as rigid prerequisite requirements, the high cost of subscription models for advanced AI learning tools, and a lack of formal recognition for informal work experience create structural barriers for non-traditional students seeking higher education pathways. Under the European Union (EU) AI Act, certain AI systems used in educational settings for admissions, evaluation or tracking are recognised as high-risk systems, emphasising that technology must not automate exclusion or exacerbate socioeconomic divides.[31]

To promote genuine social mobility, universities must actively address these participation barriers to enable adults to engage in continuous skill development without risking financial or personal instability.

4 Institutional AI governance in universities

Universities serve as key intermediaries in translating national and regional AI policies, such as the EU AI Act and ASEAN frameworks on digital transformation and AI governance, into ethical, context-sensitive institutional practices. In doing so, universities move beyond regulatory compliance to become active agents in building trustworthy AI environments that promote equity, transparency, inclusion, and educational quality. By aligning policies, strengthening equity, and addressing algorithmic challenges, universities can build trustworthy and inclusive AI environments that support equitable, high-quality and lifelong education.[32, 33]

4.1 Institutional AI policies adopted by universities

The adoption of institutional AI policies in universities has shifted from fragmented and often reactive responses to more comprehensive governance frameworks. After AI emerged, many institutions imposed restrictions, creating inconsistent policies and regulatory gaps.[34] However, these early approaches soon proved difficult to sustain, as prohibitions neither prevented AI use nor addressed the ethical, pedagogical, and equity challenges associated with its adoption.

This evolution reflects a process of policy development and layering, where universities preserve academic freedom while gradually updating their governance and academic integrity frameworks to meet the challenges posed by AI and rapidly advancing educational technologies.[35] Rather than relying solely on centralised regulation or individual discretion, most institutions have adopted governance models that balance centralised oversight of legal, data and procurement matters with faculty autonomy over the pedagogical use of AI in the classroom.[33] Regional approaches across Asia and Europe illustrate how such governance models can address equity and inclusion in different institutional and regulatory contexts.[36]

To examine how institutional AI governance takes shape across different policy environments, this white paper compares three broad regional contexts: the EU, ASEAN and selected East Asian contexts, and the United Kingdom. The analytical categories in Table 1 (see Appendices) are based on six key dimensions of AI governance in higher education: regulation, institutional authority, data protection, pedagogical flexibility, equity, and inclusion. This comparison matters for equitable lifelong learning because governance shapes how institutions address AI-related risks, access, participation and accountability.

The comparison between rights-based and decentralised AI governance highlights the need for agile institutional frameworks in higher education. The three categories capture broad policy tendencies, but considerable variation exists across countries and institutions in regulatory maturity, institutional autonomy, and implementation capacity. The purpose of the comparison is not to rank these approaches, but to identify different policy choices and trade-offs that universities and policymakers face when designing institutional AI governance. Evidence from the EU suggests that effective governance can balance legal and ethical safeguards with flexibility for faculty to adapt AI to their teaching practices.[32] In the United Kingdom, a sector-led approach promotes institutional innovation and autonomy, but

evidence also reveals equity challenges and points to gaps in AI literacy and support: although 67% of students consider digital skills essential, only 36% report receiving formal AI training.[37] Similarly, Malaysia's approach aligns university autonomy with ASEAN priorities, though disparities in infrastructure, funding and digital capacity across the region may affect the consistency and equity of AI policy implementation.[32]

4.2. Common policy priorities and areas of action

4.2.1 Towards algorithmic fairness

Ensuring that AI systems are fair, transparent and subject to meaningful human oversight should be a central policy priority for universities adopting predictive AI systems.[38] As HEIs rely on predictive analytics to identify students at risk of poor performance or dropout, algorithmic systems may inadvertently reinforce existing racial, socioeconomic and gender inequalities through biases present in historical administrative data.[39,40] Because predictive models learn from patterns present in previous institutional practices, they may reproduce systemic discrimination unless universities examine the limitations and representativeness of the datasets on which these systems are trained.[41,42] Addressing these risks requires combining pre-processing, in-processing and post-processing mitigation techniques with human-in-the-loop governance mechanisms that enable institutions to review algorithmic outcomes, assess their implications, and refine fairness objectives over time.[42]

Emerging evidence indicates that automation bias and professional complacency may lead educators to accept incorrect AI-generated outputs despite expressing limited trust in these systems. This phenomenon, often described as the "chauffeur problem", illustrates how excessive reliance on AI can gradually weaken professional judgement and critical decision-making.[43] In this sense, human oversight should function not simply as a procedural safeguard but as an active process of critical evaluation in which academic staff retain responsibility for interpreting, challenging and, when necessary, overriding algorithmic recommendations.[44]

The growing use of predictive systems, automated assessment, and learning analytics has intensified concerns regarding the opacity of black-box algorithms, whose internal reasoning often remains inaccessible to educators and students.[45] Explainable AI (XAI) provides mechanisms to make AI outputs more interpretable, strengthening institutional accountability, facilitating auditing processes, and supporting more informed human decision-making.[46] However, XAI should be understood as one component of broader institutional governance rather than as a standalone solution to algorithmic bias.

Beyond improving transparency, XAI may help identify and interpret linguistic and cultural biases in AI systems trained predominantly on English-language datasets. Such systems may disadvantage multilingual students, speakers of low-resource languages, and learners whose rhetorical conventions differ from dominant Western academic norms.[47] Similar concerns have been reported for AI-generated text detectors, which may disproportionately misclassify writing produced by non-native English speakers, potentially influencing assessment outcomes and reinforcing existing educational inequalities.[48,49] Incorporating XAI into institutional auditing procedures, therefore, enables universities not only to identify discriminatory decision patterns but also to ensure that assessment practices evaluate academic quality rather than linguistic or cultural variation.[45]

4.2.2 Data sovereignty, linguistic diversity, and decolonial governance

Building more inclusive AI environments should constitute a second strategic priority for HEIs. While AI can expand educational access and personalise learning, these benefits will remain unevenly distributed unless universities address structural inequalities in data collection, model development, and knowledge production.[41,42]

A key challenge concerns data inclusion and sovereignty. The rapid adoption of AI in higher education may widen existing inequalities when datasets, models and technological infrastructures fail to represent diverse linguistic and cultural contexts.[41] Although AI systems offer multilingual capabilities, they still rely predominantly

on English-language and other high-resource datasets, often overlooking low-resource, indigenous, and regional languages.[49] Extractive practices such as "helicopter research" and persistent Western-centric pedagogical assumptions may marginalise historically underrepresented communities, highlighting the need for more equitable, participatory, and culturally responsive AI governance.[50] Universities should therefore incorporate linguistic diversity and data sovereignty as explicit criteria in the procurement, evaluation and deployment of AI technologies, prioritising systems that support multilingual education, preserve local linguistic heritage, and respect communities' rights over their data.

Beyond linguistic representation, universities must also address questions of data governance and digital sovereignty. Traditional open-data frameworks, such as the FAIR Guiding Principles, have improved the findability, accessibility, interoperability, and reusability of research data. However, these principles alone are insufficient for contexts involving indigenous and marginalised communities because they pay limited attention to issues of ownership, collective rights, and cultural responsibility.[51] The CARE Principles complement FAIR by emphasising collective benefit, authority to control, responsibility and ethics, shifting attention from data management toward more equitable forms of data governance. Rather than viewing these frameworks as competing approaches, universities should integrate them to ensure that educational datasets are both technically interoperable and socially responsible.



5 Policy recommendations

This section organises the proposed policy reforms around six complementary functions: Enable, Transform, Govern, Connect, Include, and Implement & Evaluate. The sequence reflects the "reform rationale" developed throughout the white paper:

1. National and regional policy must first enable equitable AI-enabled lifelong learning through supportive funding, infrastructure, recognition, and quality assurance frameworks.
2. Universities must then transform their educational offer while also governing AI through explainable, inclusive, accountable, and equitable institutional arrangements.
3. These reforms must be supported by stronger external relationships that connect universities with employers, communities, and public authorities, while ensuring that implementation includes those most at risk of exclusion.

The functions reinforce one another across levels. Funding and recognition (Enable) make flexible provision possible (Transform), while quality assurance frameworks set the rules within which institutions govern AI (Govern). Governance in turn protects the learners most exposed to bias or opacity, who are the focus of Include. Partnerships (Connect) bring labour market needs into curricula and reach groups that institutions cannot reach alone, and Implement & Evaluate feeds evidence on outcomes back into every function. Taken in isolation, each function falls short, as funding without recognition stalls, governance without partners stays inside the institution, and equity without measurement remains a promise.

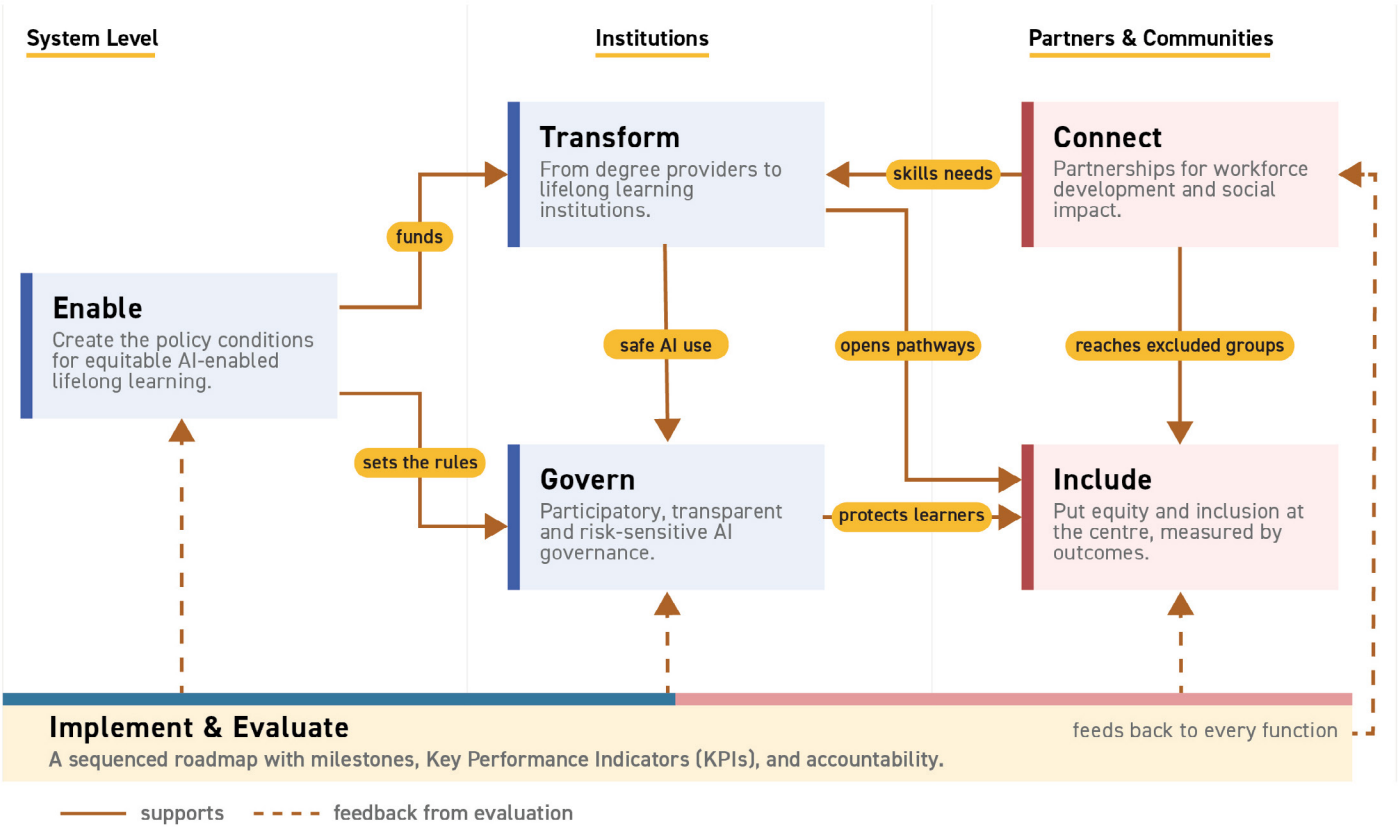


Figure 1. *Functions guiding policy reform.*

The six policy functions should be understood as interconnected dimensions of a broader reform process rather than as isolated areas of action. While each function addresses a distinct policy need, their effectiveness depends on their alignment across levels and institutional contexts. The roadmap developed in Section 6 translates these functions into a sequence of actions, responsibilities, and indicators to support implementation, monitoring, and continuous improvement.

5.1. Enabling policy conditions for equitable AI-enabled lifelong learning

As discussed in Sections 1-3, universities cannot expand equitable AI-enabled lifelong learning without supportive system-level conditions. University funding still centres on full degree programmes, even as demand for shorter, more flexible learning options continues to grow.[52] At the same time, micro-credentials are not yet consistently recognised across systems, which can limit their value for learners and employers. Before universities can count micro-credentials towards larger qualifications, clear mechanisms are needed to establish their credit value, quality, and portability. The EU has begun to address this through common definitions and shared standards for micro-credentials across its member states.[53]

Recommendations:

- Fund micro-credentials and short AI-related programmes directly, including through public funding mechanisms such as Individual Learning Accounts (ILAs);
- Establish recognition frameworks for micro-credentials, with clear credit values, common standards, and pathways for stacking towards formal qualifications and transferring credits across institutions and borders;
- Invest in shared and open AI infrastructure, including affordable access to high-quality digital tools and Open Educational Resources (OERs), particularly for disadvantaged communities;
- Align quality assurance and regulatory frameworks with flexible provision to safeguard standards, transparency and learner protection.

5.2. From degree providers to lifelong learning institutions

As outlined in Sections 1-3, universities need to move beyond conventional degree structures to more flexible lifelong learning models. Adult and non-traditional learners often face rigid entry requirements, limited recognition of prior experience, financial pressures and time constraints, while lifelong learning provision does not always reach those most excluded from higher education.[3] Modular, flexible and stackable provision can widen participation and support continuous reskilling, as reflected in European policy frameworks for lifelong learning and micro-credentials.[28,53]

Recommendations:

- Develop modular and stackable learning pathways that allow learners to build larger qualifications through shorter periods of study;
- Expand recognition of prior and non-formal learning, including professional experience and previous training to support admission and progression;
- Enable flexible entry, re-entry and progression, allowing adults to move in and out of higher education as their circumstances change;
- Ensure flexible provision remains affordable and accessible, particularly for disadvantaged learners with disabilities and other non-traditional learners.

5.3. Strengthening institutional AI governance for equity and accountability

As discussed in Section 4, universities are moving from fragmented and reactive approaches to AI towards more structured governance models, yet important gaps remain in participation, accountability, transparency, and equity. These gaps are especially significant where AI is used in high-impact areas such as admissions, predictive analytics, assessment and student support. Existing work highlights the need for participatory governance, risk-sensitive procurement, meaningful "human-in-the-loop" oversight, explainability and stronger safeguards against bias and exclusion.[48] Section 4 also shows that responsible governance must extend to data sovereignty, linguistic diversity, and the protection of underrepresented communities.[51]

Recommendations:

- Establish participatory AI governance structures that involve staff, students, technical and legal experts, and affected communities;
- Introduce risk-sensitive reviews of educational AI tools, covering bias, privacy, accessibility, and data governance before deployment;

- Require meaningful human oversight and explainability for high-impact AI decisions, with mechanisms for audit, appeal and redress;
- Integrate FAIR and CARE principles into data governance, with attention to linguistic diversity, community rights, and non-extractive data use;
- Prioritise governed autonomy, authentic assessment, and inclusive design over punitive uses of AI, ensuring that systems meet the needs of diverse and non-traditional learners.

5.4. Partnerships for workforce development and social impact

As discussed in Section 2, universities' contribution to lifelong learning extends beyond formal education to workforce development, community engagement, and social responsibility. Effective reskilling, thus, requires stronger connections between universities, employers, public authorities, and community organisations, particularly where institutions seek to identify emerging skills needs and reach groups that remain underserved by conventional provision. Community engagement is already recognised as an important driver of higher education lifelong learning, reinforcing the need to position universities within broader learning and innovation networks rather than as isolated providers.[3]

Recommendations:

- Build partnerships with employers, public authorities, and social partners to identify emerging skills needs and develop relevant reskilling and upskilling opportunities;
- Work with community and civil-society organisations to reach underserved and underrepresented groups, ensuring that lifelong learning provision responds to local needs and support social inclusion;
- Create transnational networks for AI, lifelong learning, and skills to connect universities, researchers, and policymakers across Asia and Europe, share evidence and good practices, and strengthen collaboration across regions;
- Share expertise and resources across institutions and sectors to reduce duplication and extend access to knowledge and capacity where resources are limited.

5.5. Putting equity and inclusion at the centre of implementation

As shown in Sections 3 and 4, the benefits of AI-enabled lifelong learning are unlikely to be distributed evenly without deliberate intervention. Older and displaced workers, rural populations, migrants, persons with special needs, multilingual learners and other non-traditional groups may face barriers related to affordability, accessibility, digital skills, language and institutional rigidity. At the same time, inequities can be amplified through inaccessible platforms, biased datasets, and algorithmic systems that fail to represent diverse linguistic and cultural contexts.[31,54]

Recommendations:

- Make equity an implementation criterion, assessing whether AI-enabled provision reduces participation gaps;
- Improve access to affordable and accessible digital infrastructure and AI training, particularly for at-risk workers and underserved groups;
- Prioritise underrepresented groups in public university lifelong learning provision, with attention to linguistic and cultural inclusion;
- Embed accessibility and universal design in AI procurement and service design, involving marginalised communities in policy design, implementation and evaluation;
- Monitor participation, retention, completion and outcomes across learner groups, using disaggregated data to identify and mitigate persistent inequalities.

6 Proposed implementation roadmap for equitable AI-enabled lifelong learning in higher education

This section translates the policy rationale developed in Section 5 into an implementation roadmap. Table 2 (see Appendices) shows how these functions can be sequenced, coordinated, and monitored over time, with clear milestones, responsible actors, and indicators of progress. The implementation roadmap is organised into three phases: Enable (0-12 months), Transform (1-3 years) and Sustain (3-5 years). The phases represent progression from establishing conditions, governance safeguards, and baseline equity assessments to consolidating, evaluating, and continuously improving reforms. The six functions identified in Section 5 (Enable, Transform, Govern, Include, and Implement & Evaluate) are addressed across the three phases.

Each phase is defined through key actions, milestones and KPIs that can be used to monitor implementation, assess progress, and identify areas requiring adjustment. The indicators combine measures of provision, participation, institutional capacity, equity, quality, and outcomes, recognising that increased access alone does not demonstrate equitable or sustainable change.



7 Conclusion

Equitable AI-enabled lifelong learning requires coordinated reform across multiple levels: national and regional systems must create enabling conditions; universities must redesign learning pathways and institutional practices; AI governance must become explainable, participatory, and accountable; partnerships must connect higher education with wider social and economic networks; equity must move from an aspirational principle to a measurable outcome; and these reforms must be implemented and evaluated through a sequenced roadmap of monitoring, accountability and continuous improvement. These six functions are mutually reinforcing and should not be implemented in isolation.

Meeting this challenge also requires universities to reaffirm their public mission. Lifelong learning should not be reduced to short-term employability or technological adaptation alone. It should also support social mobility, participation, agency and individuals' right to continue learning throughout their lives. In this sense, AI should be understood as an enabling resource within a broader educational and social project, rather than as the driver of reform itself. The implementation roadmap proposed in Section 6 reflects this perspective by treating reform as a phased, continuous process: first creating enabling conditions and safeguards, then transforming provision and partnerships, and finally integrating, monitoring and improving new practices over time.

In the end, the success of AI-enabled lifelong learning should be judged by a straightforward question: does it widen meaningful access to high-quality learning while reducing, rather than amplifying, existing inequalities? If this principle remains central to policy design, institutional governance and evaluation, AI can become a means to strengthening universities' capacity to provide inclusive, flexible and socially accountable lifelong learning.

Authors

Prof Lampros STERGIOULAS
The Hague University of Applied Sciences, Netherlands

Dr Cinta GALLENT-TORRES
Universitat de València, REDIE-CYTED, Spain

Dr Munir Moosa SEWANI
University of Malta, Malta

Mr Anshumaan GOEL
O.P. Jindal Global University, India

Dr Katerina MAKRI
University of Thessaly, Greece

Mr David TIMIS
Global Governance Institute, Belgium



References

1. OECD. (2023). *Public policies for effective micro-credential learning* (OECD Education Policy Perspectives No. 85). https://doi.org/10.1787/a41f148b-en

2. UNESCO. (2021). *Reimagining our futures together: A new social contract for education*. https://www.unesco.org/en/articles/reimagining-our-futures-together-new-social-contract-education

3. UNESCO Institute for Lifelong Learning. (2022). *The contribution of higher education institutions to lifelong learning: Global survey report presented to the UNESCO World Higher Education Conference (WHEC2022)*. UNESCO. https://www.uil.unesco.org/en/publication/contribution-higher-education-institutions-lifelong-learning-global-survey-report-presented-unesco

4. UNESCO. (2023). *Guidance for generative AI in education and research*. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

5. OECD. (2024). *The potential impact of artificial intelligence on equity and inclusion in education*. https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/08/the-potential-impact-of-artificial-intelligence-on-equity-and-inclusion-in-education_0d7e9e00/15df715b-en.pdf

6. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, *16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

7. Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, *21*, Article 4. https://doi.org/10.1186/s41239-023-00436-z

8. Alfiras, M. I. I., Emran, A. Q., & Mohamed, A. M. (2026). Ethics and governance of generative AI in education: A systematic review on responsible adoption. *Discover Education*, *5*, Article 37. https://doi.org/10.1007/s44217-025-01051-y

9. European Commission. (2025a). *The Union of Skills*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX%3A52025DC0090

10. ASEAN. (2015). *ASEAN Qualifications Reference Framework*. https://asean.org/wp-content/uploads/2017/03/ED-02-ASEAN-Qualifications-Reference-Framework-January-2016.pdf

11. European Union. (2022). *The digital competence framework for citizens*. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC128415/JRC128415_01.pdf

12. European Commission. (2021). *Digital Education Action Plan: Policy background*. https://education.ec.europa.eu/focus-topics/digital-education/action-plan

13. ASEAN. (2026). *ASEAN Work Plan on Education 2026-2030*. https://asean.org/book/asean-work-plan-on-education-2026-2030/

14. Apata, O. E., Ajose, S. T., Apata, B. O., Olaitan, G. I., Oyewole, P. O., Ogunwale, O. M., Oladipo, E. T., Oyeniran, D. O., Awoyemi, I. D., Ajobiewe, J. O., Ajamobe, J. O., Appiah, I., Fakhrou, A. A., & Feyijimi, T. (2025). Artificial intelligence in higher education: A systematic review of contributions to SDG 4 (quality education) and SDG 10 (reduced inequality). *International Journal of Educational Management*, *40*(1–2), 336–353. https://doi.org/10.1108/IJEM-12-2024-0856

15. Halat, R., & Rahme, L. K. (2024). *Addressing inequities in education: AI as a double-edged sword (Part II)*. Harvard Graduate School of Education. https://mepli.gse.harvard.edu/our-fellows-at-work/addressing-inequities-in-education-ai-as-a-double-edged-sword-2/

16. Knox, J. (2025). *Impact, inequality, and imagination: Envisioning a more ambitious education in the era of AI*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000393197

17. Molina, E., & Medina, E. (2025). *AI revolution in higher education: What you need to know*. The World Bank. https://documents1.worldbank.org/curated/en/099757104152527995/pdf/IDU-b1e5ef00-75ff-4ba4-a4b6-84899c3ea968.pdf

18. Pham, H., Kohli, T., Liano, E., Nokuri, I., & Weinstock, A. (2024). *How will AI impact racial disparities in education?* Stanford Law School. https://law.stanford.edu/2024/06/29/how-will-ai-impact-racial-disparities-in-education/

19. UNESCO. (2025a). *UNESCO survey: Two-thirds of higher education institutions have or are developing guidance on AI use*. https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use?hub=84636

20. UNESCO. (2026a). *Global Education Monitoring Report 2026: Access and Equity – Countdown to 2030*. UNESCO. https://doi.org/10.54676/JLKL3223

21. UNESCO. (2026b). *Lifelong learning*. https://www.unesco.org/en/query-list/l/lifelong-learning

22. Ehlers, U.-D. (2017). *Future-ready universities and graduates: Quality education beyond the horizon* (ARC6 Commissioned Paper). Asia-Europe Foundation. https://asef.org/wp-content/uploads/2020/10/ARC6_Commissioned-Paper_EHLERS.pdf

23. ILO. (2021). *Shaping skills and lifelong learning for the future of work* (International Labour Conference 109th Session). https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@ed_norm/@relconf/documents/meetingdocument/wcms_813696.pdf

24. ILO. (2026). *Lifelong learning and skills for the future: Insights for Asia and the Pacific* (Research brief). https://researchrepository.ilo.org/esploro/outputs/encyclopediaEntry/Lifelong-learning-and-skills-for-the/995699970202676/filesAndLinks?institution=41ILO_INST&index=0

25. European Commission. (2017). *European Pillar of Social Rights*. https://commission.europa.eu/system/files/2017-11/social-summit-european-pillar-social-rights-booklet_en.pdf

26. Dickson, M., Cairns, D., & Areepattamannil, S. (2020). *Why and how should educational systems support lifelong learning?* (Research paper). Regional Center for Educational Planning. https://rcepunesco.ae/en/KnowledgeCorner/PolicySummaries/PolicySummaries/05-11-2020_EN.pdf

27. United Nations. (2015). *The 17 goals*. https://sdgs.un.org/goals

28. Cedefop. (2026). *Shaping the future of lifelong learning: Policy scenarios for 2040. Transparency and transferability of learning outcomes* (Research paper). Publications Office of the European Union. https://www.cedefop.europa.eu/en/publications/5619

References

29. UNESCO. (2025b). *AI and education: Protecting the rights of learners*. UNESCO Digital Library. https://www.unesco.org/en/articles/ai-and-education-protecting-rights-learners

30. UNESCO-ICHEI. (2026). *International Women's Day 2026: Shaping an equitable future for higher education through digital empowerment*. International Centre for Higher Education Innovation under the auspices of UNESCO. https://en.ichei.org/en/news/information/4056.html

31. European Commission. (2026). *2026 European Commission work programme includes actions in disability strategy, education, quality jobs, mental health, and AI*. European Association of Service Providers for Persons with Disabilities (EPR). https://www.epr.eu/2026-european-commission-work-programme-confirms-upcoming-disability-strategy-actions-actions-on-education-quality-jobs-mental-health-and-ai/

32. European Commission. (2025b). *Commission Guidelines on prohibited artificial intelligencepractices established by Regulation* (EU) 2024/1689 (AI Act). https://ai-act-service-desk.ec.europa.eu/sites/default/files/2025-08/guidelines_on_prohibited_artificial_intelligence_practices_established_by_regulation_eu_20241689_ai_act_english_ied3r5nwo50xggpcfmwckm3nuc_112367-1.PDF

33. Şen, E., Vaněček, D., & Adnan, M. (2026). Balancing regulation and innovation: The need for agile AI governance in higher education – A cross-country study. *Studies in Higher Education*. Advance online publication. https://doi.org/10.1080/03075079.2026.2614986

34. Azevedo, L., Robles, P., Best, E., & Mallinson, D. J. (2025). Institutional policies on artificial intelligence in higher education: Frameworks and best practices for faculty. *New Directions for Adult and Continuing Education*, *188*, 70–78. https://doi.org/10.1002/ace.70013

35. Coates, H., Croucher, G., & Calderon, A. (2025). Governing academic integrity: Ensuring the authenticity of higher thinking in the era of generative artificial intelligence. *Journal of Academic Ethics*, *23*(4), 2015–2028. https://doi.org/10.1007/s10805-025-09639-7

36. Illingworth, S. (2026). *What UK university AI policies actually do: A study of 96 institutions*. Higher Education Policy Institute (HEPI). https://www.hepi.ac.uk/reports/what-uk-university-ai-policies-actually-do-a-study-of-96-institutions/

37. Freeman, J. (2025). *Student Generative AI Survey 2025.* Higher Education Policy Institute (HEPI). HEPI Policy Note 61. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/

38. Quality Assurance Agency (QAA). (2024). *Authentic assessment design and values-driven academic integrity frameworks*.

39. Jisc. (2025). Digital skills and student formal training survey insights. https://digitalinsights.jisc.ac.uk/reports-and-briefings/our-reports/2024-25-uk-higher-education-students-digital-experience-insights-survey-findings/

40. Russell Group. (2023). *Russell Group principles on the use of generative AI tools in education*. https://www.russellgroup.ac.uk/sites/default/files/2025-01/Russell%20Group%20principles%20on%20generative%20AI%20in%20education.pdf

41. OECD (2023), *OECD Digital Education Outlook 2023: Towards an Effective Digital Education Ecosystem*. OECD Publishing, Paris, https://doi.org/10.1787/c74f03de-en

42. Parnell, A., Jones, D., Wesaw, A., & Brooks, D. C. (2018). *Institutions' use of data and analytics for student success: Results from a national landscape analysis*. NASPA–Student Affairs Administrators in Higher Education; Association for Institutional Research; EDUCAUSE.

43. Passi, S., & Vorvoreanu. (2022). *Overreliance on AI: Literature review*. Microsoft Research. https://www.microsoft.com/en-us/research/wp-content/uploads/2022/06/Aether-Overreliance-on-AI-Review-Final-6.21.22.pdf

44. Dolan, E. (2026). *Teachers say they distrust AI but still accept its harsh grading mistakes, study finds*. PsyPost. https://www.psypost.org/teachers-say-they-distrust-ai-but-still-accept-its-harsh-grading-mistakes-study-finds/

45. Gillani, N., Eynon, R., Chiabaut, C., & Finkel, K. (2023). Unpacking the "black box" of AI in education. *Educational Technology & Society*, *26*(1), 99–111. https://www.jstor.org/stable/48707970

46. Sele, D., & Chugunova, M. (2024). Putting a human in the loop: Increasing uptake, but decreasing accuracy of automated decision-making. *PLoS ONE*, *19*(2), Article e0298037. https://doi.org/10.1371/journal.pone.0298037

47. Chesterman, S. (2021). *We, the robots? Regulating artificial intelligence and the limits of the law*. Cambridge University Press.

48. Ribes-Lafoz, M., Navarro-Colorado, B., & Rovira-Collado, J. (2026). Exploring strategies to detect and mitigate bias in AI in education: Students' perceptions and didactic approaches. *Trends in Higher Education*, *5*(2), Article 33. https://doi.org/10.3390/higheredu5020033-

49. Dang, Q. (2024). Beyond transparency: Why explainable AI is essential for educational equity in low-resource language communities. *Journal of Enabling Technologies*, *20*(1), 27–41. https://doi.org/10.1108/JET-07-2025-0052

50. Nyaaba, M., Wright, A. L., & Choi, G. L. (2026). Generative AI and digital neocolonialism in global education: Towards an equitable framework. *Journal of Digital Educational Technology,* 6(1), ep2608. https://doi.org/10.30935/jdet/17862

51. Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for Indigenous data governance. *Data Science Journal*, *19*, Article 43. https://doi.org/10.5334/dsj-2020-043

52. Kato, S., Galán-Muros, V., & Weko, T. (2020). *The emergence of alternative credentials* (OECD Education Working Papers No. 216). OECD Publishing. https://doi.org/10.1787/b741f39e-en

53. Council of the European Union. (2022). *Council recommendation on a European approach to micro-credentials for lifelong learning and employability*. https://eurlex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022H0627(02

54. UNESCO. (2026c). *Nuevas perspectivas sobre ética e integridad académica en la docencia y la investigación universitaria*. https://unesdoc.unesco.org/ark:/48223/pf0000399295

55. Information Commissioner's Office (ICO). (2023). *Data protection impact assessments and UK GDPR compliance guidelines*.


AI
AI
AI

Appendices

Table 1. *Comparative overview of AI governance approaches in higher education*

Dimension	European Union	United Kingdom	ASEAN and Selected East Asian Contexts
Analytical Orientation	Rights and regulation-oriented approach (e.g., Czech Technical University)[33]	Sector-led and principles-oriented approach (e.g., Russell Group / Jisc Framework)	Development- and capacity-oriented approach (e.g., Malaysia Reform Agenda)
Governance Posture	Highly structured and formalised, with institutional responsibilities shaped by the EU AI Act and related regulatory frameworks	Principles-driven, collaborative, and oriented towards innovation, with important reliance on sector guidance and institutional self-regulation[36]	Collaborative and development-oriented, with governance linked to national digital transformation and AI strategies and regional capacity building
Centralisation of Authority	Stronger central regulatory oversight, with responsibilities distributed between EU and national authorities and institutional governance structures	More decentralised and collaborative, with universities and sector bodies playing a significant role in translating principles into institutional practice	Hybrid arrangements, combining national or regional strategic direction with implementation through individual HEIs and, in some contexts, pilot "sandboxes"
Data Protection Standards	Underpinned by EU data protection and AI regulation, including requirements concerning privacy, risk management and human oversight, and strict rules on data sovereignty and consent	Governed by UK GDPR, the Data Protection Act 2018, and sector-level guidance on responsible data use; requires mandatory data protection impact assessments	Aligned with regional ASEAN AI governance principles and national data protection and digital security frameworks; significant variation across countries
Pedagogical Adaptability	Lower perceived staff adaptability due to strict regulatory compliance and fear of administrative penalties	Instructional flexibility; individual faculty retain substantial discretion to redesign authentic assessments[38]	Moderate-to-high adaptability; emphasises curriculum integration, national competitiveness, and digital literacy
Systemic Equity and Inclusion Implications	High structural protection against algorithmic bias, but risks alienating marginalised students through rigid bureaucratic barriers; strong regulatory safeguards for rights and non-discrimination	Promotes universal access and inclusive policy co-creation but risks localised fragmentation in student experiences[38]	Promotes broad digital literacy but faces systemic risks of regional resource disparities and uneven infrastructure deployment

Appendices

Table 2. *Implementation phases, milestones, and key performance indicators (KPIs)*

Phase	Timeline	Main Focus	Key Actions	Milestones	KPIs	Key Actors
Enable	0-12 months	Establish the conditions for equitable AI-enabled lifelong learning	• Develop funding and recognition mechanisms • Establish AI governance and quality assurance frameworks • Assess institutional readiness and equity gaps • Identify infrastructure, accessibility, and digital skills needs • Establish staff development and AI literacy plans	• Governance and quality assurance frameworks adopted • Equity and readiness assessments completed • Infrastructure and accessibility gaps identified • Staff development plans established • Funding and support mechanisms operational	• % of institutions with AI governance frameworks • % completing equity/readiness assessments • % of staff completing training • Infrastructure and accessibility coverage • Availability of funding/support mechanisms	National and regional authorities; higher education institutions; quality assurance agencies; academic and professional stakeholders
Transform	1-3 years	Transform educational provision, institutional capacity, and external partnerships	• Expand flexible and AI-enabled learning provision • Redesign curricula and learning pathways • Strengthen staff AI and pedagogical capacity • Develop recognition mechanisms for lifelong learning • Establish partnerships with employers, communities, and public authorities • Monitor participation and progression across learner groups	• Redesigned learning provision operational • Curriculum transformation targets reached • Staff capacity-building programmes scaled • Recognition mechanisms implemented • Partnerships operational • Equity gaps monitored and addressed	• % of programmes offering flexible/ AI-enabled provision • Participation and completion rates • % of staff participating in professional development • Number and diversity of active partnerships • Recognition/ credit uptake • Participation and completion gaps across learner groups	Higher education institutions; faculties and programme teams; employers; community organisations; public authorities; professional stakeholders
Sustain	3-5 years	Institutionalise, evaluate, and continuously improve reforms	• Integrate successful practices into quality assurance systems • Evaluate educational, equity, and employment outcomes • Review governance and provision based on evidence • Scale effective practices across programmes and institutions • Establish continuous improvement cycles and targets	• Reforms in institutional planning and quality assurance • Evaluation of outcomes completed • Improvement actions adopted • Effective practices scaled • Targets established for the next cycle	• Changes in participation and completion gaps over time • Learner progression and outcomes • Quality and satisfaction indicators • % of improvement actions implemented • Evidence of sustained institutional adoption • Scale and continuity of effective partnerships	Higher education institutions; quality assurance agencies; national and regional authorities; employers; communities; learners and other stakeholders



Humanising Learning in the Age of AI:

Policy Conditions for Pedagogy, Assessment and Learner Formation

AUTHORS
Mira Kartiwi
Andri Ioannou
Zainal Abidin Bin Sanusi
Aneela Yasmin
Joey Aviles
Bogna Liziniewicz
Vanessa Camilleri

This paper helps policymakers, higher education regulators, quality assurance agencies, and institutional leaders across Asia and Europe keep higher education qualifications meaningful as AI reshapes learning. It sets out the policy conditions for universities to redesign pedagogy and assessment so that AI strengthens, rather than replaces, what learners can do themselves.

Executive summary

Generative artificial intelligence (GenAI) can now produce polished essays, reports, and code on students' behalf. Such work is no longer reliable evidence that a student has developed the capabilities it is meant to show. As a result, the link between qualifications and the capabilities they certify is weakening, which puts graduates' readiness for work and public trust in higher education at risk.

This paper introduces the *Humanising Learning Framework* to help policymakers across Asia and Europe—including ministries, higher education regulators, and quality assurance agencies — restore that link. We argue that policy should move from policing AI-produced work to supporting learner formation: what a person becomes through education, including their judgement, agency and sense of responsibility. The framework applies four principles (AI as amplifier, not substitute; formation over performance; proportionate governance; and a cross-civilisational foundation) to pedagogy, assessment, and learner formation through three policy pillars:

- **Pedagogy and assessment reform.** Support universities to design learning in which students think for themselves before using AI and require at least one process-based or conversation-based assessment component, such as oral explanation, staged drafts, or process portfolios.
- **Proportionate oversight.** Set risk-based national rules that protect the integrity of qualifications and fair access to AI, with the strongest safeguards where AI affects grades or progression, and let institutions apply them in ways that fit their own context.
- **Long-term monitoring.** Track students' evaluative judgement, self-regulation, metacognition, and intellectual authorship (the ability to own and defend one's work) over time across cohorts and programmes to check whether qualifications still show what graduates can do.

Humanising learning means designing education so that learners develop the capabilities their qualifications certify: critical judgement, ethical responsibility, problem-solving and agency. This is consistent with UNESCO's human-centred guidance on AI.[1]

This paper contributes to 7th ASEF Higher Education Innovation Laboratory's (ASEFInnoLab7) AI for Institutional Excellence track under the theme "The Role of Universities in Advancing AI for Good," which addresses the AI governance, policies, and infrastructure that universities need to pursue institutional excellence and contribute to the common good. We argue that AI contributes to institutional excellence and the common good in higher education when institutions deploy it in

ways that: (i) strengthen every learner's own judgement, agency and responsibility; (ii) ensure its benefits reach all learners, not only the well-resourced; and (iii) maintain qualifications as trusted public guarantees of graduate capability and competence.

As a team of educators and researchers working across Asia, Europe, and Australia, we draw on complementary traditions: European traditions that emphasise independent critical reasoning and Asian traditions that emphasise ethical character and responsibility to the community. Together, they describe a graduate who can think independently and act responsibly in a world shaped by AI.

National policymakers create the conditions; institutions put them into practice. Neither can achieve humanising learning alone. The key policy question is no longer what learners can produce with AI, but what they can still do themselves.

1 The higher education landscape

Humanising learning treats the learner as a whole person. Dignity, agency, relationships, critical awareness, and judgement are human capacities. Generative artificial intelligence (GenAI) can increasingly imitate how these capacities appear in a piece of work, but it cannot develop them in the learner. The key question is therefore not whether AI is present in education, but whether its use strengthens or replaces the processes through which learners become capable, responsible, and independent.

This distinction matters because assessment is not merely a record of learning; it is also the means by which institutions claim that graduates possess particular knowledge and capabilities. Traditionally, written assignments served as evidence of students' understanding. GenAI challenges this assumption. A polished essay, report, or piece of code may demonstrate effective AI use rather than independent mastery. One study found that students using ChatGPT produced better essays but gained no additional knowledge or transfer, suggesting reduced metacognitive engagement.[2] This indicates that the quality of submitted work may no longer be reliable evidence of what a learner understands or can do alone.

Use of large language models among students was associated with lower cognitive load and shallower reasoning.[3] Moreover, knowledge workers who trusted GenAI more reported less effort in critical thinking.[4] Less effort is not a problem in itself. It becomes a problem when AI removes the intellectual work through which people develop judgement, reasoning, and domain expertise.

The evidence is not uniformly negative. A meta-analysis associated ChatGPT use with improved performance, motivation and higher-order thinking, although it also found reduced mental effort.[5] AI can therefore support or undermine learning depending on task design, user behaviour, educational context, and the role assigned to it. The policy challenge is to promote uses that strengthen learning while limiting those that replace it.

This uncertainty is a reason to act now, because it creates an asymmetry for policymakers. Gaps in capability that stay hidden throughout a degree may be difficult and costly to fix after graduation, whereas safeguards that prove too cautious can be revised. Assessment redesign can strengthen the validity of qualifications under different possible AI futures, while proportionate governance can focus oversight on risks to public trust, learning, and professional competence.

The deeper risk is that institutions respond to AI by changing their rules while leaving teaching and assessment largely unchanged. Universities should therefore rethink pedagogy and design learning with AI around the capabilities they want graduates to develop, rather than allowing external tools to determine how learning takes place.

2 Founding principles: A human-centred foundation

This section sets out four principles for using AI in higher education in ways that protect how learners develop. The recommendations in Section 4 and the indicators in Section 7 build on them.

2.1 Principle 1: AI as amplifier, not substitute

Humanising learning treats AI as a tool that supports human development without replacing the relationships and experiences through which people learn. AI acts as an amplifier when it helps learners explore, question, revise, and extend their own thinking. By contrast, it acts as a substitute when it does the work for them (Figure 1). This is consistent with UNESCO's human-centred approach to GenAI in education.[1]

A simple test applies: Can the learner explain and defend the result? Is there evidence of reasoning, revision, and judgement? Did AI extend the learner's thinking, or replace it? At the institutional level, time saved through AI should be reinvested in feedback, mentoring, and human contact, rather than treated as an opportunity to reduce teaching time.

2.2 Principle 2: Formation takes priority over performance

GenAI can improve the quality of a piece of work without developing the capability the work is meant to show. Humanising learning must therefore protect the learning process, even when this means accepting slower or less polished work. This process includes the productive struggle, judgement, and authorship through which learners become independent thinkers. It can be observed in four capacities: evaluative judgement (the ability to assess work against the standards of a discipline), self-regulation (the ability to direct one's own learning), metacognition (awareness of one's own reasoning), and intellectual authorship (the ability to own and defend one's work).

This matters because a qualification is a statement about what a person knows and can do, not about what they can produce with help. If performance becomes separated from capability, employers, professional bodies, and other education providers may no longer be able to rely on what a qualification tells them. Protecting how capability is developed is therefore essential if qualifications are to remain trusted public guarantees.

2.3 Principle 3: Governance is proportionate to stakes involved

Safeguards should match the consequences of each use of AI. They should be strongest where AI informs decisions that affect a learner's rights, progression, or results, such as grading or academic integrity cases, and lightest where the risk is low, such as using AI to brainstorm or check a draft. This avoids two failures: too little oversight where learners' futures are at stake, and too much regulation of everyday learning. The principle applies at every level, from national policy to the design of a single course.

2.4 Principle 4: A human-centred foundation is cross-civilisational

A humanising approach should draw on more than one educational tradition. European and Asian traditions agree that forming the whole person matters but understand its aims differently: some European traditions stress independent, self-directed judgement, while some Asian traditions give more weight to moral character, relationships, and responsibility to the community (Section 3.3).

Shared principles are never applied on neutral ground. How learners use AI depends on classroom expectations, assessment traditions, attitudes to authority, communication norms, language, and understandings of originality. Implementation must therefore be adapted to local cultures, disciplines, and institutions, not copied from one setting to another.

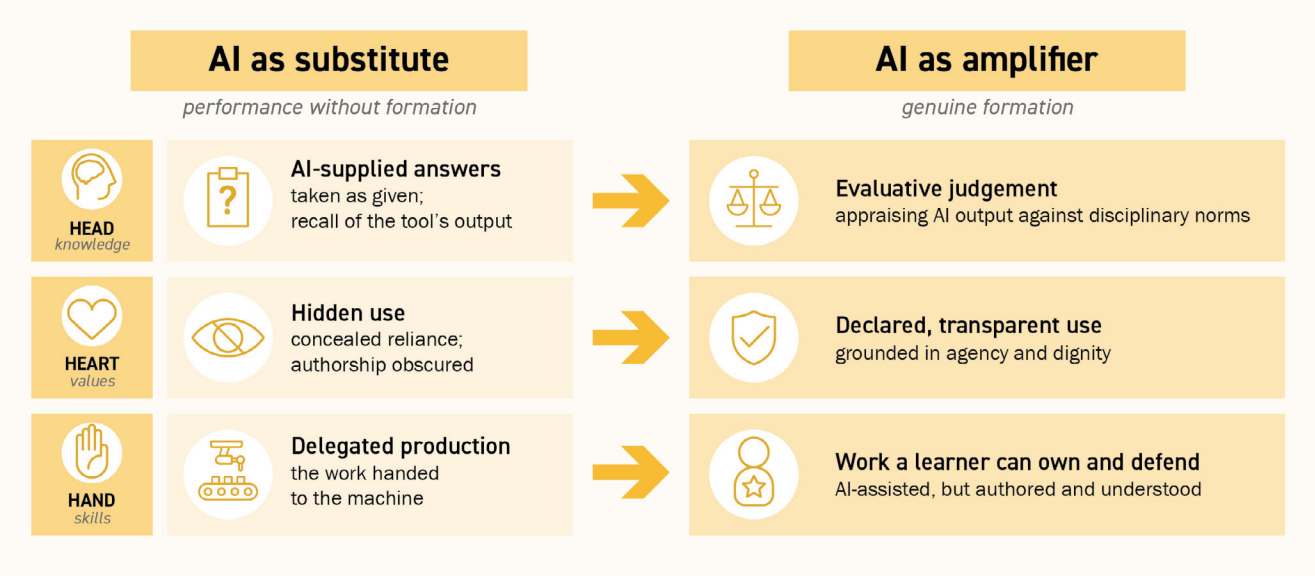


Figure 1. *The AI-Era Shift from AI as Substitute to AI as Amplifier. The figure reads across three dimensions of a well-rounded graduate: Head (what a learner knows), Heart (what a learner values) and Hand (what a learner can do). This model has its roots in Pestalozzi's 'head, heart and hand' and is widely used in holistic and transformative learning.*[6]

3 The core challenge: Pedagogy, assessment, and learner formation

The four principles apply to three connected areas. Pedagogy develops human capabilities, assessment provides evidence of them, and learner formation concerns how they grow over time. The three must change together; changing one alone weakens the whole. This section sets out what national frameworks need to protect. Decisions about course design, specific tasks, and classroom practice remain with institutions and educators.

3.1 Pedagogy

Pedagogy should deliberately build the capabilities that AI cannot develop in the learner, especially judgement, critical thinking, self-regulation, metacognition, and agency. Learners should be expected to think for themselves before using AI, engage critically with it during a task, and evaluate what it produces afterwards.

One practical approach is to ask learners to think before they prompt (pre-AI cognitive priming). They first define the problem, develop initial ideas, and explain their reasoning. AI is then used for critique, comparison, feedback, and improvement, not to replace the learner's own first thinking.

Effective pedagogy should keep productive difficulty,[7] dialogue with educators and peers,[8] reflection and justification,[9] and critical evaluation of AI outputs against the standards of the discipline.[10] Together, these processes support independent judgement and learner agency.[11]

The main risk is that AI lets learners skip the processes through which learning happens: working through uncertainty, revising attempts, discussing ideas, and reflecting on feedback. These processes are how understanding, judgement, and ownership of ideas develop.

3.2 Assessment

Assessment rests on one assumption: that submitted work shows what the learner can do. GenAI breaks that assumption, so the main challenge is now validity, not detection.[12] Responses based on rules, declarations, or authorship checks create an "enforcement illusion" whereby they demand compliance without improving the evidence of what learners can actually do.[13] Proving that a learner worked without AI also becomes harder as tools improve, and trying to do so shifts effort into surveillance.[14] Where integrity concerns arise, any review should follow fair procedures and include a supportive, face-to-face conversation between educator and learner about how the work was produced and what the assessment required.

Widespread AI use by students across many universities shows the need for system-wide assessment reform.[15] International guidance points the same way: UNESCO1 and the OECD[16] both call for assessments that value how students learn, not only what they submit. Regulators are also beginning to act. Australia's higher education regulator has required every provider to submit an institutional action plan addressing the risk GenAI poses to award integrity.[17] Assessment should make learning and judgement visible through reflection, oral examination, portfolios, peer review, staged submissions, and real-world problem-solving. It should assess how learners reason, justify decisions, and use AI critically and responsibly.[18]

One practical tool is a short AI-use statement, in which learners explain how they used AI, which outputs they accepted, changed or rejected, and how their own judgement shaped the final work.

3.3 Learner formation

Learner formation is about what a person becomes through education—that is, their intellectual character, independent judgement, ethical sense, relationships with others, and sense of responsibility. Humanising learning should therefore develop agency, dignity, relationships, the capacity to act responsibly in society, as well as competence.

European traditions of Bildung have emphasised autonomy and independent judgement.[19] Confucian and related East Asian traditions have emphasised moral self-cultivation, relationships, and responsibility to others.[20] These traditions are not opposites. Together, they suggest that humanising learning requires both the ability to think and judge independently and the ability to use that judgement in relation to others.

Recent policy examples from both Asia and Europe, some written for schools, reflect this common ground. Singapore integrates AI literacy across the curriculum and teaches students to check AI-generated information in Character and Citizenship Education.[21]Japan's guidance emphasises that GenAI outputs must undergo critical evaluation and human review.[22] Finland's guidance stresses ethical use and the protection of foundational learning.[23] European and international guidance links AI literacy with critical evaluation and ethical responsibility.[24]

These examples point to a shared principle that AI should support learning without replacing the human processes through which learners develop judgement, responsibility, relationships, agency, and a sense of self. Education should therefore do more than produce skilled users of AI. It should form learners who can question automated outputs, justify their own conclusions, and act ethically.

4 Policy recommendations for humanising learning

This section turns Sections 2 and 3 into seven recommendations for national policymakers, with institutions responsible for implementation. Recommendations 4.1 and 4.2 set the foundations; 4.3 and 4.4 address pedagogy and assessment; 4.5 addresses learner formation; and 4.6 and 4.7 apply across all three. Section 5 sets out timing, resources, and responsibilities.

4.1 Ground policy and institutional practice in a stated purpose

National frameworks should begin by stating what education is for. Each AI measure should then be justified by how it serves that purpose, so that regulation rests on educational aims rather than on reacting to each new tool. Frameworks should also require each institution to set out its own educational purpose, reflecting its culture and disciplines, and to base its AI policy on that purpose. AI policy should be led by academic and educational leadership, not left by default to IT, legal, or academic integrity offices.

Key Action: Require every national and institutional AI policy to begin with a short statement of the learning, capabilities, and learner formation it is designed to protect.

4.2 Govern proportionately

Regulation should focus on high-risk uses of AI, such as grading, progression, and academic-integrity decisions, and avoid excessive control of low-risk classroom use. Institutions should apply the same risk levels in their own policies. Section 7 sets out the minimum safeguards for each level.

Key Action: Classify AI uses as high, medium, or low risk, and match safeguards to the consequences for learners.

4.3 Assess AI tools for their impact on learner formation before procurement

Every AI tool embeds pedagogical choices about how students learn, yet decisions about procurement and digital infrastructure can put cost, features, and data compliance ahead of educational value. Ministries should require, and provide templates for, a humanising learning impact assessment based on two questions: "What kind of learning does this tool encourage? What human capability might it displace?"

Key Action: Introduce a humanising learning impact assessment for every new AI tool used in teaching and assessment to check effects on learner agency, relationships with educators and peers, intellectual authorship, and accessibility before adoption.

4.4 Enable universities to redesign pedagogy and assessment together

The right response to GenAI is redesign, not surveillance. For pedagogy, this means programmes in which learners think for themselves before using AI, work through productive difficulty, and discuss their reasoning with educators and peers (Section 3.1). National policy can support this by funding educators' time and professional development and by asking institutions, through quality assurance, to show how their programmes respond to AI, as Australia's regulator has done (Section 3.2). For assessment, quality assurance frameworks should require credible evidence of learners' reasoning, judgement and ownership of their work, while leaving institutions free to decide how that evidence is gathered. They should also prohibit the use of automated AI detection scores as the only evidence in academic integrity cases.

Key Action: Fund educators' time and professional development to redesign learning around AI. In programmes where GenAI can substantially shape the final product, require at least one process-based or conversation-based assessment component, such as oral explanation, process portfolios, staged drafts, peer review, or AI-use statements.

4.5 Fund and conduct long-term monitoring

As Section 1 shows, current evidence on AI's effects on learning is still emerging and mostly short-term. We do not yet know whether, over time, regular AI use strengthens or weakens the capacities that education aims to develop. National systems should therefore fund long-term monitoring of the four capacities set out in Principle 2: evaluative judgement, self-regulation, metacognition, and intellectual authorship.

Key Action: Agree common indicators for these four capacities and track them across cohorts and programmes, as set out in Section 7.3

4.6 Sustain cross-civilisational dialogue

Humanising learning looks different in different cultures (Principle 4). National bodies should therefore regularly exchange experience of how it is put into practice, so that no single culture's assumptions are treated as universal.

Key Action: Use the Asia-Europe Foundation (ASEF) and related Asia – Europe platforms to develop shared language, comparative case studies, and policy-learning exchanges on humanising learning in higher education.

4.7 Protect equity of access and of outcomes

The benefits of humanising learning must reach all learners, not only those with more resources. Where access to high-quality AI tools depends on what students can pay, national policy should provide tools, through shared infrastructure, that have passed the impact assessment in Section 4.3, together with AI literacy support. Institutions should then check that assessment reform and the use of AI narrow rather than widen gaps in what students achieve.

Key Action:

Give all students access to approved AI tools and AI literacy support, and require institutions to report whether gaps between student groups are narrowing or widening (see Section 7.5).

5 Implementation pathways

Recommendation	Main Implementation Mechanism	Timing	Main Resource or Cost Driver	Primary Responsibility
Ground policy in a stated purpose	Include a statement of educational purpose in national and institutional AI policies	Start now; review periodically	Policy development and consultation time	Ministry and institutions
Govern proportionately	Set risk levels and match oversight to the consequences of AI use for learners	Start now; review periodically	Governance and oversight capacity	Ministry and institutions
Assess AI tools before procurement	Build a humanising learning impact assessment into procurement and approval	Start now; then for each new tool	Evaluation and staff time	Institutions; ministry provides shared templates
Support pedagogy and assessment redesign	Support programme-level redesign towards process, reasoning, conversation and real-world tasks	Start now; ongoing	Educators' time and professional development	Institutions; ministry funds or incentivises
Monitor learner formation	Track the four capacities in Principle 2 over time	Set up now; ongoing	Instrument development, data collection, and analysis	Ministry and institutions
Sustain cross-civilisational dialogue	Use regional platforms for comparative policy learning and shared case studies	Ongoing	Coordination and convening	Ministry and regional platforms
Protect equity of access and outcomes	Provide approved AI tools and AI literacy support, and monitor gaps between student groups	Ongoing	Procurement, learner support, and monitoring	Ministry and institutions

Table 1. *Implementation pathways for humanising learning*

Table 1 shows how each recommendation can be put into practice, when, at what main cost, and who is responsible. Countries should adapt it to their own laws, capacity and stage of AI adoption; it is not a fixed sequence.

Countries should start with the low-cost steps in Table 1: a stated educational purpose, clear risk levels, and a common template for assessing AI tools, built into existing policy, quality assurance, and procurement processes. Redesigning pedagogy and assessment, and long-term monitoring, need sustained staff time and funding, which ministries can support through shared guidance and targeted incentives. Throughout, national policymakers set the conditions and basic safeguards, and institutions put them into practice in ways that fit their own pedagogy, disciplines, and culture.



6 Safeguards and accountability

Each area of the framework carries a specific risk when AI is used without safeguards. This section first matches safeguards to the level of risk, as the recommendation to govern proportionately requires (Principle 3; see Section 4.2). It then sets out the main risk to pedagogy, assessment, and learner formation.

6.1 Minimum safeguards by level of risk

Table 2 matches safeguards to risk, in line with the EU AI Act25, which classifies AI used to evaluate learning as high-risk.

Risk Level	Examples of Use	Safeguards
High	Grading and awards; progression and academic standing; academic-integrity decisions; any other assessment decision that affects a learner's rights or opportunities	Strongest: a named person responsible for each decision, checks on fairness and effects on learning, and a real right of appeal
Medium	Personalised tutoring; automated feedback on draft work; learning data that informs but does not make decisions about learners	Proportionate: learners are told when AI is used, educators oversee its use, and learners can choose a non-AI option where possible
Low	Study support; help with drafting and brainstorming; searching for information; and other everyday uses that do not feed into important decisions	Start now; then for each new tool

Table 2. *Risk levels and minimum safeguards*

6.2 Risk to pedagogy: Surveillance that harms learning

Risk: Students who feel watched are less likely to explore, test ideas, and make mistakes. This productive difficulty is how judgement develops (Section 3.1).

Policy Response: Support universities to redesign pedagogy and assessment rather than increase monitoring (Section 4.4). Prohibit AI systems that infer students' emotions; Article 5 of the EU AI Act25 already bans these in education, except for medical or safety reasons.

6.3 Risk to assessment: Unfair integrity decisions

Risk: A learner cannot examine or challenge a detector score, so an allegation based only on such a score cannot be tested, which undermines trust in assessment.

Policy Response: Through quality assurance, prohibit detector scores as the only evidence in integrity cases (Section 4.4), and require procedures that let students explain their work before any decision (Section 3.2).

6.4 Risk to learner formation: Loss of skills over time

Risk: AI tools may gradually take over the thinking work of students, weakening the four capacities in Principle 2 unless pedagogy changes too, and the effects may show only after years.

Policy Response: Require AI tools to be assessed before procurement (Section 4.3) and track these capacities over time (Section 7.3).

7 Metrics and evaluation

Monitoring is a shared task: institutions gather the evidence, and national systems set the indicators, require reporting, and use the results to review policy (Section 4.5).

The focus is the quality of learning, not the amount of AI used. The indicators follow the framework's three areas and its four principles: AI as amplifier, not substitute; formation over performance; proportionate governance; and a cross-civilisational foundation. Equity applies to all of them.

7.1 Pedagogy: Does AI extend students' thinking? (Principle 1)

Apply the Principle 1 test: can students explain and defend their work, and is there evidence of their own reasoning and revision? Check whether time saved by AI is reinvested in feedback, mentoring, and human contact.

7.2 Assessment: Validity of evidence (Principle 2)

Principle 2 holds that a qualification is a statement about what a person can do. Check whether assessment still supports it: compare results on tasks that AI cannot easily complete, such as oral defence or process portfolios,

with results on take-home work. A large gap suggests that take-home work is no longer reliable evidence. Do not rely on AI detection rates: they measure enforcement, not learning.

7.3 Learner formation: The four capacities (Principle 2)

Track students' evaluative judgement, self-regulation, metacognition, and intellectual authorship over time, and compare programmes that use AI and assess learning in different ways. Early evidence[8] suggests that learner agency and self-regulation may be at risk. We are not aware of a validated instrument that measures all four, so ministries should fund the development of such an instrument.

7.4 Governance and cultural fit (Principles 3 and 4)

Check regularly that safeguards still match the risks and adjust them as evidence grows. Institutions should also check that their understanding of learner formation works for students from different cultural traditions. Traditions differ, for example, in how far they see learning as an individual achievement or as something developed in relation to others.[20] Assessment that rewards only individual authorship may therefore disadvantage students from more relational traditions. Countries in Asia and Europe should use common indicators so that results can be compared. Where results differ, this should lead to dialogue, not pressure to follow a single model.

7.5 Equity of outcomes (All indicators)

Institutions should break down results by student group to check whether AI use and assessment reform narrow or widen gaps in access, participation, or achievement. Because this involves sensitive information, it should rely on students' consent, protect their identities, and collect only the data that is needed. This monitoring should be required, not optional.



8 Conclusion: A final note for policymakers

We have argued that GenAI can reproduce the visible products of learning without supporting the human development those products are meant to represent. In response, we have proposed the Humanising Learning Framework built on four principles: AI as amplifier, not substitute; formation over performance; proportionate governance; and a cross-civilisational foundation.

For national policymakers, the priority is to protect public trust in qualifications by setting the conditions under which institutions can act on these principles: a clearly stated educational purpose for AI; proportionate, risk-based oversight; support for redesigning pedagogy and assessment; equitable access to AI-enabled learning; and long-term monitoring of learner formation. These conditions reflect the AI for Institutional Excellence track's focus on sound AI governance, policies and infrastructure, and direct AI towards ASEFInnoLab7's overall theme "The Role of Universities in Advancing AI for Good" by fostering learners who can think and act for themselves, ensuring fair access for all, and maintaining qualifications that society can trust.

The framework deliberately draws on complementary Asian and European traditions. These traditions support a view of higher education that builds employability and technical skills, but also forms ethical, independent people who can use AI responsibly in ways that benefit the wider society. Humanising learning therefore requires institutions to design and protect the conditions in which judgement, agency, and responsibility can develop.

The bottom line.

The future of higher education will be secured when institutions use AI in ways that protect the human processes through which learners develop judgement, agency, ethical responsibility and intellectual authorship. This is what AI for institutional excellence means for higher education: AI may help produce an answer, but higher education remains responsible for forming the person who can question it, justify it, and act on it.

Acknowledgement

The work of Bogna Liziniewicz on this paper was supported by the Engineering and Physical Sciences Research Council [Grant Number EP/S023305/1].

Declaration of generative AI use

In preparing this paper, the authors used Claude[26] to support language editing and proofreading. The ideas, arguments and recommendations are the authors' own. The authors reviewed and verified all AI-assisted text and take full responsibility for the content of this paper.

Authors

Prof Mira KARTIWI

International Islamic University Malaysia, Malaysia

Prof Andri IOANNOU

Cyprus University of Technology, Cyprus

Prof Zainal Abidin Bin SANUSI

International Islamic University Malaysia, Malaysia

Dr Aneela YASMIN

EPIC College, Australia

Dr Joey AVILES, Professor

Angeles University Foundation, Philippines

Ms Bogna LIZINIEWICZ

University of Nottingham, United Kingdom

Prof Vanessa CAMILLERI

University of Malta, Malta

References

1. UNESCO (2023). *Guidance for generative AI in education and research*. https://unesdoc.unesco.org/ark:/48223/pf0000386693

2. Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, *56*(2), pp. 489–530. https://doi.org/10.1111/bjet.13544

3. Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, *160*, Article 108386. https://doi.org/10.1016/j.chb.2024.108386

4. Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3706598.3713778

5. Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, *227*, Article 105224. https://doi.org/10.1016/j.compedu.2024.105224

6. Sipos, Y., Battisti, B., & Grimm, K. (2008). Achieving transformative sustainability learning: Engaging head, hands and heart. *International Journal of Sustainability in Higher Education*, *9*(1), pp. 68–86. https://doi.org/10.1108/14676370810842193

7. Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist*, vol. 51, no. 2, pp. 289–299. https://doi.org/10.1080/00461520.2016.1155457

8. Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, *210*, Article 104967. https://doi.org/10.1016/j.compedu.2023.104967

9. Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), pp. 467–481. https://doi.org/10.1007/s10734-017-0220-3

10. Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, *54*(5), pp. 1160–1173. https://doi.org/10.1111/bjet.13337

11. Biesta, G. (2020). Risking ourselves in education: Qualification, socialization, and subjectification revisited. *Educational Theory*, *70*(1), pp. 89–104. https://doi.org/10.1111/edth.12411

12. Dawson, P., Bearman, M., Dollinger, M., & Boud, D. (2024). Validity matters more than cheating. *Assessment & Evaluation in Higher Education*, *49*(7), pp. 1005–1016. https://doi.org/10.1080/02602938.2024.2386662

13. Corbin, T., Dawson, P., & Liu, D. (2025). Talk is cheap: Why structural assessment changes are needed for a time of GenAI. *Assessment & Evaluation in Higher Education, 50*(7), pp. 1087–1097. https://doi.org/10.1080/02602938.2025.2503964

14. Perkins, M., Roe, J., & Furze, L. (2025). Reimagining the artificial intelligence assessment scale: A refined framework for educational assessment. *Journal of University Teaching and Learning Practice, 22*(7), Article e1707. https://doi.org/10.53761/rrm4y757

15. Chirikov, I., Smirnov, I., & Kizilcec, R. F. (2026). Generative AI use and misuse call for assessment reform in higher education. *Science, 392*(6800), pp. 818–820. https://doi.org/10.1126/science.aec5115

16. OECD. (2026). *OECD Digital Education Outlook 2026: Exploring effective uses of generative AI in education*. OECD Publishing. https://www.oecd.org/en/publications/oecd-digital-education-outlook-2026_062a7394-en.html

17. Tertiary Education Quality and Standards Agency (TEQSA). (2024, June 3). *Request for information: Addressing the risk of artificial intelligence.* https://www.teqsa.gov.au/guides-resources/higher-education-good-practice-hub/artificial-intelligence/request-information-addressing-risk-artificial-intelligence

18. Aruleba, K., Sanusi, I. T., Obaido, G., Ogbuokiri, B., & Mienye, I. D. (2025). *Beyond the prompt: Student strategies, ethical reflections, and learning with ChatGPT in computer science*. Technology, Knowledge and Learning. Advance online publication. https://doi.org/10.1007/s10758-025-09899-7

19. Horlacher, R. (2017). *The educated subject and the German concept of Bildung: A comparative cultural history*. Routledge.

20. Li, J. (2012). *Cultural foundations of learning: East and West*. Cambridge University Press.

21. Ministry of Education, Singapore. (2026, March 3). *We learn for life together: Building a "We First" society through education* [Press release]. https://www.moe.gov.sg/news/press-releases/20260303-we-learn-for-life-together-building-a-we-first-society-through-education

22. Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2024). *Guideline for the use of generative AI in primary and secondary education (Ver. 2.0)*. https://www.mext.go.jp/a_menu/other/mext_02412.html

23. Finnish National Agency for Education. (2024). *AI guide for teachers: Digital information literacy.* https://www.oph.fi/en/statistics-and-publications/publications/ai-guide-teachers

24. European Commission. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union. https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en

25. European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Official Journal of the European Union, L 2024/1689. http://data.europa.eu/eli/reg/2024/1689/oj

26. Anthropic. (2026). *Claude (September 2026 version)* [Large language model]. https://claude.ai



Cross-Regional Generative AI Governance in Higher Education:

From Reactive Policies to Coherent Institutional Frameworks

AUTHORS
Shalinka Jayatilleke
Akhtar Jalbani
Rashikala Weerawarna
Sunyub Kim
Julija Mironova
Andrea Mascolo

This paper helps university stakeholders to build a coherent institutional-wide GenAI governance system by diagnosing gaps in both policy coverage and coordination and mapping a staged pathway for improvements.

Executive summary

The challenge for higher education is not the absence of GenAI policy but the lack of coherence across institutional responses.

Generative AI (GenAI) is transforming teaching, assessment, research and administration in higher education. Universities have responded with new policies and guidance, but the reviewed evidence indicates that these responses can remain fragmented across institutional areas. The central problem is therefore not simply a lack of rules, but a lack of coordination: responsibilities, terminology, oversight and review processes may remain disconnected even when multiple policies exist.

This challenge directly connects to the ambition of "AI for Common Good": universities need governance that enables GenAI's educational and research benefits to be shared equitably while protecting students, staff and the wider public from harm. By connecting academic governance with equity, privacy and public accountability, the proposed maturity model offers a practical starting point for aligning institutional GenAI use with these public responsibilities.

How we examined it: We reviewed publicly available English-language GenAI governance documents from selected higher education institutions and sector bodies across Asia, Europe and Australia. The comparative findings, combined with the staged-capability logic of the Capability Maturity Model Integration (CMMI), informed a five-tier maturity model covering six governance domains: teaching and learning; assessment and academic integrity; research ethics and integrity; equity and access; data privacy and security; and public trust and accountability. The framework contains 56 indicators: 48 scored maturity indicators and eight qualitative coherence indicators. The five tiers provide developmental reference points, while the indicators can be reviewed and updated as GenAI practice, risks and regulations evolve. . The framework has been implemented as an online publicly available self-assessment tool.

Tool:
https://akjalbani.github.io/genaigov_maturitymodel_v1/

What we found: We conducted a proof-of-concept assessment involving 12 selected Higher Education Institutions (HEIs), with four from each region. Scores ranged from 30 to 71 out of 100, with a median of 56, and no institution reached the highest maturity tier. Within this small sample, the European cases had the highest regional median, while the Asian cases showed the

widest range. However, these regional differences were not statistically significant and should not be interpreted as a ranking of regions. The substantial variation within regions reinforces the importance of examining each institution's governance context and domain-level profile.

What institutions should do: Universities should move from separate policies towards coordinated governance with named ownership, consistent terminology, and aligned review processes. Governments and sector bodies can support this transition through clearer guidance, capacity-building and multilingual resources. Quality assurance bodies can explore how domain profiles and coherence evidence might support developmental review, but the model should not yet be treated as a validated accreditation threshold or a sector-wide ranking. This paper concludes with practical recommendations for these stakeholder groups and identifies priorities for broader piloting and continued refinement of the framework.

Introduction

GenAI now touches every part of university life and governing it has become a whole-of-institution task.

The rapid diffusion of Generative AI (GenAI) has made its governance a whole-of-institution challenge in higher education (HE). GenAI refers to AI systems capable of generating new content, including text, images, and other forms of media, and its growing use creates new opportunities for innovation across HE.[1] At the same time, its widespread adoption raises interconnected concerns around academic integrity, privacy, security, bias, equity, intellectual property, and accountability.

Institutional responses remain fragmented and often reactive. While many universities have focused on assessment and academic misconduct, broader issues such as data governance, staff capability, equitable access, and administrative use receive less consistent attention.[2,3,4] The challenge is therefore not simply whether policies exist, but whether governance is coherent across institutional domains. Effective AI policy in HE requires integrated attention to pedagogical, governance, and operational dimensions rather than isolated guidance.[5]

This white paper compares GenAI governance across HEIs in Asia, Europe, and Australia, where responses reflect different national, regional, and institutional contexts.[6,7,8,9] It proposes a practical GenAI Governance Maturity Model as a flexible benchmark to assess institutional capability, identify gaps, and support progress towards more coherent, accountable, and adaptive governance.

How the evidence was gathered

Cross-regional document analysis, combined with staged capability theory, grounds the model in observed practice.

Research design

This study combines two things: a comparative review of real GenAI governance documents from HEIs across Asia, Europe, and Australia, and the design of a practical maturity model that institutions can use to assess and improve their own governance. The model combines CMMI's five-stage maturity architecture with governance domains and indicators derived from the cross-regional document analysis.

Theoretical grounding

The model is based on the Capability Maturity Model Integration (CMMI), a widely used framework in which organisational capability develops through five stages, from ad hoc practice to full optimisation.[10,11] We chose CMMI over education-specific "digital maturity" frameworks because it focuses on organisational capability and coordination rather than technology adoption alone, which is what GenAI governance actually requires. CMMI[12] provides an established five-stage model of organisational capability development. We retained its five maturity stages and their progression from ad hoc practice towards increasingly managed, integrated and optimised capability, while adapting the meaning of each stage to GenAI governance in HEIs. The HEI-specific stage definitions were informed by cross-regional document analysis and the six governance domains identified in this study.

Document selection and analysis

The evidence base included publicly available GenAI policies, governance frameworks, academic integrity statements, national and regional policy documents,

and reports from recognised higher education bodies. Documents were included when they were publicly accessible, available in English, and attributable to an identifiable higher education institution, national system, or recognised sector body. Internal or restricted documents and general commentary without formal policy status were excluded.

Co-authors based in Asia, Europe, and Australia collected and reviewed documents from their respective regions, bringing familiarity with the relevant policy and regulatory contexts. The documents were examined using a shared template covering policy scope, alignment across governance domains, responsible institutional bodies, and gaps in areas such as equity and public accountability. The analysis combined themes identified from the documents with CMMI's staged-capability logic. Differences in interpretation were discussed and resolved through cross-regional consensus sessions. This process informed the six governance domains and five-tier model shown in Figure 1.

Because the review was limited to publicly available English-language documents, governance expressed mainly through internal materials, local-language policies or legislation may be under-represented. The findings therefore describe documented governance rather than the full extent of institutional practice. No human participants were involved, and ethical clearance was not required.

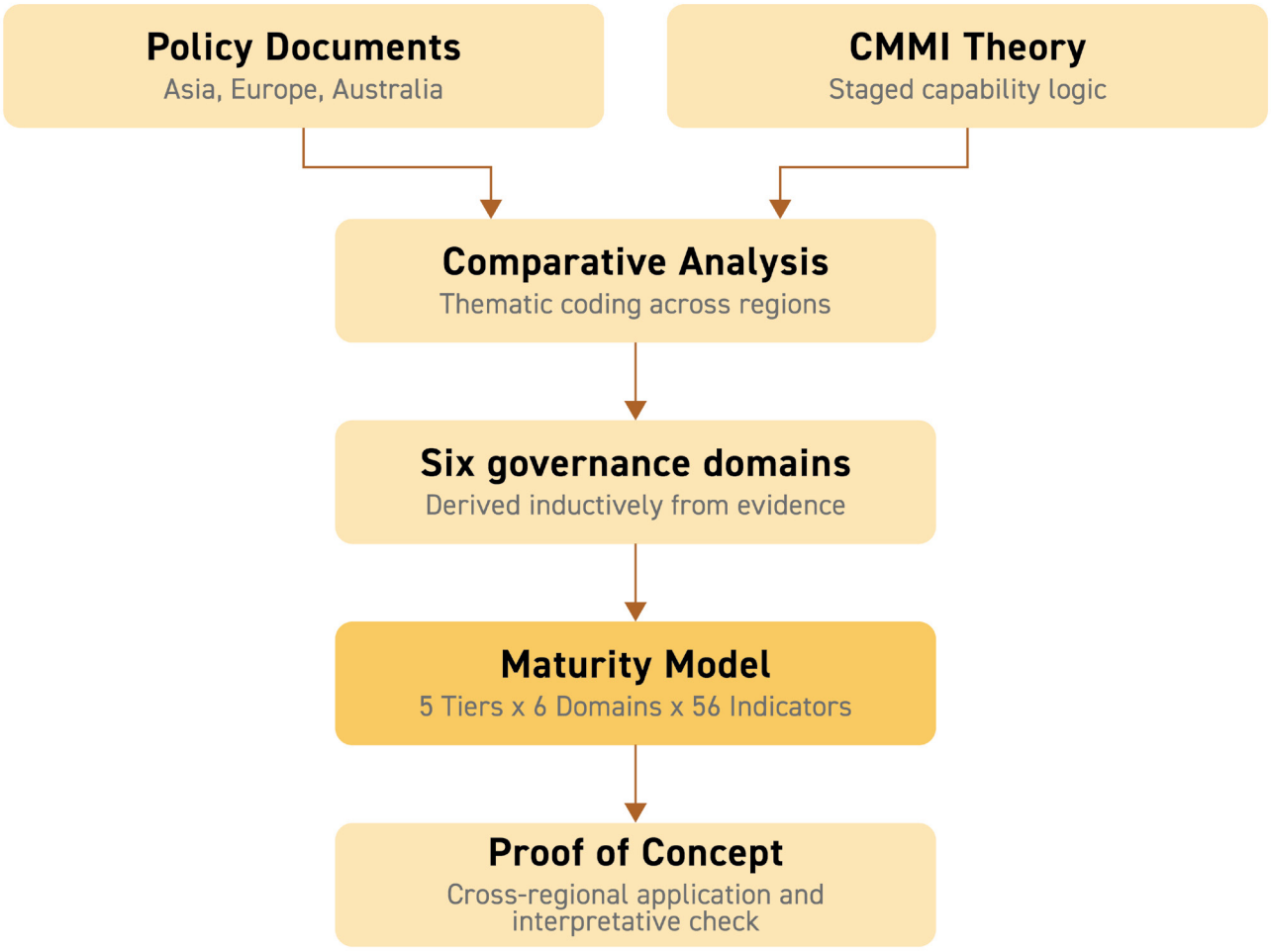


Figure 1. *How the comparative analysis and CMMI theory come together to produce the model*

What the cross-regional comparison shows

All three regions are moving beyond assessment-only responses, but along different paths.

The cross-regional analysis using a collection of publicly available policy documents across universities, examined how GenAI governance is expressed across Asia, Europe and Australia. Based on the analysis, we established four key analytical areas. Out of which, 'Governance Context and Architecture' was treated as a cross-cutting lens because it captures the regulatory, sectoral, and institutional environment shaping all areas of governance rather than representing a governance domain itself. The remaining three analytical areas map directly to six domains: Academic Governance Domains encompasses Teaching & Learning, Assessment & Academic Integrity, and Research Ethics & Integrity; Institutional Safeguards encompasses Equity & Access and Data Privacy & Security; and Public Accountability and Governance Coherence aligns with Public Trust & Accountability while also examining how governance activities are coordinated across domains. The following sections present the similarities and differences across the three regions related to these analytical areas.

1. Governance context and architecture

Across all three regions, GenAI governance operates across multiple levels and requires universities to translate broader expectations into institutional practice. In each region, there is evidence of movement away from isolated or assessment-specific responses towards more structured institutional governance. However, the balance between regional, national, and institutional influences differs: Asia shows stronger alignment with regional and national strategies; Europe combines supranational and national frameworks with institutional autonomy; and while Australia is nationally guided but largely institutionally implemented.

2. Academic governance domains

Across the three regions, assessment and academic integrity are the most consistently developed areas, with increasing emphasis on managed rather than purely restrictive AI use. Common approaches include defining acceptable use, disclosure or acknowledgement requirements, assessment redesign, and maintaining student responsibility for submitted work. Teaching and learning governance is also expanding through staff support, AI literacy and pedagogical guidance, while research ethics and integrity are increasingly recognised but are not always integrated with teaching and assessment governance.

3. Institutional safeguards

All three regions recognise that GenAI governance must extend beyond academic integrity to include privacy, security, risk, and equitable participation. Data privacy and security tend to be supported by more concrete controls, such as restrictions on sensitive information, approved systems, risk assessment and cybersecurity processes. Equity and access are also recognised but are less consistently operationalised through measurable institutional mechanisms. The regional differences therefore lie less in whether safeguards are recognised and more in what risks are prioritised and how formally safeguards are implemented.

4. Public accountability and governance coherence

Across all three regions, transparency, accountability, and coordination are increasingly important as institutions move beyond standalone AI guidance. More developed approaches connect AI governance with existing academic, research, risk, security, and institutional structures, while less integrated approaches distribute responsibility across separate policies and organisational units. A common challenge is therefore moving from the existence of multiple policies to governance in which responsibilities, terminology, oversight and review mechanisms operate coherently across the institution. This directly supports the maturity model's later distinction between multi-domain coverage and genuinely integrated governance. Table 1 compares the AI policies across the three regions on several analytical areas.

Analytical Area	Asia	Europe	Australia
Governance Context and Architecture	Strong regional harmonisation and national strategic direction; institutional governance often reflects top-down priorities	Multi-level hybrid model combining supranational regulation, national policy, sector guidance and institutional autonomy	Nationally guided but institutionally implemented, with substantial university autonomy and uneven movement towards enterprise governance
Academic Governance Domains	Strong emphasis on structured risk classification, disclosure and systematic control of AI use in academic activities	Managed-use approach with strong assessment governance and increasing integration of teaching, staff capability and research integrity	Assessment and integrity remain dominant, with growing attention to AI literacy, assessment redesign and transparent use; research governance is often separate
Institutional Safeguards	Safeguards extend beyond privacy to national intellectual property and preservation of human capability	Strong privacy and data-protection orientation; more developed institutions use procurement, security and risk controls, while equity implementation remains uneven	Increasing use of privacy, cybersecurity, risk classification and approved-tool mechanisms; equity and access are recognised but less consistently operationalised
Public Accountability and Governance Coherence	Accountability is reinforced through alignment with state requirements and international standards	Strong transparency principles, but ownership and coordination often remain distributed across organisational units	More developed institutions use named owners, committees and cross-policy links, but sector-wide coherence remains uneven
Overall Pattern	Greater emphasis on strategic alignment, structured control and regional/national coordination	Strong regulatory and ethical foundation with heterogeneous institutional implementation	Increasingly sophisticated institutional governance, but integration varies considerably between universities

Table 1. *AI policy comparison across regions*

This comparative document analysis leads us to six governance domains of HEIs:

1. **Teaching & Learning** — curriculum integration, classroom practice, and staff capability-building for GenAI.
2. **Assessment & Academic Integrity** — assessment design, disclosure requirements, and misconduct response.
3. **Research Ethics & Integrity** — authorship, data management, and research integrity training.
4. **Equity & Access** — fairness, inclusion, and equitable digital access to GenAI tools.
5. **Data Privacy & Security** — risk management, procurement standards, and information security.
6. **Public Trust & Accountability** — transparency, disclosure, and accountability to students, staff, and the public.

Proposed GenAI governance maturity model for HEIs

Five tiers, six domains, and a coherence diagnostic turn governance capability into something institutions can measure and act on.

The GenAI Governance Maturity Model for HEIs provides a structured approach for assessing and advancing coherent, consistent, and institution-wide GenAI governance, as illustrated in Figure 2. The model is designed to strengthen coordination, accountability, standardisation, and strategic alignment across institutional governance and policy domains. It is adapted from the CMMI framework and retains its five-tier maturity structure while redefining the levels to reflect the specific context of GenAI governance in higher education. The five maturity tiers, along with their corresponding score ranges and governance profiles, are presented in Table 2.

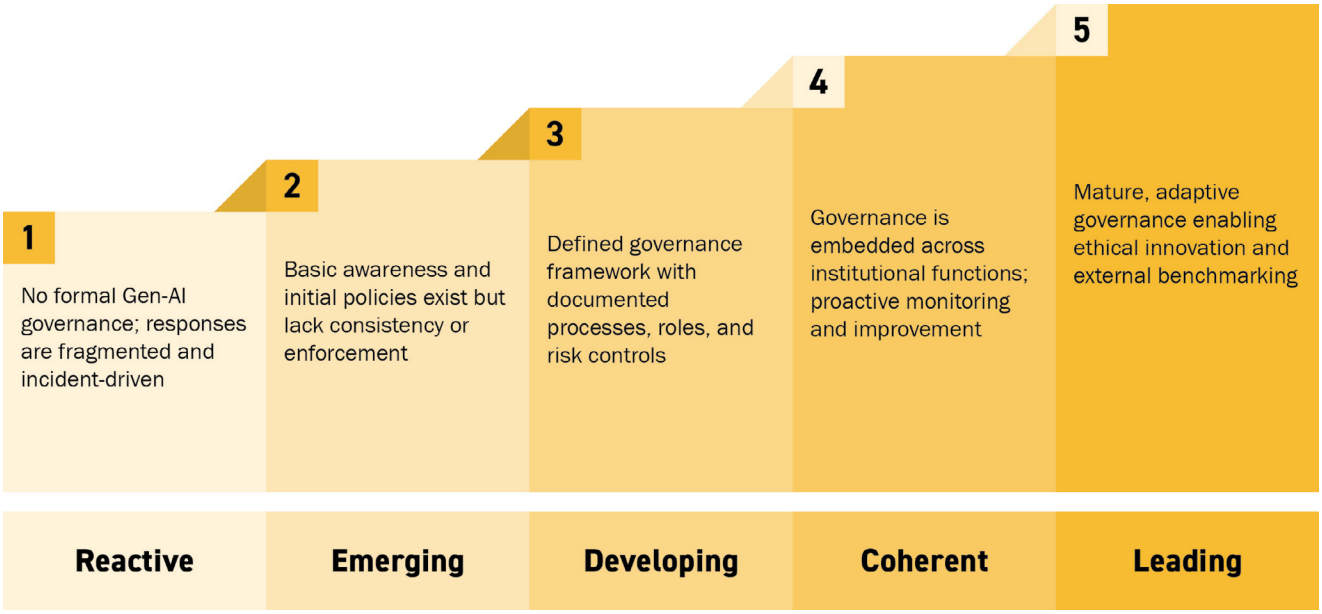


Figure 2. *Proposed GenAI Gov Maturity Model for HEIs*

Maturity Tier	Score Range	Governance Profile
1: Reactive	0-20	Ad hoc, incident-driven responses; no shared ownership or settled institutional position.
2: Emerging	21-40	Initial policies exist, concentrated almost entirely on academic integrity.
3: Developing	41-60	Multiple domains are addressed, but coordination across them remains partial.
4: Coherent	61-80	Comprehensive, integrated governance across domains with named accountability.
5: Leading	81-100	Whole-of-institution, audited, adaptive governance that shapes sector practice.

Table 2. *Tiers of the GenAI Gov Maturity Model for HEIs*

Indicators and scoring

The model comprises 56 indicators in total: 48 quantitative maturity indicators distributed evenly across the six governance domains, and eight cross-cutting coherence diagnostic indicators. The 48 maturity indicators produce the numeric score. Each is scored on a 0 to 3 scale: 0 (not in place), 1 (partial), 2 (substantially in place), or 3 (fully embedded), yielding a maximum raw score of 144. The overall maturity score is normalised to a 0 to 100 scale (the raw score divided by 144, multiplied by 100) and mapped to one of the five tiers in Table 2. The eight coherence indicators are assessed qualitatively and do not contribute to the numeric score. All indicators and domains are weighted equally. This is deliberate: equal weighting keeps the calculation transparent and reproducible, and no empirical basis yet exists for differential weights. Every score must be supported by documentary evidence; where no evidence can be produced, the indicator is scored 0. The model identifies domain-level strengths and weaknesses and is designed to support institutional, sectoral, and regional benchmarking. Provided in figure 3 is the detailed information on the GenAI governance maturity model.

GenAI Governance Maturity Indicators

Overview of the 56-indicator model across six governance domains and a cross-cutting coherence diagnostic. The model supports institutional self-assessment and improvement, not a compliance or ranking.

Scoring Scale (all 56 indicators)

0	1	2	3
Not in place	Partial/ informal	Substantially in place	Fully embedded and evidenced

Maturity Score (48 indicators)

Overall score = (raw score/144) x 100
Five maturity tiers:

0-20 Reactive
21-40 Emerging
41-60 Developing
61-80 Coherent
81-100 Leading

Coherence Indicators (8 indicators)

Assessed separately to diagnose cross-policy alignment and institutional integration. Not included in the 0-100 score (maximum raw score 144).

A1. Teaching & Learning (Indicators 01-08)

Responsible and effective integration of GenAI in teaching and learning.

01	Published teaching principles
02	Critical output verification
03	AI fluency vs evidence validity
04	Course-level AI risk scales
05	Limits on automated coaching
06	Staff development and course statements
07	AI literacy in learning outcomes
08	Supported pedagogical experimentation

A2. Assessment & Academic Integrity (Indicators 09-16)

Fair and transparent assessment with academic integrity.

09	Published assessment rules
10	Student AI-use disclosure
11	Assessment-design safeguards
12	Assistance vs work replacement
13	Disclosure transparency requirements
14	Staffreview of assessment items
15	GenAI misconduct/verification procedures
16	Consistent cross-faculty rules

A3. Research Ethics & Integrity (Indicators 17-24)

Ethical and responsible use of GenAI in research

17	GenAI in research ethics
18	Researcher AI-use disclosure
19	Bias/accountability/limitations guidance
20	Research linked to responsible AI
21	Human author liability
22	External funder policy compliance
23	Authorship and grant-writing guidance
24	GenAI research-integrity training

A4. Equity & Access (Indicators 25-32)

Inclusive, fair and equitable access to GenAI tools and opportunities.

25	Equitable AI access
26	AI literacy and inclusion
27	Responsible-use support for all
28	Unequal access addressed
29	Fair grading across access differences
30	Alternatives for AI non-users
31	Paid/free digital divide
32	Equity and cultural audit

A5. Data Privacy & Security (Indicators 33-40)

Protection of data, privacy and institutional security.

33	Privacy implications documented
34	Sensitive-data upload restrictions
35	Data-protection compliance
36	Retention/third-party risk awareness
37	Protected procurement environments
38	Institutional account security
39	Approved/prohibited tools register
40	Records and cybersecurity assessment

A6. Public Trust & Accountability (Indicators 41-48)

Transparent, accountable and publicly engaged governance.

41	Public institutional GenAI position
42	Publicly accessible governance information
43	Transparency and disclosure principles
44	Human accountability for AI decisions
45	Formal AI-use recourse process
46	Legal/administrative accountability
47	Consultation and student voice
48	Annual reporting and concerns mechanism

A7. Coherence Diagnostic (Indicators 49-56)

Cross-policy alignment, coordination and continuous review (not scored)

49	Cross-policy references	53	Policy-practice coherence
50	Consistent institutional position	54	Multidisciplinary governance structure
51	Cross-functional coordination	55	Named owner and committee
52	Regulatory/QA alignment	56	Active integrated review cycle

Notes:

- Indicators are rated on a 0 – 3 evidence scale. Only indicators 01-48 contribute to the 0-100 maturity score (maximum raw score 144).
- The coherence indicators (49 – 56) are interpreted separately.
- The operational tool remains the authoritative source for the full indicator wording and evidence capture.

Figure 3. *Detailed GenAI Governance Maturity Indicators*

As a worked example, an institution whose evidence supports a raw total of 76 points across the 48 maturity indicators receives a normalised score of 52.8, which falls in the 41 to 60 band and places it in Tier 3 (Developing). The domain subtotals then show where points were lost: if 'Equity and Access' contributed only 6 of a possible 24 points, that domain becomes the improvement priority even though the overall tier sounds respectable.

The coherence diagnostic

A qualitative coherence diagnostic checks whether the HEI's six governance domains work well together. It considers whether senior leaders are clearly responsible, whether the same terms are used across documents, whether policies refer to each other, and whether reviews are planned together. This is often the main difference between Tiers 3 and 4. An organisation may do well in separate areas but still struggle to connect everything across the whole institution.

Using the model in practice

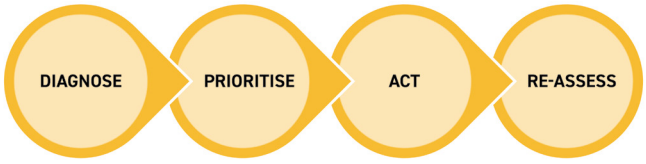


Figure 4. *Four-stage institutional adoption cycle*

The maturity model delivers the most benefit when implemented as a four-stage cycle (Figure 4). The first stage is **Diagnose**, in which each HEI assigns a team with governance experience, preferably a cross-functional group, to complete all 48 scored indicators using documentary evidence. The assessment identifies weak domains; in the second stage, **Prioritise**, the HEI team defines concrete goals to improve them. In the third stage, **Act**, the team assigns those goals to named owners, sets timelines, and links them to policy reviews, quality-assurance cycles, and committee plans. In the fourth stage, **Re-assess**, the institution repeats the assessment and improves its maturity level. As this is a cyclical process, institutional GenAI governance should be assessed annually or during major policy reviews to track progress and recalibrate priorities.

The framework described in this paper is implemented as a free, browser-based self-assessment tool. An institutional team can work through all 48 scored indicators, the eight coherence indicators, and record the supporting evidence. The result is a domain profile, an overall score out of 100 and a maturity tier.

The model is available as a working self-assessment tool.[13] No account or installation is required. No data is stored.

https://akjalbani.github.io/genaigov_maturitymodel_v1/

Navigating tier transitions

Each transition between tiers reflects a distinct stage of institutional development:

– **Tier 1 to Tier 2:** Establish a basic institutional position, assign initial responsibility, and introduce foundational guidance, often focused on assessment and academic integrity.

– **Tier 2 to Tier 3:** Expand guidance across areas such as teaching, research ethics, student support, equity, privacy, and public communication.

– **Tier 3 to Tier 4:** Move from having policies to coordinating them institution-wide through senior accountability, consistent terminology and communication, and formal governance and review mechanisms.

– **Tier 4 to Tier 5:** Embed continuous improvement through evidence gathering, stakeholder participation, equity and accessibility review, public reporting, and knowledge sharing.

Progress may not be linear, and institutions may be more mature in some areas than others. Domain-level performance is therefore more informative than relying solely on an overall tier. Overall, the tiers represent a progression from foundational policy development to broader coverage, institutional coordination, and evidence-based continuous improvement. They indicate capability rather than compliance. Scores reflect documented governance, not necessarily lived practices. As GenAI capabilities, risks, and regulations evolve, the model and its indicators will require ongoing validation, piloting, and refinement.

Evidence from a cross-regional application

Twelve institutions, three regions: institutional choices matter more than geography.

Proof of concept design

To test the model in practice, the team with one member from each region ran a proof of concept on 12 HEIs already covered in our policy dataset, four per region. We selected institutions that our own regional team members knew well enough to score confidently from documentary evidence, allowing the evidence to be interpreted with relevant regional and institutional context. This enabled documentary evidence to be interpreted with relevant regional and institutional context rather than as a decontextualised policy reading. In an operational rollout, this scoring role would normally sit with an institution's own governance team or GenAI policy committee, since they hold the same first-hand familiarity with how policy translates into practice.

Institutions are anonymised throughout, coded AU1–4, EU1–4, and AS1–4 for ethical and privacy reasons. The Appendix lists each institution's region, approximate overall score, and resulting maturity tier.

Findings

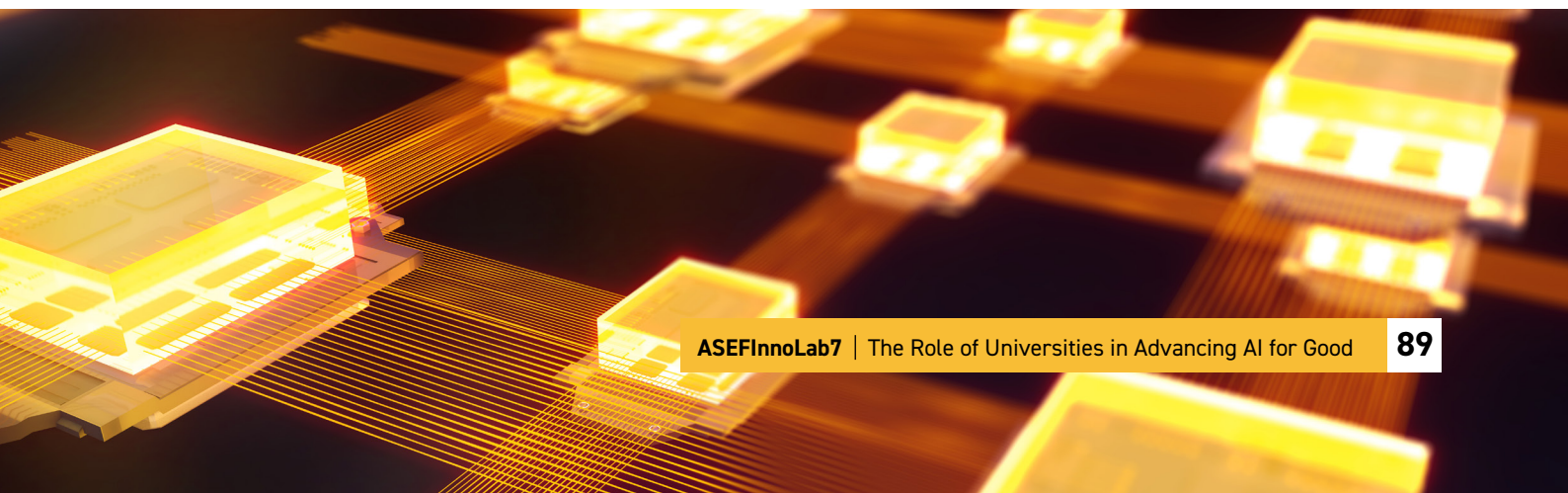
Scores across the 12 institutions ranged from 30 to 71 out of 100, with an overall median of 56. Europe recorded the highest regional median (63), followed by Australia (53.5) and Asia (45.5). A Kruskal–Wallis test[14] found no statistically significant difference in maturity scores between regions ($H = 4.25$, $df = 2$, $p = .119$); with only 12 institutions, this proof of concept cannot support claims that one region is more mature than another.

No institution reached Tier 5 (Leading); four reached Tier 4 (Coherent); three in Europe and one in Asia and two fell into Tier 2 (Emerging); one each in Australia and Asia. Variation within Asia and Australia was actually larger than the gap between regional medians, suggesting that institutional-level factors may be at least as important as geography in explaining documented governance maturity. The Appendix provides the full regional score distribution.

Despite these differences, substantial overlap was evident across regions. Several Australian and Asian institutions received similar Tier 3 maturity level scores, and the highest-scoring Asian institution received 63, the same as two European institutions. Overall, the findings indicate considerable variation in governance maturity levels within and across regions. Given the small pilot sample and the absence of statistically significant regional differences, the results should not be interpreted as supporting a simple regional ranking.

Limitations

A limitation of this study is that the policy documents used to define the governance domains and assess maturity tiers were restricted to publicly available, English-language materials. Governance activities recorded in documents with limited access, in legislation, or in local languages are therefore likely under-represented, and relying solely on English-language documents may obscure locally specific features of governance developed within distinct national and cultural contexts. In addition, regulatory frameworks cannot fully cover every corner of GenAI use in higher education, so these policy documents are limited snapshots of a rapidly evolving landscape rather than a settled state of practice. The model should therefore be regarded as a CMMI-grounded, practice-informed framework supported by an initial 12-institution proof of concept; broader cross-institutional testing is still required to empirically validate its indicators, tier boundaries and benchmarking performance across different HEI contexts.



Recommendations and call to action

Practical actions for the actors who can close the coherence gap.

The findings suggest that GenAI governance maturity depends less on geography and more on institutional choices: leadership, coordination, and evidence. The recommendations below translate the findings into action for the stakeholders who can act on them.

For higher education institutions

- **Appoint a named, senior accountable owner for GenAI governance.** Both the comparative analysis and the pilot show that coherence, not coverage, separates Tier 3 from Tier 4, and distributed ownership is the most common structural gap.
- **Assess across all six domains, not only academic integrity.** Complete the self-assessment with a cross-functional team using documentary evidence, and treat the domain profile, especially equity and access, as the improvement agenda.
- **Integrate before adding new policies.** Harmonise terminology, cross-reference policies, and coordinate review cycles before drafting new documents. Most institutions in the pilot needed integration of existing policies more than additional ones.
- **Embed re-assessment in existing machinery.** Tie annual re-assessment to quality assurance reporting, institutional risk registers, and strategic planning so the maturity profile becomes a standing evidence item rather than a parallel exercise.

For governments and policymakers

- **Use the tier framework as a shared vocabulary, not a compliance regime.** The five tiers give ministries and regulators a common language for discussing institutional GenAI readiness while leaving implementation choices to institutions.
- **Direct support to the least developed domains. Equity, access, and public accountability lag across all three regions studied.** National guidance and funding aimed at these domains will close more of the gap than further academic-integrity guidance.
- **Reach institutions that lack published governance.** Institutions without public GenAI policies are invisible in document-based evidence and likely furthest behind. Targeted capability support will reach institutions that rankings and inventories never see.

For quality assurance and accreditation bodies

- **Recognise maturity self-assessment as standing evidence.** Accepting the domain profile within existing review and audit cycles rewards institutions for honest diagnosis rather than for policy volume.
- **Enable benchmarking with context.** The common 0 to 100 scale supports comparison within and across countries, provided comparisons account for differences in institutional mission, scale, and regulatory context.

For regional and international organisations

- **Aggregate anonymised profiles to see the system.** University networks and cross-regional bodies can use pooled maturity profiles to identify sector-wide weaknesses, as this pilot did across three regions, and target shared guidance accordingly.
- **Host the model's continued refinement. The model is a versioned prototype.** Regional and international organisations are well placed to support wider piloting, cross-institutional benchmarking, empirical validation and indicator refinement as GenAI governance evolves.

The bottom Line.

AI for Common Good needs institutions that can explain who is responsible for GenAI decisions, how people are protected and how benefits are shared. Connecting policies through accountable leadership, coordinated review and evidence of improvement gives universities a practical way to work towards that goal.

Authors

Dr Shalinka JAYATILLEKE
La Trobe University, Australia

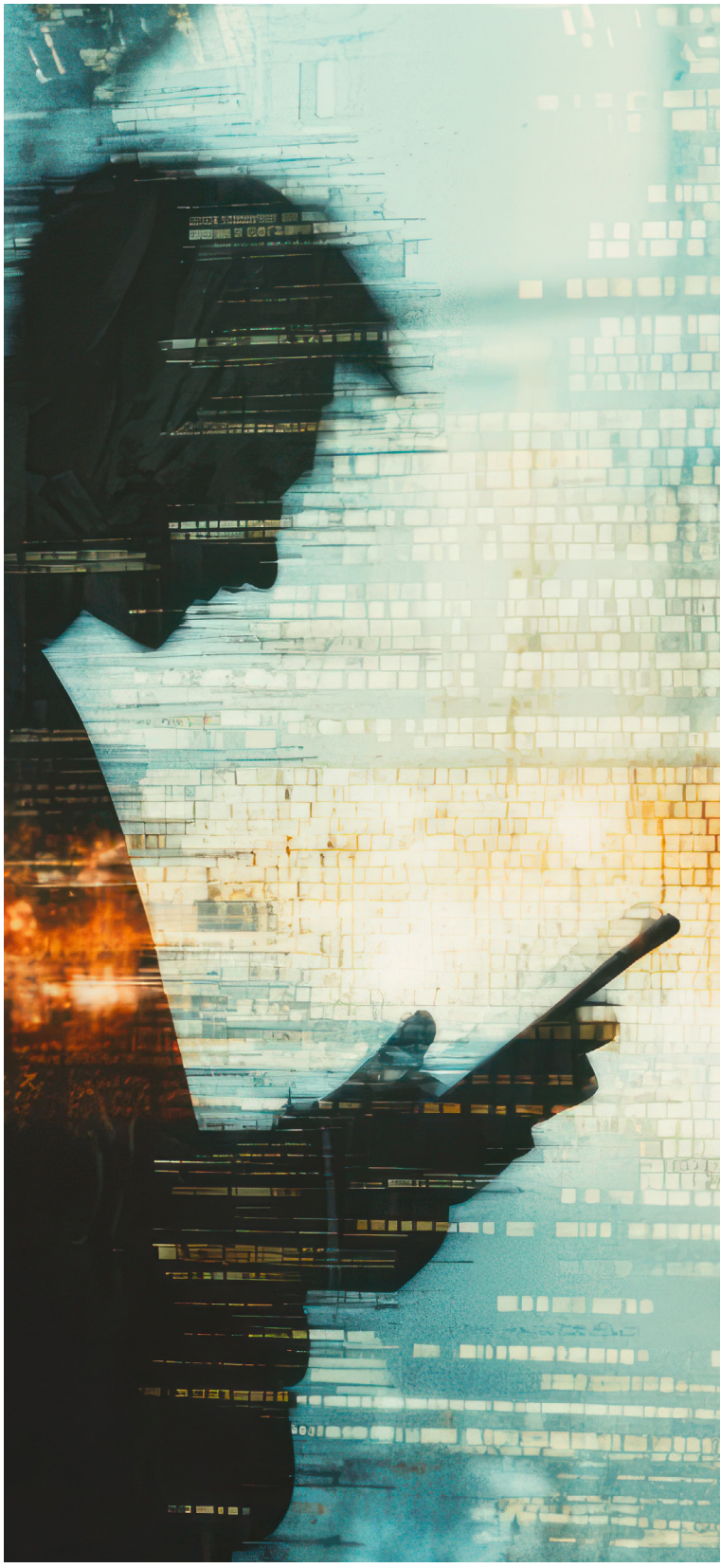
Dr Akhtar JALBANI
Holmesglen Institute, Australia

Mrs Rashikala WEERAWARNA
Holmesglen Institute, Australia

Dr Sunyub KIM
Science and Technology Policy Institute, Republic of Korea

Ms Julija MIRONOVA
Riga Nordic University, Latvia

Prof Andrea MASCOLO
European Institute for Medical Studies, Malta



References

1. Feuerriegel, S., Hartmann, J., Janiesch, C., and Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, vol. 66, pp. 111–126. https://doi.org/10.1007/s12599-023-00834-7

2. An, Y., Yu, J. H., and James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, vol. 22, Article 10. https://doi.org/10.1186/s41239-025-00507-3

3. Jin, Y., Yan, L., Echeverria, V., Gašević, D., and Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, vol. 8, Article 100348. https://doi.org/10.1016/j.caeai.2024.100348

4. Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education*, vol. 49, no. 5, pp. 651–664. https://doi.org/10.1080/02602938.2024.2309963

5. Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, vol. 20, Article 38. https://doi.org/10.1186/s41239-023-00408-3

6. Parker, L., Loper, A. J., Hayes, J., Karakas, A., White, S., and Hallman, H. (2025). Comparative analysis of artificial intelligence policies in universities across five countries. *Discover Computing*, vol. 28, Article 267. https://doi.org/10.1007/s10791-025-09745-5

7. Dai, Y., Lai, S., Lim, C. P., and Liu, A. (2025). University policies on generative AI in Asia: Promising practices, gaps, and future directions. *Journal of Asian Public Policy*, vol. 18, no. 2, pp. 260–281. https://doi.org/10.1080/17516234.2024.2379070

8. Erhardt, C., Kullenberg, H., Grigoriadis, A., Kumar, A., Christidis, N., and Christidis, M. (2025). From policy to practice: The regulation and implementation of generative AI in Swedish higher education institutes. *International Journal for Educational Integrity*, vol. 21, Article 21. https://doi.org/10.1007/s40979-025-00195-6

9. Atif, A., Jha, M., Richards, D., and Cibin, A. (2026). Institutional alignment with TEQSA's generative AI principles in Australian universities: A national document analysis with implications for ICT education. *Australasian Journal of Educational Technology*, vol. 42, no. 3, pp. 119-139. https://doi.org/10.14742/ajet.11001

10. Paulk, M. C., Curtis, B., Chrissis, M. B., and Weber, C. V. (1993, July). Capability maturity model, version 1.1. *IEEE Software*, vol. 10, no. 4, pp. 18-27. https://doi.org/10.1109/52.219617

11. Chrissis, M. B., Konrad, M., and Shrum, S. (2011). *CMMI for development: Guidelines for process integration and product improvement* (3rd ed.). Addison-Wesley.

12. CMMI Institute. (n.d.). CMMI levels of capability and performance. https://cmmiinstitute.com/learning/appraisals/levels

13. Jayatilleke, S., Jalbani, A., Weerawarna, R., Mascolo, A., Kim, S., & Mironova, J. (2026). Generative AI governance maturity model for higher education institutions (Version 1.0) [Software documentation]. Zenodo. https://doi.org/10.5281/zenodo.23363549

14. McKight, P. E. and Najab, J. (2010). Kruskal–Wallis test. In I. B. Weiner & W. E. Craighead (Eds.), *The Corsini encyclopedia of psychology*. Wiley. https://doi.org/10.1002/9780470479216.corpsy0491



Appendix

Cross-Regional Proof of Concept Results

Application of the 56-indicator model to 12 anonymised higher education institutions (HEIs) in Australia, Europe and Asia. Results illustrate distribution of maturity scores and regional patterns (not a ranking).

A. Institutional scores by domain and overall maturity

Code	Region	T&L	A&I	R&I	E&A	DPS	PTA	Overall	Tier
AU1	Australia	58	75	29	38	46	54	50	3
AU2	Australia	71	83	58	33	54	46	58	3
AU3	Australia	58	71	83	50	33	46	57	3
AU4	Australia	29	46	25	21	17	42	30	2
EU1	Europe	63	100	100	33	67	63	71	4
EU2	Europe	50	54	67	29	71	58	55	3
EU3	Europe	67	71	79	58	50	67	65	4
EU4	Europe	67	79	63	46	46	75	63	4
AS1	Asia	58	63	46	4	33	54	43	3
AS2	Asia	75	58	71	46	38	88	63	4
AS3	Asia	58	63	50	33	33	50	48	3
AS4	Asia	42	50	17	25	21	29	31	2

B. Key Findings

Overall median
56
(range 30-71)

Regional medians
Europe: 63 (range 55-71)
Australia: 53.5 (range 30-58)
Asia: 45.5 (range 31-63)

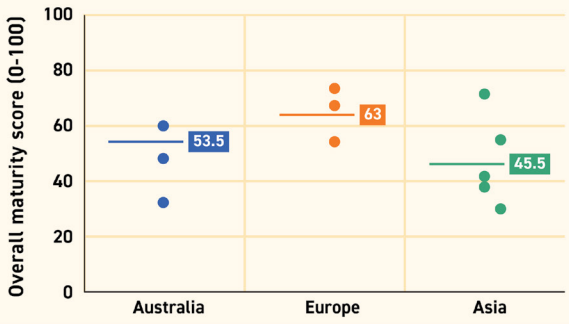
Statistical test
Kruskal–Wallis H = 4.25
df = 2, p = .119
No statistically significant regional difference.

Within-region variation
Asia: 32 points
Australia: 28 points
Europe: 16 points

Highest and lowest cases
Highest: EU1 (71, Tier 4)
Lowest: AU4 (30, Tier 2)

No institution reached Tier 5

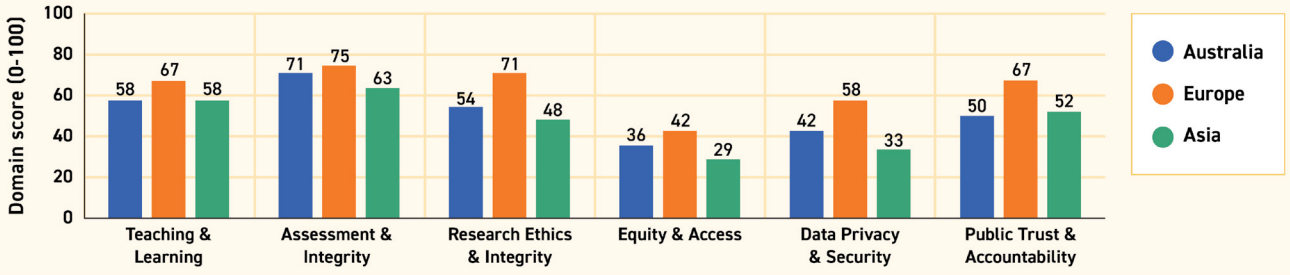
C. Overall maturity scores by region



D. Maturity tier bands

Tier 5	81-100	Leading
Tier 4	61-80	Coherent
Tier 3	41-60	Developing
Tier 2	21-40	Emerging
Tier 1	0-20	Reactive

E. Domain performance (regional medians)



F. Interpretation

- European institutions show the highest median maturity, but regional differences are not statistically significant.
- Variation within Asia and Australia exceeds the difference between their highest and lowest medians.
- The policy message is not that one region is categorically more mature, but that governance maturity varies substantially between institutions.
- Results depend on how responsibilities, policies, safeguards and review mechanisms are organised and connected.

F. Interpretation

- Proof-of-concept based on 12 anonymised HEIs (four Australia, four Europe, for Asia).
- Illustrative application to documentary evidence only.
- Not a comprehensive assessment or league table.
- Does not establish causal claims about regional maturity.
- The six-domain architecture and five-tier progression provide a stable organising logic, while indicators and evidence expectations should evolve with GenAI capabilities, regulation and institutional practice.

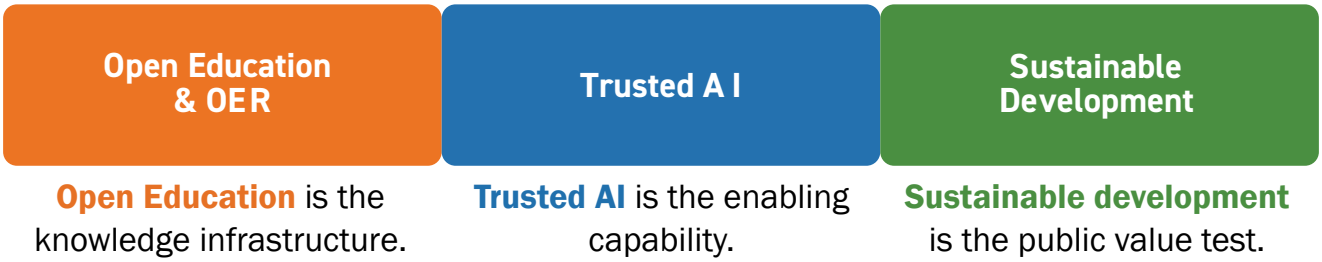
Domain abbreviations: T&L = Teachin & Learning; A&I = Assessment & Academic Integrity; R&I = Research Ethics & Integrity; E&A = Equity Access; DPS = Data Privacy & Security; PTA = Public Trust & Accountability.

The operational tool remains the authoritive source for full indicator wording and evidence capture

Why Open Education Matters More Than Ever Before

Co-creating SDG-informed Open Education with Trusted AI across the Asia-Europe Axis

AUTHORS
João Pita Costa
Anja Polajnar
Davor Orlič
Ebba Ossiannilsson
Joost Groot Kormelink
Ajita Deshmukh
Chetanpal Singh
Weiyao Meng



Executive summary

Artificial intelligence (AI) is reshaping how knowledge is created, accessed and shared. It expands the scale and speed of knowledge production, translation and adaptation, from museums[1] to universities, but can also intensify opacity, bias, commercial concentration, and misinformation. When insufficiently contextualised, AI-mediated knowledge practices may also promote homogenised learning and reduce opportunities for critical reflection,[2,3] putting the openness and quality of knowledge at risk. Its widespread adoption therefore represents a broader structural shift in the knowledge environment within which universities operate. The central policy challenge is not simply whether universities should adopt AI, but how they should govern and shape AI-mediated knowledge practices so that they remain trustworthy, inclusive and sustainable.

Open Education should be considered as a strategic public knowledge infrastructure fit for the AI era, guided by the Sustainability Development Goals (SDGs). Research institutions and ecosystems should engage in Open Education by ensuring that AI-mediated knowledge practices remain trustworthy, human-centred, ethical, inclusive and transparent. Open Education, Trusted AI, and Sustainable Development continuously reinforce one another through collaboration, innovation and feedback (see Figure 1).

This white paper presents a framework founded on a transformative principle, which is that Open Education, open education resources (OER), artificial intelligence (AI), and the SDGs are not independent priorities but interconnected components of the knowledge ecosystem. It connects these three complementary roles and positions Open Education as the strategic public knowledge infrastructure and the primary driver of quality educational transformation for sustainability. Universities occupy the centre of this ecosystem as trusted knowledge orchestrators, connecting people, knowledge, technology and communities. Three illustrative cases – the Open Education for a Better World (OE4BW) initiative; the AI4SDG programme implemented by the Gulf Information Technology Exhibition (GITEX) and International Research Centre on AI (IRCAI); and the Study Webs of Active Learning for Young Aspiring Minds (SWAYAM) platform – show different parts of this ecosystem in practice. Building on these examples, and to put the proposed framework into practice, the paper proposes an Asia-Europe University Sandbox. This would provide a shared, trusted environment in which universities and partners can collaborate for bounded AI-OER pilots, transparent evaluation, capacity building, mentoring and shared SDG challenges.

The proposed actions in this paper thus call on universities and their partners across the Global South and North to make Open Education a core infrastructure with a focus on SDGs. To close the AI–OER–SDG cycle, these principles are to be put in place by university leadership, starting with the Europe-Asia axis: (i) build multilingual knowledge commons; (ii) co-create resource, and establish governance for Trusted AI (iii) resource co-creation; establish governance for Trusted AI; (iv) impact of OE-AI on SDG; and (v) invest in people recognising open contributions; and aligning funding with open education and SDG aims.



Figure 1. *The OE–AI–SDG Circle: Universities orchestrate open, trusted, and sustainable knowledge ecosystems for the public good.*

1 Why Open Education matters more than ever before

Artificial intelligence (AI) use in higher education is no longer marginal or experimental. The 2026's Student Generative AI Survey by the Higher Education Policy Institute (HEPI)[4] found that 95% of surveyed undergraduates used AI in at least one way, while 94% used generative AI to support assessed work. Yet institutional approaches remain uneven, with students receiving inconsistent signals about whether and how such tools should be used.[5] This widespread use represents more than the rapid adoption of a new technology. AI is becoming embedded in the everyday practices through which knowledge is created, accessed, interpreted and shared.[6]

For universities, this shift raises three interrelated concerns: (i) whether they support trustworthy knowledge creation and exchange; (ii) whether they are sustainable in ways that preserve equitable participation and linguistic, cultural and epistemic diversity; and (iii) what responsibilities universities have, as knowledge institutions, to guide this transition in the public interest. At the end of the day, knowledge is a public good. Access to knowledge is essential for equitable participation in today's digital society.[7] As new technologies reshape how knowledge is created and shared, publicly funded knowledge must remain accessible and sustainable as a public good, also in regards to knowledge produced and transferred through higher education institutions. For societies to thrive, higher education – shaped by the values of inclusion and justice – should be open, accessible and equitable at both global and local levels. Open Education supports this responsibility through OER, open pedagogies, and collaborative learning practices.

The UNESCO Recommendation on OER[8] defines OER as learning, teaching and research materials in any format that are in the public domain or released under an open licence permitting no-cost access, reuse, repurposing, adaptation and redistribution. These permissions become particularly important as AI accelerates the production, translation and adaptation of educational content. They enable educators and learners to examine source materials, adapt resources to their educational contexts and share revisions. Sustaining these practices requires institutional commitment. The UNESCO Recommendation on OER identifies five action areas: capacity building, supportive policy, inclusive and equitable access to quality OER, sustainability models, and international cooperation.[8] These provide a policy basis for universities to support Open Education as AI use expands, while the SDGs orient this work towards inclusive education and sustainable development.

A key recommendation emerging from this work is the establishment of an Asia–Europe University Sandbox, providing a shared and trusted environment in which universities and partners from Asia and Europe can collaboratively experiment with the intersection of Open Education, AI, and SDGs. As we further detail in the Appendix B, universities need to move beyond simply adopting AI technologies and take a more active role in shaping how AI-supported knowledge is created, validated, shared and applied for societal benefit. The proposed AI sandbox for Open Education on SDG would translate this principle into practice by bringing universities and partner organisations together around collaborative experimentation, shared evidence and data, responsible AI pilots, OER co-creation, and SDG-focused learning. But let us first start by explaining in more debt what we mean and what we prioritise.

Open Education and OER provide knowledge accessible to the common good, education for all, with openly licensed, adaptable, and participatory knowledge resources and practices, and are part of the larger movement on openness for a better world, supported by several UNESCO initiatives.[8] **Trusted AI** refers to AI systems that support scale, translation, personalisation and analytical support under conditions of transparency, accountability, human oversight, and rights protection.[9] **SDG** provides the broader public values orientation against which technological innovation should be assessed, with universities at the centre as knowledge ecosystem orchestrators, connecting learners, educators, researchers, libraries, governments, communities, industry and international organisations.[10]

Open Education and OER are part of the larger movement on openness for a better world, supported by several UNESCO initiatives.[8] Openness matters more

than ever in the era of AI and sustainability.[11] At a policy level, "openness" in education implies embracing the philosophy of "open movement". This term can refer to several different concepts, ranging from the global socio-technical philosophy of open access and collaboration. It is focused on making human knowledge, tools, and digital resources freely accessible, transparent and reusable. Key areas include open-source software, open data, open science, open government, open education, open collaboration, open AI, open pedagogy, open research, and open educational resources. As AI reshapes these ecosystems, Open Education matters because it provides resources and practices through which universities can address these questions while keeping educational purposes and public value central.[11]



2 The challenges and opportunities to Open Education in the AI era

Universities face three connected challenges in ensuring that AI supports Open Education: enabling meaningful participation, maintaining trustworthy educational resources, and preserving knowledge diversity.

First, making content openly available does not ensure that everyone can use or contribute to it. Learners also need discoverability, accessible formats, multilingual support, digital capability, devices, connectivity and opportunities to adapt resources to local needs, contexts and challenges.[3] AI can assist translation, captioning, summarisation, and adaptation, but universities must ensure that these functions reduce rather than reproduce exclusion. AI needs to support meaningful public participation and inclusive access.[12]

Second, the growing availability of AI-generated content increases the importance of provenance and human judgement. Such content can be persuasive without being reliable. OER ecosystems can make sources, licences, authorship, revisions and review processes visible. AI should take a trusted, human-centred approach[2,9,13] to augment these practices by supporting traceability and quality checks, while final educational and ethical responsibility remains human and institutional.[14,15] Localisation, in this sense, is a human-centred approach that protects knowledge diversity.

Third, concentration around a small number of AI models and platforms risks limiting the diversity of educational knowledge. Open licences and participatory OEP enable communities to translate, contextualise, challenge and improve resources. This is essential across the Asia–Europe axis, where educational priorities, languages, cultures and institutional capacities vary widely. AI should support linguistic diversity and contextualisation of educational content across different cultural and institutional settings.[13]

Table 1 connects these challenges to design requirements for AI-supported Open Education, showing how universities can address barriers to participation, strengthen trust, and preserve opportunities for local adaptation.

AI-era Tension	Open Education Opportunity
Abundance without confidence	Transparent provenance, open review and revision
Efficiency without understanding	Human judgment, pedagogy and contextualisation
Access without equality	Accessibility, multilingualism and support
Innovation without accountability	Open governance, rights and public oversight
Scale without diversity	Local adaptation and distributed knowledge creation

Table 1. *AI-era tensions and opportunities for Open Education to address them.*

3 The OE–AI–SDG Triangle: Open Education as knowledge infrastructure

Going back to the triangular framework presented in Figure 1, which express the organic relationship between Open Education, AI, and Sustainable Development, we argue or their complementary roles. The framework connects how educational knowledge is made available and shared, how technology supports these practices, and the societal outcomes they should advance.

Open Education provides the foundation. Through OER, open pedagogies, and collaborative learning practices, it enables knowledge to be accessed, adapted, reused and shared. Treating Open Education as public knowledge infrastructure places these activities within universities' long-term responsibilities for educational equity, lifelong learning and international knowledge exchange.

AI provides supporting capabilities. It can assist learning, research, accessibility, multilingual knowledge creation, and evidence-informed decision-making through ethical, transparent, accountable and human-centred technologies. These capabilities should operate under conditions of transparency, accountability, human oversight, and rights protection – the principles of Trusted AI.[9] Their contribution to Open Education relies on institutional governance that supports quality education and responsible use.

Sustainable Development provides the purpose. The SDGs orient educational activity towards benefits for individuals, communities, institutions, and society.[16] They provide a basis for assessing whether AI-supported Open Education contributes to inclusive learning and addresses societal needs.

Human wisdom guides the relationships among all three components. Ethical judgement, critical thinking, cultural diversity, creativity, and human values ensure that AI and Open Education continue to serve humanity rather than replace it.

Within this ecosystem, universities act as trusted knowledge leaders that integrate research, education, policy, technology, and community engagement to maximise public benefit and strengthen international collaboration.[10]

The three case studies summarised in Table 2 and extended in Appendix A illustrate practices relevant to this framework. Each contributes to understanding open education builds capacity, research generates knowledge and innovation networks translate knowledge into AI solutions for sustainable development.

The community of practice OE4BW[17] demonstrates an AI-enhanced open educational ecosystem in which AI can support the creation, translation, adaptation, discovery and personalisation of OER while connecting educators, mentors and learners across linguistic, cultural and institutional contexts. Its relevance lies in examining how AI can expand openness, accessibility, inclusion and knowledge reuse while preserving human agency and quality.

The GITEX-IRCAI AI4SDG programme[18] focuses on the innovation ecosystem linking universities, companies and governments to match scientific and technological capabilities with concrete challenges across the 17 SDGs, providing a model for collaborative AI development, deployment and impact assessment oriented toward societal needs.

The SWAYAM platform[19] demonstrates how publicly supported open education can operate as shared digital infrastructure at scale, combining distributed content creation, open access, institutional curation, learner choice, assessment and academic recognition. Its relevance in the AI era is particularly strong because it illustrates the transition from universities as knowledge providers to knowledge stewards, where AI can support discovery, explanation and recommendation while institutions retain responsibility for contextualisation, validation, curriculum alignment and trust.

Use case	Ecosystem Role	Main Focus	Potential Contribution
OE4BW	Education and knowledge sharing	OER and educational ecosystems	Create, translate, adapt, discover and personalise open resources
GITEX-IRCAI AI4SDG	Innovation and deployment	Universities-companies-governments	Match challenges, expertise and solutions for the 17 SDGs
SWAYAM	Research and development	AI for Sustainability & Sustainable AI	Generate knowledge, methodologies and technologies

Table 2. *Case studies and their role in the ecosystem, focus and contribution*

From these case studies as described in Table 2, we propose an Asia–Europe University Sandbox as a practical means of implementing the triangular framework. The sandbox provides a shared, trusted environment where universities and partners can co-develop and test AI-enabled Open Education addressing SDG-related challenges. Its activities would follow UNESCO's ethical principles for AI, including human agency, fairness and inclusion, transparency and explainability, accountability, sustainability and respect for human rights.

The proposed sandbox would organise collaboration around five interconnected actions, implements responsible AI–OER pilots across diverse Asia-Europe institutions:

i. publishing and openly evaluating the evidence generated;
ii. building capacity among educators, students, staff and leaders in AI literacy, ethics and OER;
iii. embedding and scaling successful practices through networks such as OE4BW; and
iv. establishing a shared challenge agenda focused on priority areas where AI-enabled Open Education can contribute to the SDGs.

4 SDG by trusted AI for university ecosystem leadership

4.1 Universities as leaders of quality, inclusive, scalable Open Education to accelerate the SDGs

Universities occupy a distinctive position because they combine education, research, innovation, lifelong learning, knowledge validation, public engagement, and international networks. Their task is to create the conditions under which knowledge remains open, credible, adaptable, and socially relevant.[10,9] The dimensions in Table 3 translate the AI–OER–SDG framework into a coherent institutional role.

Quality and Inclusion. Universities should establish assurance processes for AI-assisted OER that document sources, licences, human contributions, model use, limitations, and revision histories. Quality requires transparent provenance, academic integrity, disciplinary and pedagogical review, explainability, and mechanisms for correction. Inclusion should be designed from the outset, ensuring accessibility, multilingualism, cultural

Dimension	Key Indicators	Policy Questions
OPENNESS	• Share of resources with open licenses • Reuse and adaptation • Metadata completeness	**Can others legally and practically find, understand, and improve the resource?**
TRUST	• Provenance disclosure • Review completion • Correction time • Reported harms	**Can users verify outputs, question the system, report problems, and obtain accountability**
INCLUSION	• Languages • Accessible formats • Participation and representation of knowledge and data from underserved groups • Low-bandwidth availability	**Who benefits, who is missing and who bears the risks?**
CAPABILITY	• Staff/student training • Mentoring • Communities of practice • Governance maturity • Shared compute resources • SDG data	**Do institutions have the human, organisational and technical capacity to sustain the work?**
SUSTAINABILITY	• **Financial**: recurring funding, repository continuity, vendor independence • **Environmental:** environmental assessment, energy use and efficiency	**Will the system remain viable and publicly accountable? Can the system remain financially viable over time while minimising its environmental impact?**
SDG IMPACT	• Use in SDG curricula and community projects • Partnerships • Documented and monitored educational, social, or sustainability outcomes	**What measurable public value has been created, and for whom?**

Table 3. *Indicators and Policy Questions*

relevance, affordability, gender equality, and meaningful participation by marginalised communities. AI-supported education should be tested across different languages, abilities, connectivity levels, and forms of digital literacy, making open education central to an inclusive ecosystem.

Scalability and Sustainability. Scaling should extend access to quality learning beyond university enrolment while preserving local adaptation rather than imposing uniformity. Shared repositories, open standards, reusable workflows, and cross-institutional mentoring can expand reach, while AI can reduce translation and adaptation costs with appropriate human and regional oversight. Sustainability encompasses environmental, financial, institutional and social dimensions and aligns with the SDGs and OER principles. Initiatives require durable funding, recognised workloads, staff development, repository maintenance, and governance, while AI choices should consider energy use, vendor dependence, data protection, and long-term public value.

Universities should therefore evolve from knowledge providers into knowledge ecosystem orchestrators that curate evidence, enable participation, support learner agency, govern AI, connect stakeholders, and evaluate public impact.

4.2. Co-creation and governance of Trusted AI for Open Education

Trusted AI is an institutional property created through governance, evidence, accountability, and sustained human capability, not a product label. In Open Education, trust requires AI-assisted resources to be traceable, contestable, correctable and continuously improvable. Universities should define responsibility for approving, publishing, monitoring and withdrawing AI-supported resources. Co-creation should involve shared problem definition, representation of affected learners and communities, transparent AI use, access to evidence, appropriate licensing and attribution, contextual review, and post-release feedback. Participation should be adequately resourced, recognised and ethically governed.



Open	Human-centred	Equitable	Accountable
licensing, provenance, interoperability	agency, judgment, redress	accessibility, language, representation	roles, audit, correction
Sustainable	**Participatory**	**Evidence-led**	**International**
funding, capacity, environmental care	co-design, review, community voice	pilots, evaluation, transparency	shared learning with local ownership

Table 4. *Guiding principles*

The precautionary principle and Collingridge dilemma highlight the need for anticipatory governance: early impacts may be uncertain while design choices remain flexible; later, impacts become clearer but institutional dependence makes change harder. Universities should therefore use impact assessments, bounded pilots, documented assumptions, protected alternatives, and defined review points before systems become embedded. Governance should address data protection, copyright and licensing, accessibility, bias, environmental impacts, academic integrity, security, explainability, human oversight, and redress, while distinguishing low-risk support from high-stakes decisions affecting assessment, progression, admission or student welfare. These principles, summarised in Table 4, provide guiding principles for responsible AI development aligned with this vision.

5 Conclusion and call to action

The AI era does not reduce the importance of Open Education; it makes openness a condition of trustworthy knowledge. Universities now face a choice: They can adopt AI through fragmented tools and proprietary platforms, or they can shape an open knowledge ecosystem in which technology remains accountable to educational purpose, human rights and sustainable development. this white paper proposes Open Education and OER as strategic public knowledge infrastructure; Trusted AI as a governed enabling capability; and Sustainable Development as the public-value test. The practical agenda is achievable, requiring institutional ownership, multilingual knowledge commons, inclusion by design, explicit and resourced co-creation, risk-based AI governance, measures of public impact, recognition of open work and funding conditions that protect openness and interoperability. ASEF[20] and OE4BW[17] can help connect policy, mentoring, and demonstrator communities[21] pioneering across Asia and Europe.

The call to action is therefore direct: make Open Education core infrastructure now, before AI-mediated knowledge systems become too embedded to reshape. By acting early, universities can turn the Collingridge dilemma into an opportunity – guiding technological development while choices remain open and building sustainable knowledge commons reimagined in a post-2030 world. The argument leads us to eight actions, each of which hints at a practical shift rather than a general aspiration:

1. **Make Open Education a core infrastructure,** embedding the Dubai recommendations15 and OER/OEP in institutional strategies for AI, curriculum, research, libraries, digital transformation, internationalisation and sustainability. Assign ownership, recurring budgets, and repository support.
2. **Build multilingual knowledge commons,** developing cross-border OER collections for SDG challenges, with open recognition, licences, interoperable metadata, regional review, and AI-assisted translation followed by human validation. Go beyond translation by enabling contextual adaptation and incorporating local examples, policies, data and knowledge traditions, with validation in the intended educational context.
3. **Design inclusion by default**, applying accessibility standards, low-bandwidth options, low energy compute, AI-ready datasets, multilingual interfaces and participation by underrepresented groups from the beginning of every initiative. Involve responsible stakeholders.
4. **Resource co-creation**, defining who participates, at what stages and with what authority. Providing mentoring, recognition, attribution and compensation where appropriate.
5. **Establish governance for Trusted AI**, creating risk-based review, impact assessment, disclosure, human oversight, monitoring and redress processes for AI-assisted education and OER with verifiable sources where appropriate.
6. **Measure public and SDG impact**, tracking openness, reuse, recognition, adaptation, accessibility, learner participation, cross-border collaboration and contributions to relevant SDG targets aligned to locally defined SDG aims.
7. **Invest in people and reward openness**, developing Open Education and AI literacy for leaders, educators, librarians, researchers, and students. Recognise OER, OEP, mentoring, international collaboration, open science, and community engagement in workload and promotion.
8. **Align funding and procurement**, making openness, interoperability, accessibility and FAIR-aligned stewardship conditions of public funding and technology procurement, while avoiding vendor lock-in.

Moreover, five multistakeholder actions – implement, publish, build capacity, embed and scale, and set the challenge agenda (see Appendix B) – position the proposed Asia–Europe University Sandbox as more than a technology-testing environment. It becomes a mechanism through which universities can exercise their wider responsibility as trusted knowledge institutions, ensuring that AI innovation contributes to open, inclusive, ethical and sustainable development. Table 5 describes the prioritisation of responsibilities and contributions of a wide range of stakeholders that need to be part of this discussion and hands-on implementation.

Level and actors	Primary responsibility	Priority contribution
International: UNESCO, OECD (OECD, 2024) UN agencies	Norms, rights and shared standards	OER policy, AI ethics, SDG indicators and financing
Asia-Europe: ASEF, ASEM networks, European Commission	Convening and evidence exchange	Policy dialogue, demonstrators, shared metrics, and equitable partnerships
Regional networks: OE4BW and partners	Mentoring, localisation, and community building	Mentor-developer matching, multilingual adaptation, and thematic OER hubs
National: ministries, regulators, funders	Policy, finance, and enabling regulation	Open licensing, repository investment, quality frameworks, and recognition
Institutional and local: universities, libraries, communities industry	Implementation, validation, and social relevance	Governance, staff support, co-creation, feedback, and responsible technology

Table 5. *Stakeholders responsibilities*



Authors

Dr João PITA COSTA
International Research Centre on AI under the auspices of UNESCO & Journal of AI for Sustainable Development, Slovenia

Dr Anja POLAJNAR
International Research Centre on AI under the auspices of UNESCO & Open Education for a Better World, Slovenia

Mr Davor ORLIČ
International Research Centre on AI under the auspices of UNESCO, Slovenia

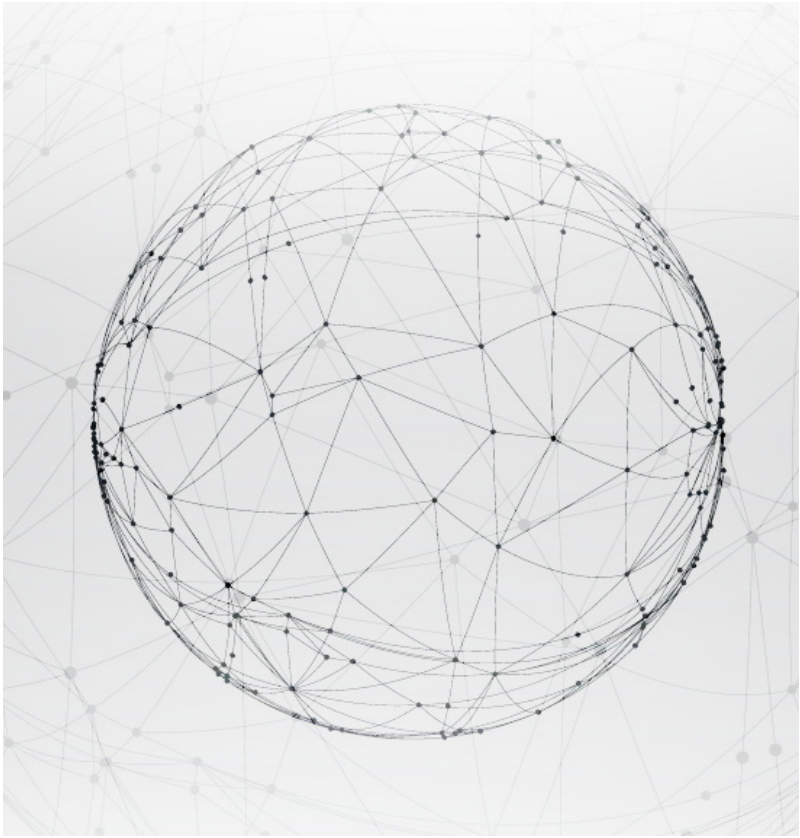
Prof Ebba OSSIANNILSSON
Ossiannilsson Quality in Open Online Learning Consultancy, Sweden

Ir Joost GROOT KORMELINK
TU Delft, Netherlands

Dr Ajita DESHMUKH
MIT Art Design and Technology University, India

Dr Chetanpal SINGH
RMIT University, Australia

Dr Weiyao MENG
University of Nottingham, United Kingdom



References

1. UNESCO. (2026b). *UNESCO-ICOM 2026 report on the use of artificial intelligence in museums*. UNESCO Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000399452_eng

2. European Commission. (2026, July 12). *AI Act*. Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

3. UNESCO. (2024). *Dubai Declaration on Open Educational Resources (OER): Digital public goods and emerging technologies for equitable and inclusive access to knowledge*. UNESCO Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000392271

4. Stephenson, R., & Armstrong, C. (2026, March 12). *Student Generative AI Survey 2026*. Higher Education Policy Institute. https://www.hepi.ac.uk/wp-content/uploads/2026/03/HEPI-Report-199-Gen-AI-Survey-2026.pdf

5. LeBlanc, P., Gamby, T., & Siemens, G. (2026). *Reclaiming purpose: the university in an AI World*. John Wiley & Sons, Limited. https://www.wiley.com/en-fr/shop/general-introductory-business-management/reclaiming-purpose-the-university-in-an-ai-world-p-9781394407477

6. Akinwalere, S., & Chang, K. (2026). The symbiotic roles of artificial intelligence and human intelligence in advancing knowledge ecosystem. *Journal of Knowledge Management*, 1—42. https://doi.org/10.1108/JKM-03-2025-0362

7. UNESCO. (2026a). *The right to know: open knowledge ecosystems and AI*. UNESCO Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000399308

8. UNESCO. (2019, November 25). *Recommendation on Open Educational Resources (OER)*. https://www.unesco.org/en/legal-affairs/recommendation-open-educational-resources-oer

9. OECD. (2024, July 1). *AI principles*. OECD. https://www.oecd.org/en/topics/sub-issues/ai-principles.html

10. Commonwealth of Learning (COL). (2025, January 24). *Harnessing AI for a brighter educational future*. https://thecommonwealth.org/news/harnessing-ai-for-education

11. Zawacki-Richter, O. (2026). What can Open, Distance, and Digital Education (ODDE) contribute to a sustainable future? Implications for Leadership. In D. C. Lim (Ed.), *Critical Perspectives on AI, Digitalisation, and Human-Centred Learning: Towards Reclaiming Digital Education Futures* (pp. 65-77). Springer Nature Singapore. https://doi.org/10.1007/978-981-92-3113-3_5

12. World Bank Group. (2026). *Digital public infrastructure and services*. World Bank. https://www.worldbank.org/ext/en/topic/digital-and-ai/digital-public-infrastructure-and-services

13. UNESCO. (2023, September 7). *Guidance for generative AI in education and research*. UNESCO. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

14. Jermol, M., & Costa, J. P. (2026). Onwards to an ethical and bias aware education for sustainability through AI. Special collection on AI and Education for Democracy. *Journal of AI for Sustainable Development*, (to appear).

15. UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. UNESCO. https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence

16. Urbančič, T., Polajnar, A., & Jermol, M. (2019). Open Education for a Better World: A mentoring programme fostering design and reuse of Open Educational Resources for Sustainable Development Goals. *Open Praxis*, 11(4). https://doi.org/10.5944/openpraxis.11.4.1026

17. Open Education for a Better World (OE4BW). (2026). *International online mentoring programme: Open education for the UN Sustainable Development Goals*. https://oe4bw.org/

18. Zaouini, M., Costa, J.P., Rahmani, Y., Kassis, R., Stopar, L., Souss, S., Lamgari, A., & Mochariq, O. (2025). Addressing water sustainability challenges in North Africa with artificial intelligence. *Proceedings of the SIKDD 2025*. https://doi.org/10.70314/is.2025.sikdd.17

19. Swayam (2017). *Framework for Universities to conduct examinations for SWAYAM Courses*. https://swayam.gov.in/about_university

20. Asia-Europe Foundation. (2026). *7th ASEF Higher Education Innovation Laboratory: The role of universities in advancing AI for good*. https://asef.org/projects/asefinnolab7/

21. Bong, Y., & Yunus, S. (2025, February). White paper for universities on navigating artificial intelligence innovation ecosystems in AI for sustainable development. *ASEF Higher Education Innovation Laboratory*. https://doi.org/10.25878/QWXR-RP84

Appendices

This white paper is founded on a transformative principle: Open Education and OER, AI, and the SDGs are not independent priorities, but interconnected components of a single knowledge ecosystem. The appendices translate this principle into practice in two complementary ways. First, it illustrates how the proposed knowledge ecosystem can operate through three complementary case studies, each demonstrating different dimensions of the interaction between Open Education, OER, AI, and the SDGs. Second, it moves from illustration to implementation by proposing concrete actions for universities, policymakers, communities, and other stakeholders to strengthen this ecosystem and turn its principles into sustained institutional and societal impact.

A. The OE–AI–SDG case studies and open sustainable knowledge infrastructure

We position Open Education as strategic public knowledge infrastructure, with universities at the centre as trusted knowledge orchestrators connecting people, knowledge, technology, governance, communities, and society to create long-term public value across the Asia–Europe region. To explore how AI can support the evolution of Open Education and contribute to sustainable development, three complementary case studies (A1–A3) are particularly relevant. They represent different but interconnected levels of the emerging ecosystem: open educational resources and learning networks, collaborative AI innovation, and research and development infrastructures for sustainable AI.

A1. OE4BW: Open educational resources and learning ecosystems

The Open Education for a Better World (OE4BW) network provides a first case study for examining how AI can strengthen OER and the wider educational ecosystems in which they are created, shared and reused.[9,20] Beyond the production of individual OER, the network connects educators, mentors, institutions and learners across countries, creating an ecosystem for knowledge exchange and capacity building. Participants, mentors or mentees, transform their knowledge and ideas into openly accessible educational solutions with tangible social impact.

At the heart of OE4BW is a structured mentoring process. Selected developers work with experienced international mentors who support them in designing, improving and implementing their projects. Participants are also connected through thematic and regional hubs, creating communities in which knowledge, experience, and good practices can be exchanged across countries and disciplines. The programme therefore goes beyond the creation of educational materials: it builds international partnerships and strengthens participants' capacity to use open education as an instrument for sustainable development.

OE4BW is based on the principle that knowledge should be accessible and reusable, particularly when it can contribute to solving pressing societal challenges. Projects may address areas such as quality education, environmental sustainability, health and well-being, digital skills, social inclusion, gender equality, and opportunities for underserved communities. By combining open education, mentorship, international collaboration, and the SDGs, OE4BW enables participants to turn local ideas into resources that can be shared, adapted and used globally.

AI is enhancing this ecosystem by supporting the creation, translation, adaptation, discovery and personalisation of OER, while also helping to connect educational resources with different cultural, linguistic and institutional contexts. The case study is therefore relevant for investigating how AI can contribute to more accessible, multilingual and reusable open education while preserving the principles of openness, quality, human agency and inclusiveness. The OE4BW case is explored here as a model of AI-enhanced open educational ecosystems, where AI acts not simply as a content-generation tool but as an infrastructure for connecting knowledge, educators and learners.

A2. GITEX–IRCAI AI4SDG: Matching universities, companies and governments

The Gulf Information Technology Exhibition (GITEX) – International Research Centre on AI (IRCAI) AI4SDG programme provides a second case study focused directly on the development and deployment of AI solutions for the 17 SDGs. Its distinctive feature is the creation of connections

between universities, companies and governments, matching research and technological capabilities with concrete societal and policy challenges. This model moves beyond the traditional transfer of knowledge from academia to industry by creating a collaborative environment in which public-sector challenges can be translated into AI research and innovation opportunities. Universities contribute scientific expertise and talent, companies contribute technological and implementation capacity, and governments contribute real-world challenges, institutional knowledge and pathways to deployment.

The Network for Artificial Intelligence, Knowledge and Sustainable development (NAIXUS) provides a research collaboration base for this use case to take place, centred on the research and development dimension of the ecosystem. NAIXUS brings together institutions and their SDG experts collaborating on research related to AI for Sustainability and Sustainable AI, creating a network through which researchers can exchange knowledge, develop joint research agendas and contribute to the scientific foundations of responsible and sustainable AI. The network is particularly relevant because it connects two complementary perspectives. *AI for Sustainability* considers how AI can accelerate progress toward environmental, social and economic sustainability, while *Sustainable AI* examines how AI itself can be developed and deployed in ways that minimise environmental and societal costs and respect principles of responsible innovation.

The case study is particularly relevant for exploring how AI innovation ecosystems can be organised around the SDGs, including mechanisms for identifying challenges, matching stakeholders, developing solutions, evaluating impact and scaling successful applications. It also provides an opportunity to examine how AI can be directed toward measurable contributions to sustainable development rather than being developed independently of societal priorities. Moreover, NAIXUS can support the development of a research and innovation layer with universities at the core of the conversation, connecting scientific knowledge, emerging technologies, sustainability challenges and policy needs. Its potential contribution extends from fundamental and applied research to the development of methodologies, indicators, assessment frameworks, and technologies that can support evidence-based decision-making.

A3. SWAYAM

The Study Webs of Active Learning for Young Aspiring Minds (SWAYAM) platform[19] demonstrates how Open Education can function as shared public digital infrastructure. An initiative of the Ministry of Education, it embodies a public-good principle: public investment in knowledge should create public value beyond the institution that produces it. SWAYAM brings together courses developed by universities and national institutions through 10 National Coordinators – AICTE, NPTEL, UGC, CEC, NCERT, NIOS, IGNOU, IIM Bangalore, NITTTR and INIs. This distributed model enables co-creation and sharing of academic expertise. The platform currently reports 20,370+ courses, 6.6 crore+ enrolments, 62.8 lakh+ certifications and 205 partnering institutes. Courses are free to access, with certification through proctored examination. Its sustainability rests on shared infrastructure, distributed content creation and reuse:

Create once → improve collectively → reuse many times → reach millions

Publicly supported academic expertise can therefore reach learners across institutions and geographies without every institution having to independently develop comparable courses. The critical university layer turns open access into meaningful academic choice. Universities can map SWAYAM courses to their curricula, learning outcomes and semester structures, identify suitable courses, inform students and enable them to choose courses aligned with their academic pathways. Under the UGC framework, institutions can recognise SWAYAM courses for credit transfer, with up to 40% of programme courses in a semester permitted through SWAYAM. Since 2024, universities that have adopted SWAYAM for credit transfer can also conduct SWAYAM examinations for their own students, aligned with their academic calendars, increasing flexibility and integration with regular university processes. This institutional integration is particularly important in the age of generative AI. If students can learn almost anything from AI, why would they choose universities or SWAYAM? SWAYAM offers more than information: structured learning, identified academic sources, assessment, certification and academic credit. AI can explain and recommend; universities determine whether a course is appropriate for a particular programme, semester and learner.

This is human wisdom in practice. Universities exercise judgement about what is relevant, which knowledge is credible, how it fits the curriculum, what choices students should have and what learning merits academic recognition. AI can support these decisions; universities provide context, accountability and trust. SWAYAM is also expanding access to AI knowledge, with 100+ AI-related courses and more than 50 lakh enrolments reported in 2026. Its larger significance, however, is the model itself: as AI expands the volume of available knowledge, the need for trusted curation, contextualisation and academic validation grows.

Public investment → Shared infrastructure → Co-creation → Open access → University curation → Learner choice → Academic recognition → Public value

SWAYAM therefore illustrates the shift from Universities as Knowledge Providers → Universities as Knowledge Stewards: institutions need not produce every course – they can create, contribute, curate, validate, connect and recognise knowledge within an open ecosystem.

B. Recommendation: Building an Asia–Europe University Sandbox for AI, Open Education and the SDGs

A key recommendation emerging from this work is the establishment of an Asia–Europe University Sandbox, providing a shared and trusted environment in which universities and partners from Asia and Europe can collaboratively experiment with the intersection of Open Education, AI, and SDGs. Universities need to move beyond simply adopting AI technologies and take a more active role in shaping how AI-supported knowledge is created, validated, shared and applied for societal benefit. The proposed AI sandbox for Open Education on SDG would translate this principle into practice by bringing universities and partner organisations together around collaborative experimentation, shared evidence and data, responsible AI pilots, OER co-creation, and SDG-focused learning.

The first priority is implementation. Recommendations arising from ASEFInnoLab7 should be translated into practical initiatives across ASEF member states and participating research universities, with mechanisms for monitoring progress and sharing outcomes. Rather than institutions experimenting independently, the sandbox would provide a common environment in which approaches can be tested, evaluated and adapted across different institutional and cultural contexts. This would also create an evidence base for identifying what works, what does not, and under what conditions responsible AI and Open Education initiatives can be successfully scaled.

The second priority is to publish and strengthen the evidence base. Collaboration with IRCAI to develop a dedicated journal collection on the OE–AI–SDG nexus would provide an opportunity to convert practical experimentation into research evidence. Universities participating in the sandbox could contribute case studies, empirical findings, conceptual frameworks and evaluations of responsible AI initiatives. This creates an important connection between experimentation and scholarship: universities would not only implement innovative approaches but also critically evaluate and openly communicate their outcomes.

A third component is capacity building. Collaboration with the UNESCO OER Dynamic Coalition on an online mini-course and a VideoLectures.Net channel could make knowledge developed through the initiative accessible beyond participating institutions. Capacity building is particularly important because responsible AI adoption cannot depend solely on technological infrastructure. Academics, students, professional staff, and institutional leaders require the knowledge and skills to understand AI capabilities and limitations, apply ethical principles, develop and adapt OER, and assess the societal consequences of AI-supported educational practices.

The sandbox should subsequently provide a pathway to embed and scale successful practices. Integration with the OE4BW programme could support international co-creation and mentoring while extending successful initiatives to wider communities. This is important for ensuring that promising projects do not remain isolated pilots. Approaches that demonstrate educational, social or sustainability value could instead be adapted, reused and transferred across institutions and regions.

Finally, the initiative should establish a shared challenge agenda by identifying ten priority challenges where AI can strengthen Open Education and accelerate progress towards the SDGs. This would provide universities with common problems around which interdisciplinary and international collaboration can develop.



Universities as Catalysts for Aligning AI Innovation with the Sustainable Development Goals

AUTHORS
Sharina Yunus
Francesco Amighetti
Emese Belenyesi
Nimisha Parashar
Muhammad Sohail
Nur Lisa Zaharin

Executive summary

Universities report their contributions to the Sustainable Development Goals (SDGs) and their research and innovation in artificial intelligence (AI). Far less visible is where these strengths come together to address public problems, and whether the work reaches implementation. Without a clear route from university expertise to public use, valuable capability stays scattered while SDG progress lags.

This paper helps university leaders and public agencies decide which university strengths to draw on for an SDG challenge, which implementation functions need strengthening, and how to connect the work to public use.

It introduces a two-stage diagnostic to examine that link. Stage 1 maps publicly visible AI and SDG activity; Stage 2 looks for evidence of pathways into policy, real-world testing, practitioner preparation, independent evaluation and adoption. Applied to 12 universities across Europe and Southeast Asia, it found that seven showed broad activity in both AI and the SDGs. Yet only one showed strong public evidence of an implementation pathway, and that university was not among the seven. Stage 2 shows three gaps in the pathway: no defined role in delivery, unclear testing route to adoption, and lack of visibility in evaluation and attribution.

These findings are exploratory. The mapping does not rank institutions or represent national systems, and limited public evidence does not mean that implementation is absent. It does show why visible capability should not be assumed to translate into public benefit. Universities can contribute in different ways: by developing AI, defining public problems, bringing sector and community knowledge, training practitioners or independently evaluating systems in use.

The paper proposes four actions to close these gaps, and these are to be explored and chosen according to the gaps a country finds rather than completed in sequence: make university capabilities visible; build challenge-based collaboration; set up a national AI-for-SDG platform; and share methods and evidence across the region. Examples from seven countries illustrate how universities can take defined roles in challenge-based collaboration, shared infrastructure and independent evaluation. For any pilot initiative arising from these actions, it is important to name a public owner, plan independent evaluation from the start, and set a clear rule to adopt, adapt or end it. Without these conditions, AI-for-SDG capability risks staying untapped.

1 Connecting university AI-for-SDG capability to national delivery

With four years remaining until 2030, only 36% of assessable Sustainable Development Goal (SDG) targets are on track or making moderate progress.[1] Artificial intelligence (AI) may help, but its effects depend on how it is chosen and used: research suggests that AI could support progress on 134 SDG targets while inhibiting progress on 59.[2]

For governments, the challenge extends beyond funding AI tools. Public agencies need the capacity to define the problem, decide whether AI is appropriate, test systems in real settings, manage risks, prepare practitioners and evaluate whether the results benefit people.[3] Without a route through these steps, promising work may remain a pilot or fail when introduced into public services.[3,4]

Universities can contribute across that route, including institutions that do not develop AI systems themselves. Their strengths may lie in sector knowledge, community relationships, practitioner training, real-world testing or independent evaluation. They can also bring government, industry and communities together around a defined public challenge.[5] The question is how to identify those complementary strengths and give them a clear role in delivery.

1.1 Why SDG 4 and SDG 17 matter

This paper considers AI applications across the SDGs, with particular attention to two goals that help make implementation possible. SDG 4 concerns education and lifelong learning: people need the skills to use, question and govern AI responsibly.[6] SDG 17 concerns partnerships: knowledge, technology and resources must connect across institutions, sectors and borders.[7] Universities contribute to both, through education and research and by helping partners apply and evaluate AI in public settings.

1.2 What this paper does

Universities often document SDG contributions for reporting and rankings while recording AI research and innovation separately. Neither account alone shows clearly whether AI expertise is being applied to SDG priorities or connected to public implementation. This paper brings the two together in a two-stage diagnostic (Section 2).

The diagnostic is intended to help university leaders and public agencies identify available strengths, investigate missing connections and decide which partners and implementation functions a public challenge requires. We test it with public evidence from 12 universities across Europe and Southeast Asia, then draw out four actions to close the gaps it reveals.



2 What the mapping reveals about the delivery gap

How the mapping was done

The two-stage framework at a glance

Stage 1 describes publicly visible AI and SDG activity and Stage 2 assesses evidence of a pathway into implementation. A strong Stage 1 profile need not mean a strong Stage 2 result.

Stage 1 places each university in one of four indicative profiles (Figure 1). Profiles were assigned by each author, using the Stage 1 bands (High 18–24, Moderate 12–17, Limited 6–11 on each axis) as a guide, together with qualitative evidence of where activity is concentrated, and were checked in peer validation. Following the team's scoring guide, they are broad positions, not precise categories derived from small score differences.

Stage 2 rates evidence of an implementation pathway as limited, emerging, or strong, drawing on seven implementation functions: adoption support, formal advisory role, testing in real settings, policy and decision support, independent evaluation of AI systems, a named national or sector role, and recognition in implementation. These are recorded at institution level and may come from different projects, so several functions at one university do not necessarily form a connected pathway within one project.

Independent evaluation means assessing whether an AI system used for a public purpose works as intended, and fairly, by a party that did not build or sell it. Publications, rankings and general partnerships show activity but do not by themselves demonstrate a pathway.

Stage 1 activity map and possible consortium roles

Four indicative profiles based on publicly visible AI and SDG activity

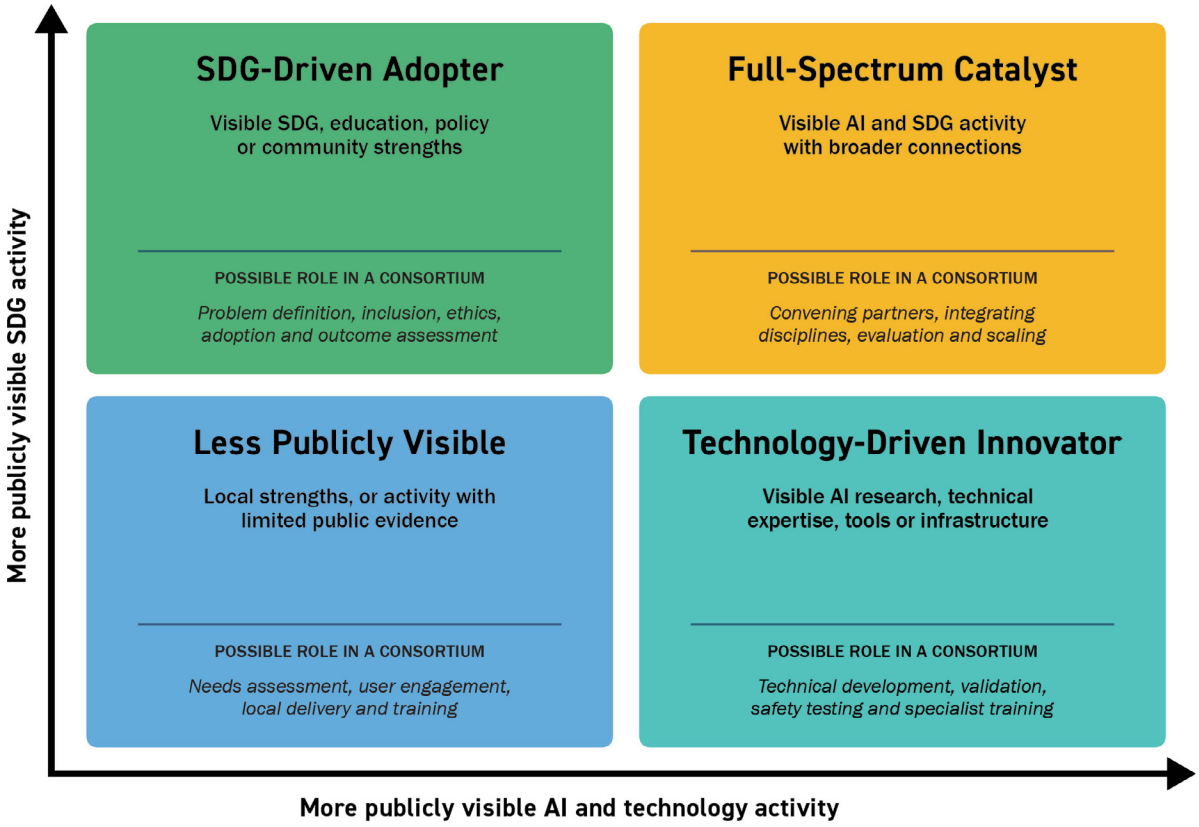


Figure 1. *Stage 1 activity map and possible consortium roles showing four indicative profiles based on publicly visible AI and SDG activity. Profiles show where visible activity is concentrated; they are not rankings or stages of maturity. Source: Authors' two-stage framework.*

Evidence collection and review. Public evidence of how universities apply AI to SDG priorities is scattered across institutional pages, project announcements and local-language documents. We tested the framework on 12 universities the authors know well: their own universities, where applicable, and familiar comparators with visible activity in AI, the SDGs, or both. They are listed in the Appendix for transparency, not as a ranking. Only these 12 universities were scored and profiled, but the examples shown in Section 3 were not scored—they only serve to illustrate the actions. Evidence was collected and peer-validated between June and August 2026. The sample is not representative and familiarity could favour authors' own institutions; in practice these did not score higher than comparators (average Stage 2 total scores of 5.8 and 6.1 out of 14, respectively).

Each score was linked to a public source and checked by a second team member, with differences resolved through discussion.

Interpretation limit. Limited evidence may reflect weak public visibility rather than limited capability. Figures such as "7 of 12" or "1 of 12" relate to our sample of 12 universities; they are not estimates for universities. Partnership or investment decisions should be confirmed through direct engagement with the universities concerned.

AI use disclosure. Figures 1 to 3 were produced with Claude (Anthropic). The authors defined the content, structure and data, reviewed and edited all outputs for accuracy, and retain responsibility for the final material.[8]

2.1 AI and SDG capability is widespread

Of the 12 universities mapped, seven showed visible activity in both AI and the SDGs ("Full-Spectrum Catalysts"), three leaned towards technical AI ("Technology-Driven Innovators"), one towards SDG activity ("SDG-Driven Adopter"), and one had less visible evidence overall ("Less Publicly Visible") (Figure 1). These profiles describe where public activity is concentrated and do not indicate institutional maturity or ranking.

2.2 Only one of 12 showed strong public evidence of a delivery pathway

Only one university showed strong public evidence of a pathway into implementation, while six showed emerging evidence, and five had limited evidence (Figure 2). The strong result came from the sole SDG-Driven Adopter. Of the seven Full-Spectrum Catalysts, four showed emerging and three limited evidence. In this sample, the breadth of visible AI and SDG activity was therefore not a reliable proxy for strong pathway evidence.

Central interpretation.

The mapping points to a gap in the mechanisms that connect university capability to adoption and sustained delivery, or in their public visibility.[9]

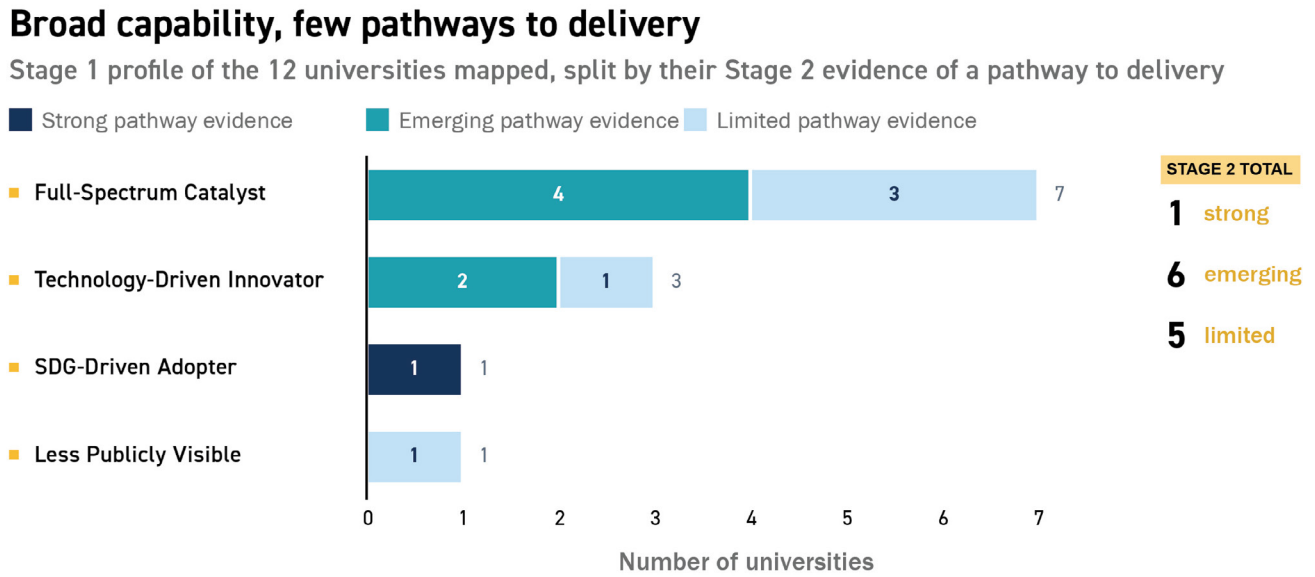


Figure 2. *Stage 1 activity profiles and Stage 2 evidence of implementation pathways across the 12 universities mapped. Categories are based on accessible public evidence and broad validated profiles. Source: Authors' mapping of publicly available evidence.*

Key finding. 7 of 12 showed broad AI and SDG activity, but only 1 of 12 showed a strong pathway to delivery.

Limited public evidence does not establish that delivery activity is absent. None of the seven Full-Spectrum Catalysts showed strong pathway evidence; the one strong result came from an SDG-Driven Adopter.

2.3 Three gaps in the pathway

Figure 3 below counts universities with visible or partial public evidence for each implementation function. Further activity may exist that was not captured and the numbers do not reflect effectiveness or a complete delivery pathway. The counts also differ from the Stage 2 categories in Figure 2.

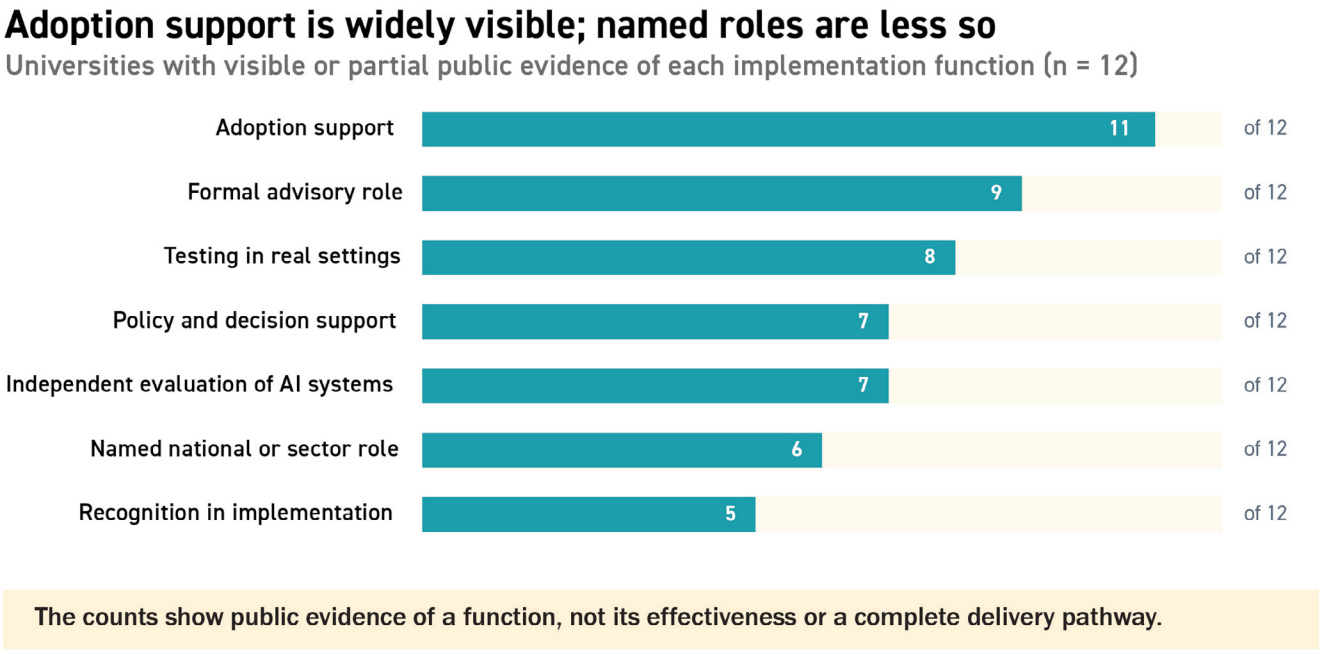


Figure 3. *Implementation functions visible across the 12 universities mapped. Source: Authors' mapping of publicly available evidence.*

The seven functions cluster into three weak links in the implementation pathway (Table 1). The last column shows which of the four actions in Section 3 responds to each. These are evidence gaps in the public record, not proof that a university lacks the function.

The weakest visible link is recording who contributed to implementation (5 of 12). Independent evaluation was visible in seven universities, but the mapping does not establish whether it was built into a pilot early enough to guide adoption.[3] These patterns warrant direct checks with universities and testing in other national systems.

Gap	Stage 2 Functions (universities with evidence, of 12)	What the Mapping Shows	Possible Response
1. Role - No defined role in delivery	Named national or sector role (6); policy and decision support (7); formal advisory role (9)	Half showed no public evidence of a named role where public decisions and delivery happen	Action 1: make university capabilities visible (Section 3.1) Action 2: build challenge-based collaboration (Section 3.2)
2. Route - Testing to adoption route is unclear	Adoption support (11); testing in real settings (8)	Adoption support is widely visible, while testing in real settings is less so. The mapping cannot trace a complete route.	Action 3: set up a national AI-for-SDG platform (Section 3.3)
3. Results - Evaluation and attribution are less visible	Independent evaluation of AI systems (7); recognition in implementation (5)	Seven show evaluation evidence; only five are named in accounts of what was delivered.	Action 3: build independent evaluation into the platform (Section 3.3) Action 4: share methods and evidence across the region (Section 3.4)

Table 1. *Three gaps in the pathway, from the Stage 2 data. Numbers show universities with visible or partial evidence, of 12. Source: Authors' mapping of publicly available evidence (Figure 3).*

3 What now: Four actions to close the gaps between capability and delivery

The policy choice is whether to leave university AI-for-SDG capability as reported activity, or to give universities defined roles in testing, independent evaluation and public delivery. Reporting remains useful; it becomes more valuable when it informs who does the work and how the results will be used.

The mapping identifies publicly visible strengths and weak links. Four actions are proposed to address them: make university capabilities visible (Section 3.1); build challenge-based collaboration (Section 3.2), set up a national AI-for-SDG platform (Section 3.3); and share methods and evidence across the region (Section 3.4). The gaps come from the 12-university mapping while the actions are proposals, illustrated by examples rather than tested by the mapping. SDG 17 underpins the partnerships in Actions 2 and 4, and SDG 4 underpins practitioner training in Action 3 and Box 7.

How to use the diagnostic. Identify which implementation functions have limited public evidence, then check those gaps with universities and public agencies. Choose the actions that address the confirmed gaps, agree on first-year signals, and repeat the diagnostic to see what has changed. The actions are options, not a required sequence.

Minimum conditions. Every AI-for-SDG pilot arising from these actions needs a named public owner, independent evaluation planned from the start, and criteria for adopting, adapting or ending it.

Learning from national and institutional cases. The country examples serve different purposes. Malaysia shows a possible role that is not yet filled; Hungary and the United Kingdom show different ways of connecting university capability to public problems and policy needs; Italy, Lithuania and Greece show infrastructure that could support public challenges; a second example from the United Kingdom shows what independent evaluation can reveal; and Brunei shows how a regional



programme can be delivered locally. They illustrate parts of the pathway, not proof of the complete model.

Two boxes involve the authors' own institutions: Ludovika University of Public Service in Box 2 and Universiti Teknologi Brunei in Box 7. The others are external examples, selected by the authors from public sources to illustrate a particular gap or action, not through a systematic scan of AI-for-SDG activity. They show what is possible, not a representative picture of country practice.

3.1 Make university capabilities visible

Why this action. Half the universities showed no public evidence of a named role in delivery (Gap 1). One likely reason is that relevant expertise cannot be commissioned if agencies cannot find it: our own evidence search had to piece together institutional pages, project records and local-language sources, often relying on knowledge of each system (Box 1).

How it works. Universities keep a verified, searchable record of expertise, facilities, datasets, implementation functions and past results, indexed by sector and public challenge. The ministry hosts or links the records and helps keep claims current. Stage 1 profiles can help identify complementary partners, but are not a substitute for searchable expertise. Universiti Brunei Darussalam's Expert Directory and University and Industry Platform already offer a starting point for industry.[10]

University role. Provide and maintain accurate records, including what has not yet been tested.

Who leads. The higher education or research ministry, with university research offices.

First-year signal. A first directory within six months, used by at least one agency to find a partner.

Box 1 A national AI system with room for a university role

Malaysia: SiPKPM early-warning system (SDG 4)

External example

The Ministry of Education and the United Nations Children's Fund (UNICEF) co-developed SiPKPM, an AI system that identifies students at risk of dropping out. After a six-month pilot in which 670 principals, teachers, counsellors and officials were consulted and trained, it now operates in more than 10,000 schools and helped more than 9,000 identified students return in 2025. The published account names no specific university partners involved in the pilot. Research on local dropout factors, independent evaluation of the system and practitioner training are roles universities could take on. A searchable record of university expertise would make those roles easier to fill.[11]

3.2 Build challenge-based collaboration

Why this action. Advice is common, delivery roles are not. Nine universities showed a formal advisory role but only six named a role in delivery (Gap 1). Visibility helps an agency find potential partners, while a challenge agreement turns their expertise into work the agency can use. It must specify who defines the problem, tests the system and evaluates the result.[5,9]

How it works. The agency that owns a public problem forms a consortium around it and gives each partner a named task: problem definition, data governance, testing, evaluation or maintenance. Roles, intellectual property and the route to adoption are agreed in writing at the start. The Stage 1 profiles guide who does what: a Technology-Driven Innovator may develop the AI, while an SDG-Driven Adopter brings sector and community knowledge (Figure 1). Few teams combine advanced AI with deep sustainability expertise on their own.[12]

University role. A named task with defined deliverables, not a general advisory role (Boxes 2 and 3). Smaller universities may be vital where community trust matters.

Who leads. The agency that owns the problem, with one university coordinating.

First-year signal. A signed agreement within six to nine months naming each partner's role, a public owner, an evaluation plan and the adoption route.

Box 2 A university consortium organised around a national challenge

Hungary: National Laboratory for Water Science and Water Security (SDG 6)

Involves an author's institution

Led by the University of Pannonia, the laboratory brings together six universities, including Ludovika University of Public Service, one of the authors' institutions, with research institutes and the General Directorate of Water Management. Its HUF 8 billion programme includes AI-based water monitoring, forecasting and early-warning systems, responding to problems such as Lake Balaton's 2019 algal bloom, which existing monitoring had not foreseen. Not every component is AI-based. It shows the structure Action 2 calls for; what public sources do not yet show is how the water authority will use the results, or measures such as warning lead time that would show whether the work improves public decisions.[13]

Box 3 Connecting university evidence capability directly to policy

United Kingdom: Mobilising Evidence through Artificial Intelligence and User-Informed Synthesis (METIUS) (multiple SDGs)

External example

METIUS is a GBP 11.5 million UKRI-funded infrastructure led by Queen's University Belfast with university and evidence partners. It combines AI with human expertise to synthesise scientific evidence more quickly and make it accessible to policymakers, with exemplar work spanning education, crime and justice, climate and environment, and international development. Its design includes close collaboration with UK and international policymakers and links to UN SDG evidence networks. It illustrates Action 2 because the collaboration is organised around a defined policy function rather than a general research partnership. What is not yet publicly established is how often the evidence changes specific policy or implementation decisions, or how those effects will be evaluated.[14]

3.3 Set up a national AI-for-SDG platform

Why this action. Testing in real settings was visible at 8 of the 12 universities, against 11 for adoption support (Gap 2), and independent evaluation at 7 (Gap 3). The mapping cannot show whether the same projects move from testing into routine use. A shared national mechanism is one proposed way to connect secure testing, independent evaluation and a public decision across several challenges.[3,4]

How it works. A named ministry uses or adapts an existing innovation or digital mechanism around a priority challenge. It supplies secure data access, test environments including an AI sandbox where appropriate, legal and ethical review, and procurement advice.[15] The developer builds, a structurally independent team evaluates, and the public agency decides on that evidence whether to adopt, adapt or end the pilot. Expand only when public value is demonstrated (Boxes 4 to 6).

University role. Universities may test a system and train practitioners, or independently evaluate a system they did not build or sell. These roles must be kept separate within a project.

Who leads. A named ministry or agency, using existing innovation or digital budgets.

First-year signal. One challenge running within 12 months, with a public owner, an independent evaluation plan and criteria to adopt, adapt or end the pilot.

Box 4 A university-led network that reaches industry and outlasts its first funding

Italy: Future Artificial Intelligence Research (FAIR)

External example

Set up under Italy's recovery plan, FAIR is a hub-and-spoke partnership in which universities lead research spokes alongside national research bodies and private partners. Through cascade funding calls, 76 companies received a total of EUR 11.8 million to develop projects with the partnership's researchers. The recovery-funded programme closed in December 2025, and the FAIR Foundation has since started follow-on projects, including the Intelligent and Secure Cyber-Physical Systems for Healthcare, Industry and Society (SINTESI) project in Southern Italy. The test now is whether the network keeps serving public challenges, not only research.[16]

Box 5 Shared AI computing led or hosted by universities

Lithuania and Greece: LitAI and Pharos AI Factories

External example

LitAI, Lithuania's national AI Factory, is led by Vilnius University with Kaunas University of Technology, VILNIUS TECH, Vytautas Magnus University and state agencies including the State Data Agency and the Innovation Agency. The consortium has EUR 65 million in EU co-financing towards a EUR 130 million centre serving priority sectors such as cybersecurity, green energy, smart industry and digital health, but still depends on final agreements. Greece's Pharos, coordinated by the National Infrastructures for Research and Technology (GRNET) and hosted at the National Technical University of Athens' Lavrion park, is due to offer services in 2026. Neither is an SDG programme as such; both could deliver public benefit if SDG-linked pilots and usage targets are attached from the start.[17,18]

Box 6 A university as independent evaluator of a national AI programme

United Kingdom: National Health Service (NHS) AI Lab (SDG 3)

External example

The NHS AI Lab was launched in 2019. In March 2024, after GBP 170 million of investment, NHS England commissioned an interdisciplinary University of Edinburgh team to evaluate it independent of the developers. The findings were published in 2025. The evaluation found early evidence of cost savings and better health outcomes from some technologies, and showed what held adoption back: procurement, integration with existing systems and fit with service needs. Government now has evidence available on what to scale and what to fix.[19]

3.4 Share methods and evidence across the region

Why this action. Only five of the 12 universities were named in accounts of what was delivered (Gap 3): limited public recording of results is a finding of the mapping. The case for regional cooperation is an additional policy judgment: common methods and records let countries compare lessons from pilots that each system may run only rarely.[20,21]

How it works. Countries agree a shared protocol for evaluating AI-for-SDG pilots, apply it through existing national programmes, and record which pilots were adopted, adapted or ended, why, and who contributed. The Asia-Europe Foundation (ASEF) could convene the protocol and joint challenge calls,[20,22] while the Asia-Europe for AI (AE4AI) Network and regional programmes could carry methods and training across borders while respecting national law, data sovereignty, language and inclusion.[23]

University role. Act as independent evaluators applying the shared protocol at home, and train others to use it through regional networks (Box 7).

Who leads. A regional body such as ASEF, with each partner funding its own participation.

First-year signal. At least two countries apply the shared protocol to one pilot each.

Box 7 A regional programme delivered locally through a university

Brunei Darussalam: AI Ready ASEAN (SDG 4)

Involves an author's institution

AI Ready ASEAN, run by the ASEAN Foundation with support from Google.org, aims to equip 5.5 million people across the ten ASEAN Member States covered at its launch with AI skills. In Brunei, Universiti Teknologi Brunei, one of the authors' institutions, is one of two local implementing partners and aims, with the Big BWN Project, to reach 20,000 youth, educators and parents through workshops, training and "Hour of Code" sessions. It shows that regional programmes already reach communities through local universities. It also shows the Gap 3 problem: We found no published results as of September 2026. A shared protocol would let each country report reach and results in the same way, so the region could compare what works.[24]

Besides regional programmes such as AI Ready ASEAN (Box 7), existing networks offer a second route. The AE4AI Network, launched by ASEFInnoLab alumni in October 2023, now has 270 members across Asia and Europe, with working groups on AI governance, AI in education, AI for sustainable development, and AI for lifelong learning[25,26] Its outputs so far are mainly publications and events; linking it to national platforms would give it a policy function.[25] The authors are members.

The examples support distinct claims. METIUS shows infrastructure designed to connect AI-enabled evidence synthesis with policymaking; the United Kingdom NHS AI Lab example shows what an independent university evaluation can reveal after a national programme is underway. The other examples show arrangements or opportunities whose public outcomes still need to be measured. Planning evaluation at the start would make future results easier to judge.

Tracking progress. Alongside the first-year signals, track strategies naming a university role, pilots adopted or ended on evidence, and independent evaluations published.

4 Assign responsibility for turning university capability into public benefit

The first step for every actor is to name who will do the work, fund it and judge whether it benefits the public.

University leaders should appoint one contact or coordination unit that keeps an accessible record of expertise, facilities and results and responds to public agencies. They can use the two-stage framework to identify strengths and missing implementation functions, then invite external peers to review the evidence. Leaders must also protect the freedom to publish unfavourable evaluations, declare conflicts of interest, and decline work that compromises research integrity.

This role also needs sustained support, such as promotion criteria that recognise policy work, as in the Netherlands;[27] workable intellectual property agreements such as the UK's Lambert Toolkit;[28] and core funding for coordination and independent evaluation.

National governments should name the agency that owns each priority challenge, use the diagnostic to specify which university functions are missing, and fund the route to adoption and ongoing ownership beyond a single pilot.

Regional organisations and Asia–Europe networks should lead the shared evaluation protocol and record in Section 3.4, and share training and computing access.[20,21,23]

Industry partners should commit to the data access, maintenance and skills transfer needed for responsible use. Communities should help define the problem and judge whether the outcomes benefit them.[29] Public agencies remain accountable for decisions, with fairness, privacy, safety, transparency and independent evaluation built in from the start, and routes for explanation, review and redress.[3,30]

5 Conclusion: From AI-for-SDG capability to delivery

Universities across Asia and Europe bring different strengths to AI-for-SDG work. Some develop AI systems; others understand public needs, train practitioners, work with communities or independently assess what works. Our mapping shows why these strengths should be considered together: seven of the 12 universities had broad visible AI and SDG activity, yet only one showed strong public evidence of a pathway into implementation. The finding calls for closer engagement, not a judgment that the others lack capability. The examples point the same way: where universities hold a defined step, research capability can be connected directly to policy needs, as in METIUS, or used to independently evaluate a national programme, as in the United Kingdom NHS AI Lab.

The route forward is practical, and countries need not do everything at once: the diagnostic shows which gaps to close. Governments can name a public challenge and its owner; universities can make their strengths visible and take defined roles; and partners can agree from the start how a solution will be tested, evaluated and adopted. Regional networks can help share methods and learning. If these connections are made, university capability can move beyond promising activity to AI-for-SDG work that is tested, trusted and delivers public benefit.

Acknowledgements

What began as a project among people who had never met became a steady connection across countries and cultures. Over almost six months, the team continued to exchange ideas and contribute through busy schedules and holiday breaks, with generosity, care and professionalism. Meeting face to face in Nottingham was a highlight. By then, working together had built trust that made the conversation feel natural.

"All for one, and all for the SDGs" captures the spirit of this collaboration: different perspectives brought together for one shared purpose. As team leader, Sharina thanks every team member for their commitment. The team also thanks ASEF and the University of Nottingham for bringing them together and making this paper possible.

Nimisha Parashar acknowledges support from the Engineering and Physical Sciences Research Council (EPSRC) [Grant Number EP/S023305/1].

Authors

Dr Sharina YUNUS

Universiti Teknologi Brunei, Brunei Darussalam

Mr Francesco AMIGHETTI

THINK.TANK Plus, Greece

Dr Emese BELENYESI

Ludovika University of Public Service, Hungary

Ms Nimisha PARASHAR

University of Nottingham, United Kingdom

Mr Muhammad SOHAIL

Mykolas Romeris University, Lithuania

Dr Nur Lisa ZAHARIN

Universiti Malaya, Malaysia

Sharina Yunus served as team leader and paper lead. All authors contributed equally to the development of this white paper.

References

1. United Nations. (2026). *The Sustainable Development Goals report 2026*. United Nations Department of Economic and Social Affairs. https://unstats.un.org/sdgs/report/2026/

2. Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, *11*(1), Article 233. https://doi.org/10.1038/s41467-019-14108-y

3. OECD. (2024). *Governing with artificial intelligence: Are governments ready?* (OECD Artificial Intelligence Papers, No. 20). OECD Publishing. https://doi.org/10.1787/26324bc2-en

4. Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., A'Court, C., Hinder, S., Fahy, N., Procter, R., & Shaw, S. (2017). Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research*, *19*(11), Article e367. https://doi.org/10.2196/jmir.8775

5. Trencher, G., Yarime, M., McCormick, K. B., Doll, C. N. H., & Kraines, S. B. (2014). Beyond the third mission: Exploring the emerging university function of co-creation for sustainability. *Science and Public Policy*, *41*(2), 151–179. https://doi.org/10.1093/scipol/sct044

6. United Nations Department of Economic and Social Affairs. (n.d.). *Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*. https://sdgs.un.org/goals/goal4

   Ferk Savec, V., & Jedrinović, S. (2025). The role of AI implementation in higher education in achieving the Sustainable Development Goals: A case study from Slovenia. *Sustainability*, *17*(1), Article 183. https://doi.org/10.3390/su17010183

7. United Nations Department of Economic and Social Affairs. (n.d.). *Goal 17: Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development*. https://sdgs.un.org/goals/goal17

8. Anthropic. (2026). *Claude* [Large language model]. https://claude.ai

9. Trencher, G., Nagao, M., Chen, C., Ichiki, K., Sadayoshi, T., Kinai, M., Kamitani, M., Nakamura, S., Yamauchi, A., & Yarime, M. (2017). Implementing sustainability co-creation between universities and society: A typology-based understanding. *Sustainability*, *9*(4), Article 594. https://doi.org/10.3390/su9040594

10. Universiti Brunei Darussalam. (n.d.). *Driving innovation through sustainability research*. UBD Sustainability. https://sustainability.ubd.edu.bn/sustainability-research/

11. UNICEF. (2026, June 1). *Future ready and resilient: How AI is keeping Malaysian students in school*. https://www.unicef.org/digitaleducation/stories/future-ready-and-resilient-how-ai-keeping-malaysian-students-school

12. Gohr, C., Rodríguez, G., Belomestnykh, S., Berg-Moelleken, D., Chauhan, N., Engler, J.-O., Heydebreck, L. V., Hintz, M. J., Kretschmer, M., Krügermeier, C., Meinberg, J., Rau, A.-L., Schwenck, C., Aoulkadi, I., Poll, S., Frank, E., Creutzig, F., Lemke, O., Maushart, M., . . . von Wehrden, H. (2025). Artificial intelligence in sustainable development research. *Nature Sustainability*, *8*(8), 970–978. https://doi.org/10.1038/s41893-025-01598-6

13. National Research, Development and Innovation Office. (n.d.). *Víztudományi és Vízbiztonsági Nemzeti Laboratórium* [National Laboratory for Water Science and Water Security]. https://www.nemzetilaborok.nkfih.gov.hu/viztudomanyi-es-vizbiztonsagi-nemzeti-laboratorium/laboratorium-bemutatasa

    University of Pannonia. (n.d.). *Víztudományi és Vízbiztonsági Nemzeti Labor (RRF-2.3.1-21-2022-00008)* [National Laboratory for Water Science and Water Security]. https://fpi.uni-pannon.hu/projektek/rrf/140-viztudomanyi-es-vizbiztonsagi-nemzeti-labor

    University of Pannonia. (n.d.). *National Laboratory for Water Science and Water Security (RRF-2.3.1-21-2022-00008)*. https://pen.uni-pannon.hu/en/national-laboratory-for-water-science-and-water-security-rrf-2-3-1-21-2022-00008/

    National Research, Development and Innovation Office. (n.d.). *National Laboratory for Water Science and Water Security* [Programme overview]. https://www.nemzetilaborok.nkfih.gov.hu/green-transition/files-to-download-240826/national-laboratory-for

14. UK Research and Innovation. (2025, September 22). *AI investment to transform global policy with scientific evidence.* https://www.ukri.org/news/ai-investment-to-transform-global-policy-with-scientific-evidence/

    UK Research and Innovation. (2024, September 19). *Transforming global evidence: AI-driven evidence synthesis for policymaking.* https://www.ukri.org/opportunity/transforming-global-evidence-ai-driven-evidence-synthesis-for-policymaking/

15. OECD. (2023). *Regulatory sandboxes in artificial intelligence* (OECD Digital Economy Papers, No. 356). OECD Publishing. https://doi.org/10.1787/8f80a0e6-en

    European Union. (2024). *Regulation (EU) 2024/1689 (Artificial Intelligence Act), Article 57: AI regulatory sandboxes*. Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

16. Fondazione FAIR. (n.d.). *FAIR – Future Artificial Intelligence Research*. https://fondazione-fair.it/en/

    Fondazione FAIR. (2026, September 25). *SINTESI kicks off: Building a stable, distributed AI innovation hub in Southern Italy*. https://fondazione-fair.it/en/news/sintesi-kicks-off-building-a-stable-distributed-ai-innovation-hub-in-southern-italy/

    Fondazione FAIR. (2025, December). *Rome hosts the closing General Conference of the PNRR-funded FAIR project: Scientists and companies from across Italy come together to present three years of national research on artificial intelligence*. https://fondazione-fair.it/en/news/rome-hosts-the-closing-general-conference-of-the-pnrr-funded-fair-project-scientists-and-companies-from-across-italy-come-together-to-present-three-years-of-national-research-on-artificial-intelligen/

References

17. EuroHPC Joint Undertaking. (n.d.). *Lithuania*. https://www.eurohpc-ju.europa.eu/lithuania_en

    Ministry of the Economy and Innovation of the Republic of Lithuania. (2026, June 22). *Skelbiamas kvietimas dirbtinio intelekto gamyklai „LitAI“ steigti* [Call opened to establish the LitAI artificial intelligence factory]. https://eimin.lrv.lt/lt/ziniasklaidai/naujienos/skelbiamas-kvietimas-dirbtinio-intelekto-gamyklai-litai-steigti-EayI/

    Lithuania.lt. (2025). *Lithuania launches €130 million national AI centre LitAI*. https://lithuania.lt/governance-in-lithuania/lithuania-launches-e130-million-national-ai-centre-litai/

18. Pharos AI Factory. (n.d.). *Pharos AI Factory*. https://www.pharos-aifactory.eu/

    GRNET. (2025, March 26). *The implementation of the new European supercomputer “DAEDALUS” and the data centre at the Lavrion Technological and Cultural Park (LTCP) begins*. https://grnet.gr/en/2025/03/26/daedalus-dc-ylopoihsh-lavrio/

19. University of Edinburgh Usher Institute. (2025). *Independent evaluation highlights impact and future potential of NHS AI Lab*. https://usher.ed.ac.uk/news-events/news/independent-evaluation-highlights-impact-and-future-potential-of-nhs-ai-lab

    Digital Health. (2025, April 29). *Evaluation of NHS AI Lab calls for national support for AI initiatives*. https://www.digitalhealth.net/2025/04/evaluation-of-nhs-ai-lab-calls-for-national-support-for-ai-initiatives/

20. Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2024). Global AI governance: Barriers and pathways forward. *International Affairs*, *100*(3), 1275–1286. https://doi.org/10.1093/ia/iiae073

21. Xu, J., Lee, T., & Goggin, G. (2024). AI governance in Asia: Policies, praxis and approaches. *Communication Research and Practice*, *10*(3), 275–287. https://doi.org/10.1080/22041451.2024.2391204

22. Asia-Europe Foundation. (n.d.). *About us: History of the Asia-Europe Foundation*. https://asef.org/about-us/our-mission/

23. Zhu, Y., Zhu, Z., & Xu, W. (2025). Cross-border higher education cooperation under the dual context of artificial intelligence and geopolitics: Opportunities, challenges, and pathways. *Frontiers in Education*, *10*, Article 1656518. https://doi.org/10.3389/feduc.2025.1656518

24. Universiti Teknologi Brunei. (2025, June 5). *Brunei launches AI Ready ASEAN programme to empower Brunei citizens with essential AI skills*. https://www.utb.edu.bn/news/brunei-launches-ai-ready-asean-programme-to-empower-brunei-citizens-with-essential-ai-skills/

    Borneo Bulletin. (2025, June 6). *‘AI Ready ASEAN’ to equip 20,000 citizens with digital skills*. https://borneobulletin.com.bn/ai-ready-asean-to-equip-20000-citizens-with-digital-skills/

25. Asia-Europe Foundation. (2023). *ASEFInnoLab4 launches the Asia-Europe for Artificial Intelligence (AE4AI) Network*. https://asef.org/?p=24012

    Asia-Europe Foundation. (2024). *Asia-Europe for Artificial Intelligence (AE4AI) Network*. https://asef.org/projects/asia-europe-for-artificial-intelligence-ae4ai-network/

26. AE4AI Network. (n.d.). *About us*. https://www.asiaeurope4ai.org/about-us

    AE4AI Network. (n.d.). *Members*. https://www.asiaeurope4ai.org/members

    AE4AI Network. (n.d.). *Home*. https://www.asiaeurope4ai.org/

27. VSNU, NFU, KNAW, NWO, & ZonMw. (2019). *Room for everyone’s talent: Towards a new balance in the recognition and rewards of academics*. https://www.nwo.nl/sites/nwo/files/media-files/2019-Recognition-Rewards-Position-Paper_EN.pdf

28. UK Intellectual Property Office. (2018). *University and business collaboration agreements: Lambert Toolkit*. GOV.UK. https://www.gov.uk/lambert-toolkit

29. Fobbe, L. (2020). Analysing organisational collaboration practices for sustainability. *Sustainability*, *12*(6), Article 2466. https://doi.org/10.3390/su12062466

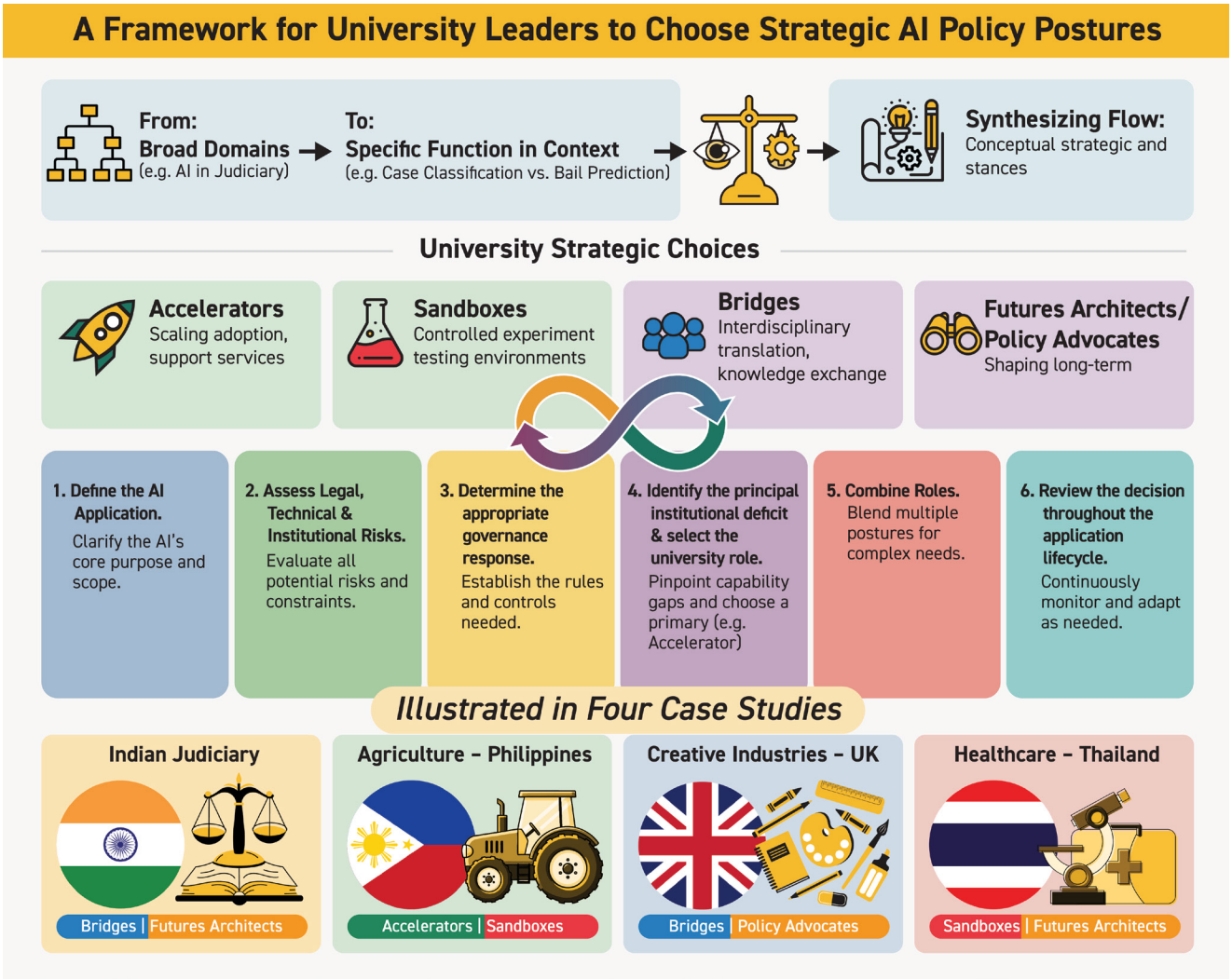
30. OECD. (2024). *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

    ASEAN. (2024). *ASEAN guide on AI governance and ethics*. ASEAN Secretariat. https://asean.org/book/asean-guide-on-ai-governance-and-ethics/

A Framework for University Leaders to Choose Strategic AI Policy Postures

A Sandbox, An Accelerator, A Bridge and A Futures Architect / Policy Advocate

AUTHORS
Spriha Bhandari
Tejpavan Gandhok
Zeynep Yaradanakul
Marion Lux Castro
Pavlos Panagiotidis
Antonio Pelegrina Jimenez
Rustam Shariq Mujtaba



Executive summary

This paper is aimed at senior leaders of leading universities and key artificial intelligence (AI) policymakers or influencers considering how universities can contribute strategically to the responsible adoption and governance of AI by external stakeholders.

Through our observations of the dominant practitioner and academic discourse on AI governance and of how universities are engaging with AI, we identified a recurring gap: existing approaches do not help guide universities on what strategic role they should play in relation to a specific AI application.

Domain-level classifications of AI systems are often too broad to determine what institutional role a university should play in a particular AI application. For example, "AI in law" or even the subset of "AI in judiciary" is too broad to be useful for the abovementioned purpose. AI systems used for judicial case classification, case allocation, judgment drafting, or bail prediction perform distinct functions and have different implications for fairness, accountability and rights.

Therefore, we propose a pragmatic solution to bridge this gap: a framework in which universities can choose four complementary roles:

- **Accelerators** by supporting adoption and scaling;
- **Sandboxes** by facilitating controlled experimentation;
- **Bridges** by enabling interdisciplinary translation and understanding; and
- **Futures Architects & Policy Advocates** by shaping long-term institutional and societal responses.

This framework is not prescriptive or rigid; it needs to be used with agility and dynamism given the rapid and complex nature of AI adoption. A single AI system may require multiple governance approaches at different stages of its usage lifecycle.

We illustrate this framework through four appended case studies: the Indian Judiciary, Creative Industries in the United Kingdom (UK), Healthcare in Thailand, and Agriculture in the Philippines.

Our framework can be applied at two levels:

- **To work with external stakeholders on a specific AI Initiative.** The framework proposes six steps to determine which role(s) would suit the university best when engaging with external stakeholders to support them in their AI initiative deployment. It can help the initiative's proposers think through key risks vs rewards, consequences, choice of key ecosystem partners, evolution pathway etc.; and thus, the more appropriate strategic stance(s) to support the given initiative over time.
- **To shape the university's portfolio of such AI initiatives.** Senior leaders can use it to map where different strategic AI initiatives stand; to iterate and shape the overall strategy for where the university should focus its resources and capabilities; as a communication tool; and most importantly, to review and reflect if important portfolio gaps exist.

We believe there is significant opportunity for leading national and global universities to play more proactive roles such as *Bridges and Futures Architects / Policy Advocates*, given the increasing debate and concerns about AI safety and governance.



1 Context

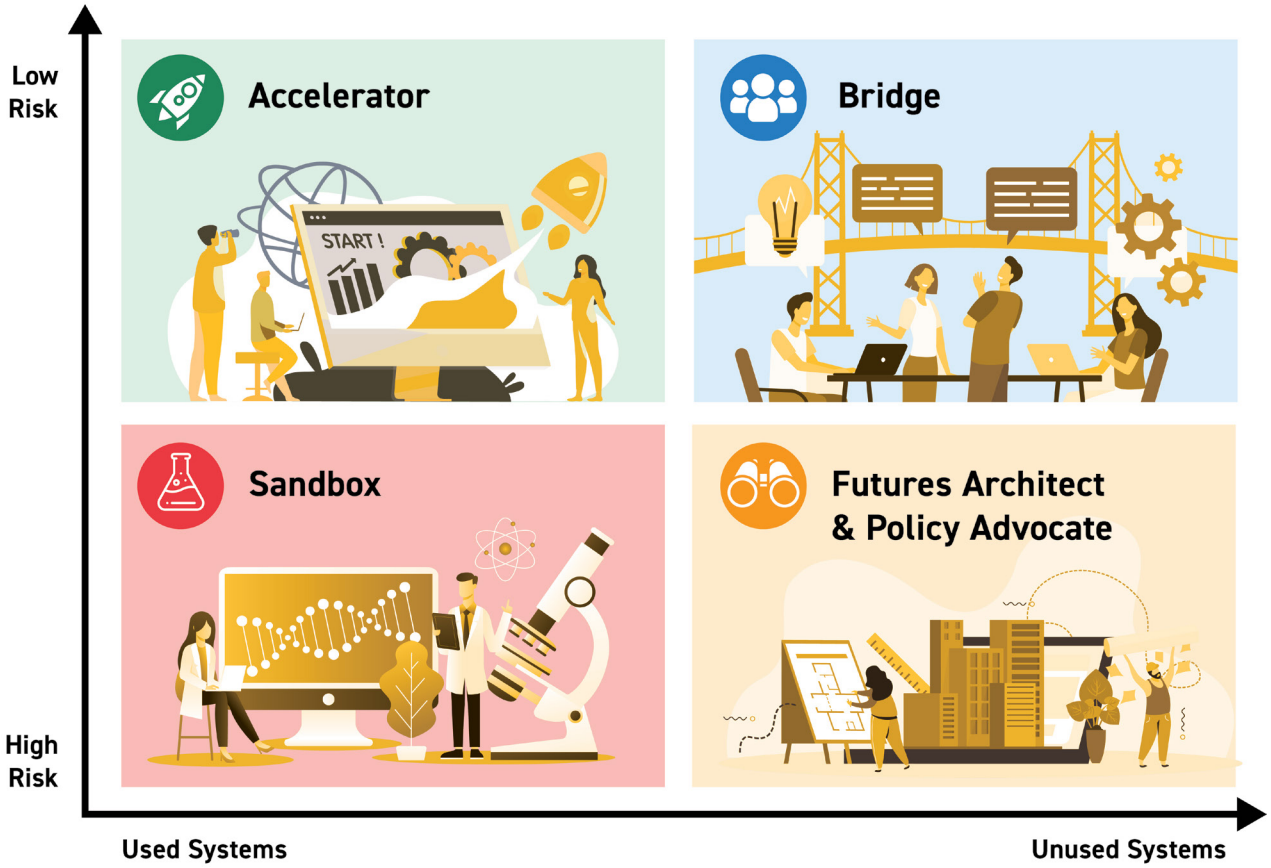
Artificial intelligence (AI) is increasingly embedded in public services, professional practice, business operations and critical infrastructure. As AI systems become integrated into decision-making processes, governance can no longer focus solely on technical performance. Effective governance must address how systems are designed, tested, procured, deployed, monitored and, where necessary, withdrawn.

This paper argues that universities can help fill this missing institutional layer. The proposition is not that universities should govern AI in place of governments, regulators, courts, industry or civil society; nor should university participation be interpreted as evidence that a particular AI application is safe, effective or socially beneficial. Universities have their own incentives, funding dependencies, intellectual property interests, and reputational considerations.

For the purposes of this framework, we therefore propose four conditions for credible university participation:

- Relevant expertise
- Operational independence
- Transparent methods
- Meaningful stakeholder participation

Under these conditions, universities can make four distinct but complementary contributions to AI governance:



- **Bridge**. Universities can translate between technical developers, public institutions, professional communities, and affected populations. They can help establish shared problem definitions, surface competing priorities and strengthen legitimacy in decision-making.
- **Sandbox**. Universities can generate independent, context-specific evidence through controlled testing and evaluation. This role helps decisionmakers assess whether systems perform as intended, for whom they work, and what risks they create.
- **Accelerator**. Universities can support responsible implementation when an application has demonstrated public value and sufficient evidence exists to justify deployment. Their role is to transfer capability, support integration, and strengthen institutional readiness, not simply to accelerate adoption.
- **Futures Architect & Policy Advocate**. Universities can help institutions anticipate longer-term changes resulting from emerging technologies. This includes assessing impacts on professional roles, public infrastructure, institutional authority, rights and future policy choices.

These roles are not mutually exclusive and may change across the lifecycle of an AI application. A university that helps design or test a system must remain willing to recommend re-design, restricted use, further evaluation, or non-deployment if the evidence supports those conclusions.

1.1 The missing institutional layer in AI governance

AI governance is often framed as a question of rules: which uses should be permitted, what obligations should apply, and who should be accountable when harms occur. These questions are necessary but not sufficient, they represent only part of the governance challenge.

Between the adoption of governance principles and the deployment of an AI system lies a series of operational decisions about problem definition, system design, procurement, evaluation, professional practice, oversight and remedy. These decisions determine whether principles such as fairness, accountability, transparency and safety become meaningful safeguards or remain aspirational.

Many organisations understand the principles of what responsible AI requires. Fewer possess the discipline to implement it consistently. Institutional challenges are not solely regulatory failures or inertia. Lack of adequate capacity and informed consensus are also at play. AI systems do not operate in isolation. They function within institutions, workflows and professional environments. Their impacts depend not only on technical performance but also on how they are implemented, governed and used. The same system may serve as a low-risk administrative tool in one context while significantly affecting access to healthcare, welfare benefits, employment or justice in another.[1] As a result, effective AI governance requires more than regulation alone. It requires institutions capable of translating governance principles into operational practice.

1.2 Why universities matter in this context

Universities occupy a distinctive position within the AI ecosystem. Unlike regulators, universities do not set legal rules. Unlike technology firms, they are not primarily responsible for commercial deployment. Unlike civil society organisations, they are not solely focused on advocacy. Their contribution lies in their ability to combine expertise, credible evidence generation, interpretation, stakeholder engagement, capability building, and long-term institutional thinking.

Research on innovation ecosystems and public interest governance suggests that universities are uniquely positioned to:

- Convene stakeholders who would not otherwise work together;[2]
- Generate independent evidence through testing and evaluation;[3]
- Support knowledge transfer across sectors;
- Build public sector and professional capability;
- Preserve institutional memory across political and commercial cycles; and
- Examine long-term societal impacts rather than immediate deployment outcomes.[4]

Universities should not be viewed as arbiters of whether AI systems should proceed. Their role is more limited, and often more valuable. Their involvement should not be assumed to favour deployment. Evidence may support adoption, but it may also support redesign, restriction or non-deployment. Their role should be to strengthen

society's ability to make informed, accountable and evidence-based choices about how AI is developed and used.

2 The University Role-Selection Framework

Foundational literature provides important starting points for understanding universities as actors in innovation and governance. Yet through our observations of the dominant practitioner and academic discourse on AI governance and our explorations of how universities are engaging with AI, we identified a recurring gap: existing approaches do not help guide universities on what strategic role they should play in relation to a specific AI application. Indeed, there appears a significant opportunity for leading universities to step more proactively into the increasing debate and concerns about AI safety and governance.

The framework we propose is our synthesis and is offered as a pragmatic solution to this gap we observed. It is inspired by similar roles universities have played in various other major impact technologies (e.g. gene editing, modern vaccines, nuclear physics, quantum computing, and even decades earlier foundations of AI and machine learning).

The four-role stances framework we propose translates these insights, together with our practical experience and observations, into a relevant tool for university leaders in this context.

Universities should determine their role in AI governance using a problem-led rather than institution-led approach. The question is not where a university wishes to participate, but what institutional deficit prevents a sound governance decision.

The University Role-Selection Framework follows six steps:

1. Define the AI application
2. Assess legal, technical and institutional risks
3. Determine the appropriate governance response
4. Identify the principal institutional deficit and select a role
5. Combining roles
6. Review the decision throughout the application lifecycle

This sequence separates the role of the university from the decision about the technology itself. University participation should not automatically be presumed to result in deployment. In many cases, the most valuable contribution a university can make is providing credible evidence that a system should not proceed.

2.1 Step 1: Define the AI application

The first step is to define the AI application with sufficient context-specific clarity.

Broad descriptions such as "AI in courts", "AI in education", or "AI in agriculture" provide little guidance for governance decisions. Within a single sector, AI may support transcription, research, classification, translation, summarisation or decisions that directly affect rights and entitlements. These applications create materially different risks and governance requirements.

The relevant unit of assessment is therefore a specific AI application performing a defined function within an identifiable workflow and its potential consequences – intended or unintended, direct vs next order.

Institutions should assess:

- The problem the application seeks to solve
- Its intended function
- Where it enters the workflow
- Who operates it
- Who remains accountable for its use
- Whose interests may be affected
- What non-AI alternatives exist

This prevents a common failure mode in which organisations begin with an available technology and subsequently search for a problem to justify its use.

2.2 Step 2: Assess legal, technical and institutional risks

Applications should be evaluated across five dimensions:

- Public stakes and rights implications
- The role the system plays in decision-making
- The reversibility and contestability of its effects
- The quality of available evidence
- Institutional capacity to govern and oversee its use

The framework deliberately avoids numerical scoring systems. Composite scores can obscure serious legal, rights-based, or safety concerns. Strong performance in one area should not offset unacceptable risks in another.

Instead, findings are classified as **Red**, **Amber** or **Green**.

Red

A red finding identifies a controlling concern that prevents deployment in its current form.

Examples include:

- Unlawful uses
- Unacceptable rights impacts
- Lack of meaningful contestability
- Unjustifiable exposure to risk
- Insufficient governance capability

A red finding cannot be averaged away through stronger results elsewhere. Appropriate responses include redesign, restriction, pause, withdrawal or non-deployment. Universities may still contribute under these circumstances. They may act as a Bridge to explain the basis for restriction or as a Futures Architect to explore alternatives. They should not create a Sandbox solely to justify proceeding with a fundamentally unsuitable application.

Amber

An amber finding indicates that important questions remain unresolved but can potentially be answered through ethically and legally defensible testing. In these circumstances, the application may enter a controlled *Sandbox*.

Testing should focus on specific uncertainties, including:

- Whether the system performs as intended
- Whether professionals can appropriately interpret outputs
- Whether affected individuals can challenge decisions
- Whether the system improves on existing approaches

Sandboxes should include:

- Defined objectives
- Independent oversight
- Success and failure criteria
- Incident reporting
- Conditions for early termination

Green

A green finding does not constitute approval. It simply indicates that no controlling barrier has been identified at that stage and that the institution may proceed to the next question: *What is preventing a sound, legitimate and sustainable decision about this application?* This question serves as the bridge between application assessment and university role selection.

2.3 Step 3: Determine the governance response

Depending on the evidence, the appropriate response to an AI application may be:

- Adoption
- Controlled pilot
- Conditional deployment
- Enhanced oversight
- Restricted use
- Redesign
- Pause
- Withdrawal
- Non-deployment

In some cases, the most valuable contribution a university can make is providing credible evidence that a system should not proceed.

2.4 Step 4: Identify the institutional deficit and select a role

Our proposed framework is built on four recurring deficits.

▶ Translation Deficit → Bridge

Technical developers, policymakers, professionals and affected communities often lack a shared understanding of the problem being addressed, the risks involved, or the conditions for acceptable use. Without effective translation, discussions become fragmented. Technical performance measures may fail to capture legal, ethical or operational concerns, while community experience may remain invisible within formal evaluation processes.

The *Bridge* role of universities addresses this challenge through:

- Stakeholder engagement
- Participatory problem definition
- Translation across technical, legal and operational domains
- Accountability design

Evidence Deficit → Sandbox

Decisionmakers often lack independent, context-specific evidence regarding how an AI system performs in practice. Vendor claims, benchmark results, and laboratory testing rarely provide sufficient information about real-world effectiveness, institutional impact, or distributional consequences.

The *Sandbox* role of universities addresses this challenge through:

- Controlled testing
- Comparative evaluation
- Context-specific evidence generation
- Assessment of technical, legal and social impacts

The objective is not validation, but reduction of uncertainty sufficient to support a defensible decision.

Adoption Deficit → Accelerator

Even where an AI application demonstrates value and sufficient evidence exists, organisations may lack the capability to implement and govern it responsibly. Challenges frequently include procurement, workflow integration, staff training, maintenance, monitoring and vendor dependency.

The *Accelerator* role of universities addresses this challenge through:

- Capability transfer
- Workforce development
- Workflow redesign
- Procurement support
- Monitoring and governance arrangements

Acceleration does not mean rapid deployment. It means responsible implementation within defined legal, evidentiary and institutional limits.

Foresight Deficit → Futures Architect & Policy Advocate

Some AI applications do more than improve existing processes. They reshape institutions themselves. They may alter professional responsibilities, change how authority is exercised, create new infrastructure dependencies, or narrow future policy choices in ways that are not immediately visible. Addressing these impacts requires foresight rather than implementation planning alone.

The *Futures Architect* role of universities addresses this challenge through:

- Strategic foresight
- Technology assessment
- Participatory scenario development
- Institutional design
- Long-term policy and infrastructure planning

The focus is on preserving future options and improving institutional preparedness

2.5 Step 5: Combining roles

The four roles are complementary but should not be combined automatically. The framework requires institutions to identify the principal deficit and designate a lead role. Additional roles should be introduced only where they are necessary to ensure legitimacy, reliability or sustainability.

Examples include:

- **Bridge + Sandbox**: Stakeholders help shape evaluation criteria and determine what should be tested.
- **Sandbox + Accelerator:** Independent testing provides sufficient evidence to support implementation.
- **Bridge + Sandbox + Accelerator:** Complex public-sector deployments requiring participation, evidence generation, and capability transfer.
- **Bridge + Futures Architect & Policy Advocate:** Emerging technologies that require both public engagement and institutional redesign.
- **All Four Roles:** Mission-oriented programmes addressing complex public challenges over extended time horizons.

Permutation (i.e. the order of roles) is often as important as their combination.

Where the problem itself remains contested, Bridge activities should precede Sandbox activities. Where evidence remains uncertain, Sandbox activities should precede Accelerator activities. Where implementation may create long-term dependencies, Futures Architect activities should begin before significant deployment occurs.

2.6 Step 6: Reviewing choices across the lifecycle

Role selection is not a one-time decision. The appropriate university contribution may change as an AI initiative moves through:

- Problem definition
- Design
- Testing
- Deployment
- Scaling
- Monitoring
- Withdrawal

Bridge and Futures Architect functions are often most valuable during problem definition. Sandbox activities become more relevant during evaluation. Accelerator functions should only occur once legal, evidentiary and public value thresholds have been met.

Monitoring may reactivate Sandbox activities where performance changes or new evidence emerges. Bridge activities remain important throughout the lifecycle to surface user and community experiences.

Reassessment is particularly important when:

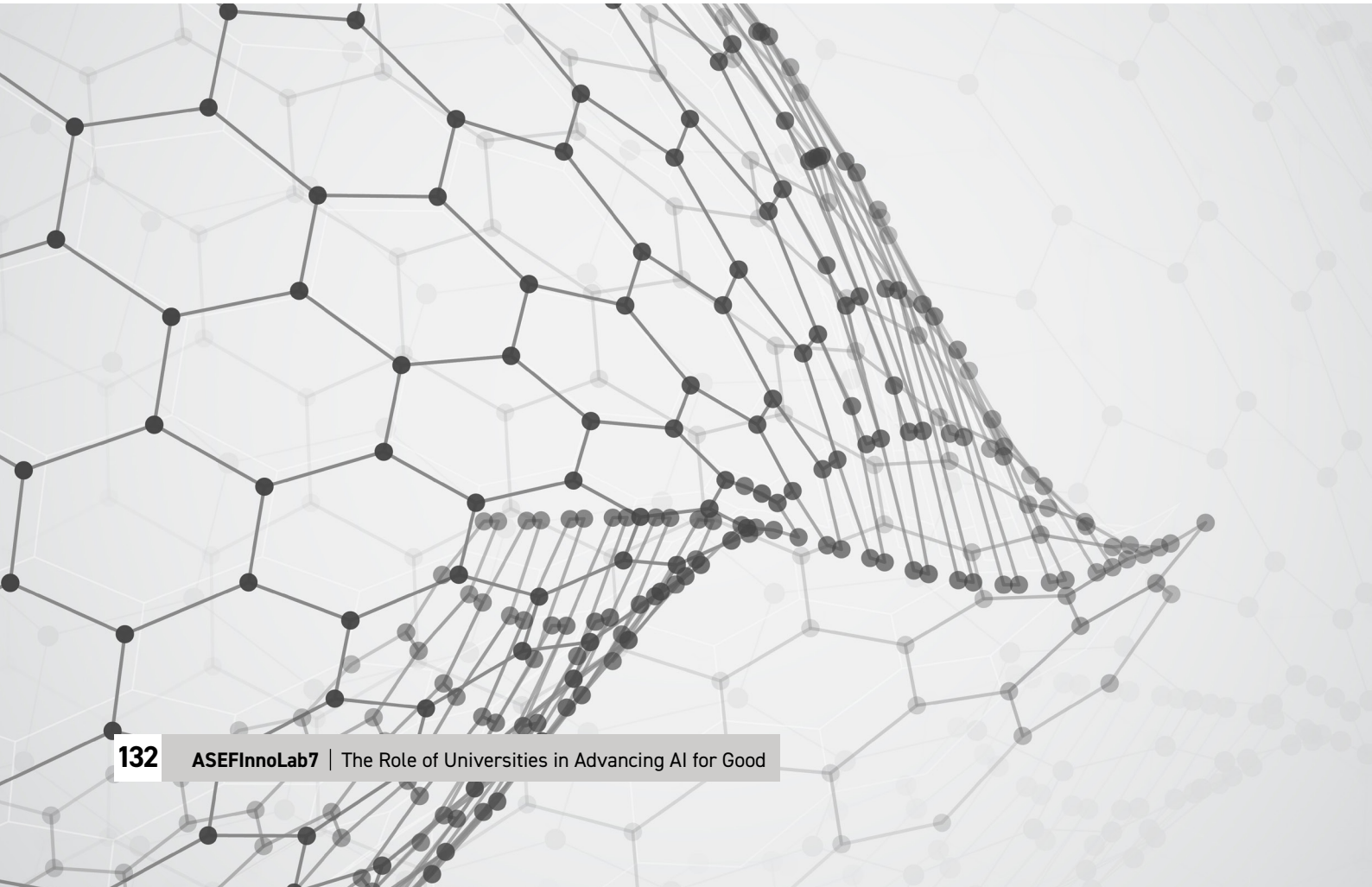
- A serious incident occurs
- The model or underlying data changes materially
- The system expands to a new population
- Evidence of unequal performance emerges
- Human oversight weakens
- Vendor arrangements change
- The system's role evolves

Governance should respond to how a system functions in practice, not simply how it was described at procurement. Withdrawal should also be treated as a legitimate governance outcome rather than a failure of innovation.

Before scaling an application, institutions should understand:

- How it can be paused.
- How past decisions can be reviewed.
- How records can be corrected.
- What alternatives remain available?
- How remedies will be provided.

A university acting as governance infrastructure must be as capable of supporting responsible exit as it is of supporting implementation.

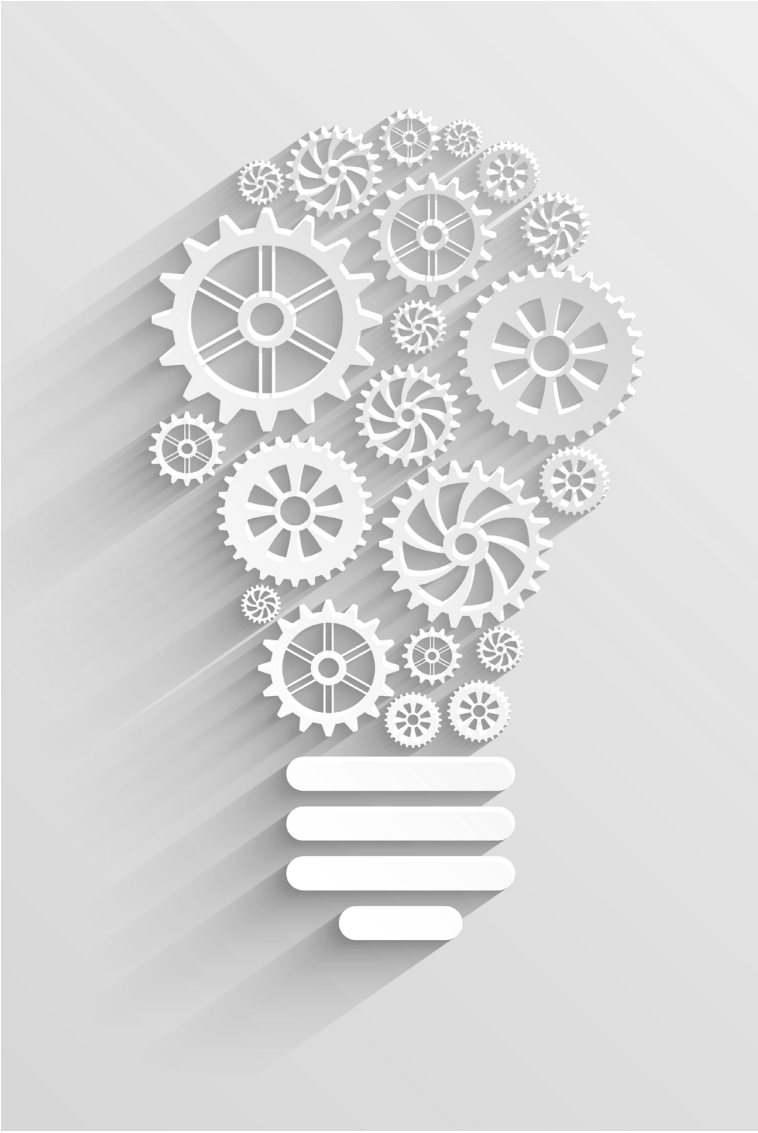


3 Conclusion

AI governance should start with a more specific vocabulary: the technology, use-case context, and its risk-reward consequences, including higher order and unintended effects. The framework proposed in this paper shifts the question from “How can universities support AI?” to “What deficit is preventing a sound governance decision?” Universities can help address the space between high-level governance principles and real-world implementation, but their role should respond to the specific needs of each application and context.

Universities create value not by endorsing innovation, but by improving the quality, legitimacy and accountability of decisions about innovation. In an increasingly AI-enabled society, that capability represents an important component of governance.

We encourage university leaders to use this framework to assess their AI initiatives and portfolios, identify where they can add the most value, and make deliberate decisions on how to engage. Leading global and national universities should have a good set of Accelerator and Sandbox projects and may aspire for more challenging initiatives in Bridge and Futures Architect quadrants. Others at more developing phases in their journey with a blank slate may wish to get started with partnerships for Accelerator or Sandbox projects.





Acknowledgements

The work of Pavlos Panagiotidis on this paper was supported by the Engineering and Physical Sciences Research Council [Grant Number EP/S023305/1].

Authors

Prof Spriha BHANDARI
O.P. Jindal Global University, India

Prof Tejpavan GANDHOK
O.P. Jindal Global University, India

Ms Zeynep YARADANAKUL
Erasmus Mundus Association, Finland

Dr Marion Lux CASTRO
University of the Philippines Los Baños, Philippines

Mr Pavlos PANAGIOTIDIS
University of Nottingham, United Kingdom

Mr Antonio PELEGRINA JIMENEZ
Universidad de Granada, Spain

Mr Rustam SHARIQ MUJTABA
Health Sciences Authority, Singapore

References

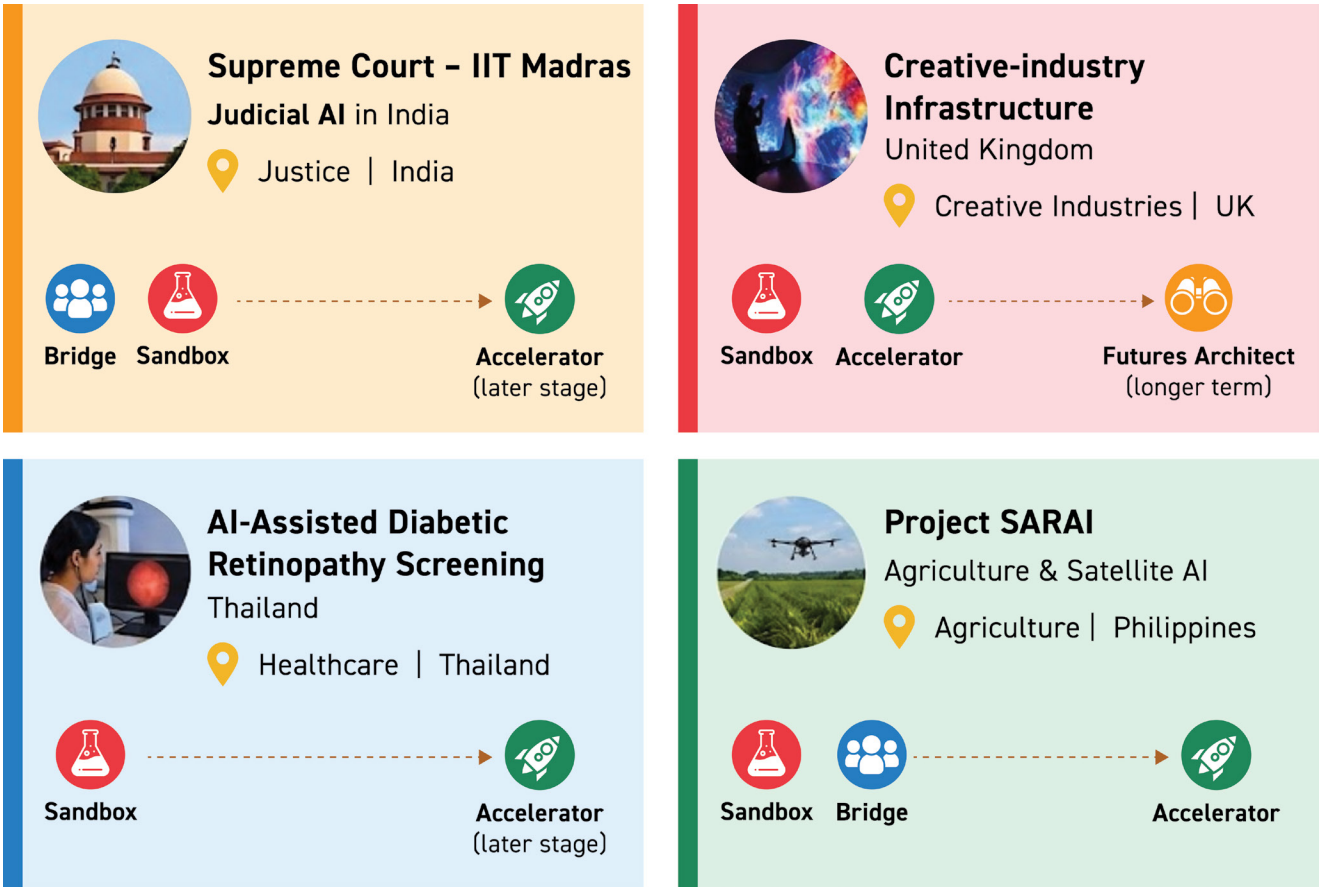
1. Veale, M., Van Kleek, M., & Binns, R. (2018). Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. *Proceedings of the 2018 CHI Conference on Human Factors in Computing* Systems - CHI '18. https://doi.org/10.1145/3173574.3174014
2. Guston, D. H. (2001). Boundary organizations in environmental policy and science: An introduction. *Science, Technology, & Human Values, 26(4),* 399–408. https://doi.org/10.1177/016224390102600401
3. Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research Policy, 29(2),* 109–123. https://doi.org/10.1016/S0048-7333(99)00055-4
4. Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy, 42(9),* 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Appendices

The Indian Judiciary; UK Creative Industries; Thai Healthcare; Philippine Agriculture

Case Studies: The Framework in Action

The appended four case studies illustrate that university roles should vary by application and context. They also show how roles can be combined and sequenced to support evidence-based AI governance across different sectors.



Appendices

1 SUVAS, SUPACE, and the Supreme Court – IIT Madras collaboration: A bounded pathway for AI use by the judiciary in India

Situation	Justice systems often face challenges relating to language accessibility, increasing caseloads, administrative complexity and delays in information processing. In a linguistically diverse country such as India, access to judgments and legal documents can be constrained by language barriers while judges, lawyers and court staff must navigate large volumes of case materials. AI technology use within courts raises significant concerns around accuracy, fairness, transparency, accountability and procedural justice.
Challenge	Judicial systems are high-stakes environments where errors can affect legal rights, access to justice, and public trust. AI-supported tools may produce inaccurate translations, incomplete summaries or incorrect retrieval of legal precedents. The challenge is not simply technical adoption but ensuring that AI assists court processes without displacing human judgment or undermining institutional accountability.
Response	India's emerging court-AI ecosystem has adopted a bounded and assistive approach to implementation. The **Supreme Court Vidhik Anuvaad Software (SUVAS)** supports the translation of judgments and orders into Indian languages, while the **Supreme Court Portal for Assistance in Court Efficiency (SUPACE)** was developed as a research support tool to help users understand case facts and identify potentially relevant precedents rather than to support judicial decision-making itself. In October 2023 the Supreme Court of India and the Indian Institute of Technology Madras (IIT Madras) entered into a collaboration covering AI and emerging technologies for transcription, transcript summarisation, translation, a streaming platform for court trials, process automation, legal-domain language models, capacity building, training, e-learning, and ICT skills. One of the clearest documented outputs of this collaboration is an AI/ML-based e-filing defect-identification tool developed by the Supreme Court Registry in coordination with IIT Madras. Prototype access was provided to 200 Advocates-on-Record for testing and feedback. The Supreme Court and IIT Madras have also tested prototypes for curing filing defects and extracting data and metadata, with integration into the electronic filing system and Integrated Case Management and Information System contemplated.
Impact	• An e-filing defect-identification prototype was made available to a limited group of 200 Advocates-on-Record for testing and feedback. • Collaboration between the Supreme Court and IIT Madras strengthens knowledge exchange between technical researchers and judicial practitioners
Takeaways	The Supreme Court – IIT Madras collaboration provides a strong example of the *Bridge* role because it connects technical expertise with judicial and registry workflows. The bounded testing of e-filing and related prototypes also illustrates the *Sandbox* role: systems are being exposed to real users and operational conditions before wider integration. The case also illustrates an important distinction within the Sandbox concept: a university that co-develops a system may support controlled experimentation, but co-development does not automatically constitute independent evaluation. Independent validation may require separate evaluators, institutional firewalls, or other arm's-length arrangements. Training, process integration, and wider roll-out to other courts and jurisdictions may eventually become an *Accelerator* role, but only where applications have been adequately validated. Public information remains limited regarding translation error rates, retrieval accuracy across different legal contexts, user experience, accessibility and mechanisms for addressing incorrect AI-assisted outputs. The transferable lesson is not that courts should rapidly automate legal processes, but that universities can help public institutions experiment responsibly, generate evidence and build capability while preserving human oversight. In a judicial context, defining clear limits on AI use, as well as sandboxing the development and testing of alternate AI systems that can speed up the massively clogged backlogs of Indian courts are examples of potential higher value yet higher risk use cases and policy stances for universities.

Appendices

2 Creative industry infrastructure in the UK: Accelerating capability

Situation	Much of the creative sector is composed of small and medium-sized enterprises (SMEs) and independent practitioners that lack access to AI infrastructure, technical expertise, and sustained research and development capacity.
Challenge	Creative sector organisations often face significant barriers to AI adoption, including limited access to advanced facilities, AI compute, technical support, and testing environments. At the same time, AI applications in creative fields raise broader concerns around authorship, intellectual property, creative labour, accessibility, cultural diversity, and equitable distribution of benefits. Without appropriate support structures, both innovation and its benefits risk becoming concentrated among larger organisations.
Response	The United Kingdom has developed a university-centred innovation ecosystem that connects frontier AI research with creative sector adoption. The Creative Industries Clusters Program established nine regional partnerships between universities and creative businesses, supported by GBP 56 million in public investment. The Convergent Screen Technologies and Performance in Realtime (CoSTAR) network provides GBP 75.6 million of shared infrastructure, including specialist laboratories, AI-enabled production environments, and a Foresight Lab examining future skills, intellectual property arrangements, and industry structures.
Impact	• Immersive Arts UK's first funding round attracted 2,517 applications and awarded GBP 1.18 million to 83 projects. • Comparable UK creative industry programmes have reported approximately £5 of co-investment for every GBP 1 of public funding, alongside thousands of jobs created or safeguarded and thousands of creative professionals trained. • Foresight and governance capabilities encourage consideration of long-term implications for skills, intellectual property, creative work, and responsible AI use.
Takeaways	The lead university role is *Accelerator*, because the principal deficit is unequal access to infrastructure and implementation capability. However, acceleration cannot operate credibly without two supporting roles. Universities act as a *Bridge* when they translate between technical researchers, creative professionals, cultural institutions, and policy actors. They could act as a *Sandbox* where shared facilities are used to test technologies within realistic production environments before wider adoption. The Foresight Lab could potentially also add a *Futures Architect* function by examining how emerging technologies may reshape creative work and the institutional conditions under which it is performed. This case therefore illustrates that potentially all four roles could meaningfully be played by universities. Yet it is not a general licence to accelerate every application developed within it. CoSTAR remains too recent for confident claims about its long-term effects. Investment secured, businesses engaged, products created, or jobs reported may indicate activity, but they do not independently establish cultural value, fair remuneration, equitable access, or responsible use of AI. The appropriate governance response is controlled experimentation followed by conditional implementation at the level of the application. Universities could separately examine who gains access to the infrastructure, who owns the resulting intellectual property, how creative workers are affected, and whether public investment produces capabilities that remain broadly accessible.

Appendices

3 Health education and adoption at scale: AI-assisted diabetic retinopathy screening in Thailand

Situation	Thailand's Ministry of Public Health set a national target of screening 60% of registered diabetes patients for diabetic retinopathy (DR) annually. With estimations of 4.5 million diabetes patients and 200 retinal specialists nationwide, manual reviews created week-long waits.
Challenge	An initial attempt driven only by the AI global tech Titan and a local lab and a human-centred field study across 11 clinics in Pathum Thani and Chiang Mai found the tool struggled in real-world scenarios: about 1 in 5 images were rejected due to low-quality data. Nurses could screen only around 10 patients per two-hour session, and a tool meant to speed up triage instead created new bottlenecks and frustrations.
Response	In a revised, subsequent initiative sometime later, a leading Thai university was roped in for a credible *Sandbox* role. GH researchers ran participatory design workshops with stakeholders to redesign the screening workflow around real clinic conditions rather than lab assumptions.
Impact	• Prospective, multisite interventional study across 13 health regions (7,940 patients screened), the deep-learning system achieved 94.7% accuracy, 91.4% sensitivity, and 95.4% specificity for vision-threatening DR – comparable to a panel of retinal specialists (93.5% accuracy; 84.8% sensitivity). • Refreshed programme has subsequently scaled toward a public commitment of 1 million free AI-assisted DR screenings in resource-constrained Thai communities over the next decade, alongside a related licensing agreement extending the model to India. • Instant, same-visit results replaced multi-week waits for grading, changing the patient experience from open-ended uncertainty to immediate triage and referral.
Takeaways	The renewed partnership with a strong university collaboration allowed for a concerted approach which allowed for long-term digital and AI transformation. With the university effectively playing a *Sandbox* role initially for the re-launched initiative and an *Accelerator* role for subsequent scale up. Due to foresight deficit, the governance gaps were not considered in the initial attempt. An AI model becomes deployable not with lab accuracy, but when it has been credibly tested, tempered and trusted with partners in the real world.

Appendices

4 Project SARAI in the Philippines: University research advancing agriculture with satellite imagery and AI

Situation	Traditional survey-based estimation of crop yield is labour intensive, slow and prone to respondent recall bias, and time-lags between actual harvest and release of final statistics. This leaves planners, policy and decisionmakers, and other stakeholders blind to shortages or supply chain bottlenecks.
Challenge	The lack of near-real time and high-quality ground truth data together with the susceptibility of optical satellite imagery to cloud cover and atmospheric disturbances can lead to unreliable and flawed crop yield estimates.
Response	To address survey-related lag and recall bias, the Smarter Approaches to Reinvigorate Agriculture as an Industry in the Philippines (SARAI) Project introduced tools that agricultural extension workers can use for digitising and recording of geotagged crop data. The tools developed address cloud cover and atmospheric interference with satellite data for accurate crop yield estimates and even flood damage assessment.
Impact	• 99.2% accuracy in monitoring standing crops and 99.7% accuracy in estimating post-flooding and cyclone events represent major improvement as farmers underreport yields by about 568 kg/ha or misreport based on plot size. • Highly accurate farm-level yield forecasts help Local Government Units (LGUs) to prepare for extreme weather, deploy targeted interventions (e.g. early harvesting, mobile dryers), and coordinate hauling logistics early to prevent post-harvest waste. • LGUs can generate science-based, site-specific crop advisories transitioning their roles as passive information receivers into proactive, evidence-based decisionmakers.
Follow up Projects for Universities	*Research and Extension.* Building on similar pioneering successes in countries like India, the United States, and European Union member states, advances in crop health monitoring, yield prediction, climate risk modelling and early warning systems are prioritised, along with transdisciplinary research and co-design of advisory systems with farmers, LGUs and other industry stakeholders.
Takeaways	*University Roles.* The project demonstrates a classic *Sandbox* role for universities, generating evidence from testing satellite against unmanned aerial vehicle (UAV) and local data. For subsequent phases, the university could play wider roles to act as *Bridge* – connecting researchers, government agencies, and local agricultural units. Subsequent deployment of systems and tools, training, and regional implementation could lead to an *Accelerator* role by transferring capability beyond the original research team. *Governance and Scaling.* Governance response depends on conditional and locally validated use. Models should be reassessed across crops, regions, seasons and extreme-weather conditions, and local officials should constantly question estimates where they conflict with field evidence. The transferable lesson is that a university can connect public research, open data, local knowledge and institutional training to reduce dependence on commercial systems, demonstrate epistemic discipline of disclosing what has been validated, and what remains uncertain.
References	Faustino-Eslava, D., Cui, L., Rosemarie, L., Areglado, M., Regina, V., Regalado, M., Ericha, V., Montecillo, M., Dorado, Leo, N., Viado, C., Khan, J., Hao, A., Salazar, M., Victoria, O., & Espaldon. (n.d.). *Mutually exclusive benefits and challenges in AI and Big Data for agriculture: Leveraging the SARAi experience.* Retrieved September 28, 2026, from https://www.the-easia.org/jrp/pdf/w23/presentation/invited_speaker/agriculture/decibel_faustino_eslava.pdf Dehaven, A., & University, S. D. S. (n.d.). *Algorithm allows farmers to monitor crops in real time.* Phys.Org. Retrieved September 28, 2026, from https://phys.org/news/2023-09-algorithm-farmers-crops-real.html

This edition is organised by

Asia-Europe Foundation

ASEF is an intergovernmental not-for-profit organisation located in Singapore. Founded in 1997, ASEF promotes understanding, strengthens relationships and facilitates cooperation among the people, institutions and organisations of Asia and Europe. ASEF enhances dialogue, enables exchanges and encourages collaboration across the thematic areas of culture, education, governance, sustainable development, economy, public health and media. For more information, please visit https://asef.org/.

In partnership with

Fudan University

Fudan University is a major public research university in Shanghai, People's Republic of China. Founded in 1905, today it is widely considered as one of the most prestigious and selective universities in the country. The QS University Rankings 2021 ranked Fudan as the 7th most reputable university in Asia, while it is classified as a Double First Class University by the Ministry of Education in China. Fudan also actively incubates high-tech industries and encourages them to convert knowledge to power. In return, the multi-pattern development of the high-tech industries helps the University to industrialise the research outcomes. For more information, please visit https://www.fudan.edu.cn/en.

University of Nottingham

University of Nottingham is a leading public research university renowned for world-class research, innovation, and interdisciplinary education. The university's mission emphasises excellence in education, research, and civic engagement, with a strong international outlook and commitment to improving lives worldwide. It serves a large and diverse student body, with tens of thousands of students from more than 150 countries enrolled across undergraduate, postgraduate, and research programmes spanning humanities, sciences, engineering, business, and health disciplines. They have several campuses hosting a comprehensive range of academic pathways in the United Kingdom, China, and Malaysia. For more information, please visit https://www.nottingham.ac.uk/.

Made possible with the support of

unesco Centre Under the auspices of UNESCO

IRCAI International Research Centre on Artificial Intelligence under the auspices of UNESCO

IGF Internet Governance Forum

INTERNATIONAL ASSOCIATION OF UNIVERSITIES INTERNATIONAL UNIVERSITIES BUREAU

DAAD Deutscher Akademischer Austauschdienst German Academic Exchange Service

Asia-Europe Meeting the ASEM Education Secretariat

AE4AI NETWORK

And funding from

People's Republic of China

Earmarked Funder



Flags represent countries that contributed to the ASEF General Pool for the previous year, as of 9 Sep 2026.