Preprint (author's submitted version). This chapter is forthcoming in: Büber Kaya, H. (Ed.), Responsible AI Use and Digital Solidarity: Human-Centered Approaches for a Sustainable Planet, Nova Science Publishers. Please cite the published version once available.

Responsible AI in Education and Governance: Cross-Cultural Perspectives

Hatice Büber Kaya Assist. Prof. Dr., Department of Software Engineering, Faculty of Engineering, Kırklareli University, Kırklareli, Türkiye, hatice.buberkaya@klu.edu.tr, ORCID: 0000-0002-4726-4412

Abstract

Artificial intelligence has become part of the routine operation of education systems, from admission and assessment algorithms to generative tools in classrooms, and the question of how its use should be governed has moved to the centre of educational policy. This chapter examines the governance of responsible AI in education through a cross-cultural lens. Drawing on a synthesis of international policy instruments, comparative regulation scholarship, and empirical studies of public attitudes, it traces a consistent pattern: ethical principles converge at the level of vocabulary while their interpretation, institutional form, and enforcement diverge along cultural and institutional lines. The chapter maps the three-level governance architecture that has formed around AI in education, comprising supranational instruments, national strategies, and institutional policies, and argues that the gap between principle and practice is produced across these levels rather than within any single one. It then examines why governance frameworks travel unevenly across cultural settings, and grounds the analysis in the case of Türkiye, a bridging country subject to European regulatory pull while sharing capacity constraints with much of the Global South. The chapter closes by outlining a set of commitments for culturally responsive AI governance in education, directed at policymakers, institutional leaders, and educators.

Keywords: responsible artificial intelligence; educational governance; cross-cultural perspectives; artificial intelligence policy; higher education; Türkiye

1. Introduction

Artificial intelligence has settled into the everyday machinery of education. Algorithms inform admission decisions, flag students considered at risk, score examinations, and personalise the sequence of learning materials, while generative tools have entered classrooms faster than most institutions have been able to write rules for them. What was until recently a specialist debate about emerging technology has therefore become a governance question in the ordinary sense of the word: who decides how these systems are used, under which rules, and who answers when they fail a learner. The stakes are high because educational decisions are formative in a way that few other algorithmic decisions are: a recommendation engine that misjudges a consumer loses a sale, whereas an assessment algorithm that misjudges a student may alter an educational trajectory for years.

Policy responses have accumulated quickly. International organisations, national governments, and individual institutions have produced an extensive body of principles, strategies, and guidelines for the responsible use of AI, and education figures prominently in nearly all of them. Yet this productivity conceals two difficulties. The first is a gap between principle and practice: reviews of the guideline corpus find broad agreement on principles such as transparency, fairness, and accountability alongside persistent disagreement about what those principles require and who is responsible for realising them (Jobin et al., 2019). The second difficulty is less frequently acknowledged. The documents that define responsible AI have been produced overwhelmingly in a small number of economically dominant countries, which raises the question of whose values, institutional assumptions, and educational traditions are built into frameworks that present themselves as universal. Both difficulties suggest that the study of responsible AI in education cannot remain a study of texts; it must become a study of how texts meet contexts.

The present chapter takes up that task as a conceptual and policy synthesis. It offers no new empirical data. Its contribution lies instead in bringing two discussions together that usually run in parallel: the mapping of the global governance architecture for AI in education, and the cross-cultural analysis of how governance frameworks are produced, transferred, and translated. The chapter argues that responsible AI in education converges at the level of principle and diverges at the level of practice, and that this divergence follows cultural and institutional lines rather than random variation. If that argument holds, then the central design problem of the coming years is not the formulation of better principles but the construction of governance arrangements that can absorb cultural difference without abandoning shared commitments. The idea of digital solidarity that frames this volume names the normative ground on which such arrangements would rest, and the conclusion returns to it explicitly.

The chapter proceeds in five steps. It first clarifies its conceptual foundations, distinguishing responsible AI from adjacent vocabularies and specifying what governance means across supranational, national, and institutional levels. It then maps the global governance architecture that has accumulated around AI in education, from the UNESCO Recommendation to the EU Artificial Intelligence Act, and identifies where its apparent consensus begins to thin. The third step applies a cross-cultural lens to this architecture, asking why the same principles land differently in different settings. The fourth step grounds the analysis in the experience of Türkiye, a country whose position between regulatory traditions makes the tensions of transfer unusually visible. The final step draws these threads into a set of commitments for culturally responsive AI governance and considers what they imply for policymakers, institutional leaders, and educators.

2. Conceptual Foundations: Responsible AI in Education

Debates about the governance of artificial intelligence in education draw on three partially overlapping vocabularies: ethical AI, trustworthy AI, and responsible AI. Ethical AI generally

denotes the alignment of AI systems and their uses with moral principles, and it is the language most visible in the wave of guidelines issued by governments, companies, and professional bodies since the late 2010s (Jobin et al., 2019). Trustworthy AI, a term closely associated with European policy, frames the problem as one of warranted public trust and specifies that AI systems should be lawful, ethical, and technically reliable throughout their life cycle (High-Level Expert Group on Artificial Intelligence [AI HLEG], 2019). Responsible AI, in the sense developed by Dignum (2019), shifts the emphasis from the properties of systems to the obligations of the people and institutions that design, deploy, and regulate them. This chapter adopts responsible AI as its working term precisely because of that shift: a governance perspective is concerned less with whether an algorithm can be labelled ethical than with who answers for its consequences, under which rules, and before which authority.

At the level of principles, the three vocabularies appear to converge to a considerable degree. In a review of 84 AI ethics documents produced by public and private bodies worldwide, Jobin et al. (2019) identified an emerging global convergence around five principles: transparency, justice and fairness, non-maleficence, responsibility, and privacy. Hagendorff (2020) reached a broadly similar conclusion from a separate analysis of 22 major guidelines, in which accountability, privacy, and fairness appeared almost universally. Both reviews, however, also report substantive divergence beneath this apparent consensus, since the documents differ in how the principles are interpreted, why they are considered important, and to whom their implementation is assigned (Jobin et al., 2019). Hagendorff (2020) goes further and argues that most guidelines lack enforcement mechanisms altogether, so that adherence remains largely voluntary and deviation carries few tangible consequences. The pattern that emerges from these analyses, shared vocabulary at the level of principle and persistent uncertainty at the level of practice, indicates that the interesting questions about responsible AI are now questions of governance rather than of principle formulation.

Translating this general debate into education requires a clear picture of what AI systems actually do in educational settings, because different applications generate different governance problems. Systematic review evidence offers a usable typology. Zawacki-Richter et al. (2019), synthesising 146 studies on AI in higher education published between 2007 and 2018, grouped applications into four areas: profiling and prediction (including admissions decisions and the identification of students at risk of dropping out), intelligent tutoring systems, assessment and evaluation, and adaptive systems and personalisation. Policy guidance extends this learner-facing picture with system-level uses, such as education management information systems and the automation of administrative processes (Miao et al., 2021). Generative applications, which reached classrooms mainly after these documents were prepared, have since broadened the range of uses without changing the underlying categories of concern.

Each application type carries a distinct governance burden. Profiling and prediction depend on the large-scale collection and reuse of learner data, which raises questions of privacy, consent,

and purpose limitation, particularly where the data subjects are minors. Intelligent tutoring and adaptive systems embed pedagogical models in software, so their governance concerns transparency: teachers and learners may be unable to see, let alone question, the assumptions about learning that shape what the system recommends. Automated assessment attaches algorithmic judgment to consequential decisions about progression and certification, which places fairness, accountability, and the possibility of contesting an outcome at the centre of concern. Administrative automation, finally, imports into education the due process questions long familiar from public administration, including who reviews a machine-generated decision and on what grounds. It is telling that Zawacki-Richter et al. (2019) found an almost complete absence of critical reflection on ethics and risks in the research they reviewed, which suggests that scholarship on educational AI has tended to lag behind the governance questions its own applications generate.

The educational context also adds obligations that generic AI ethics does not capture. Holmes et al. (2022) argue that the ethics of AI in education cannot be reduced to the ethics of data or of computation, because education involves learners who are often children, who participate under varying degrees of compulsion, and who stand in a markedly unequal power relation to the institutions and systems that assess them. Decisions embedded in educational AI are, in this reading, pedagogical decisions with long-term formative consequences, not merely technical ones. UNESCO's policy guidance points in the same direction when it frames the use of AI in education around a human-centred approach oriented toward equity and inclusion rather than efficiency alone (Miao et al., 2021).

Against these definitional foundations, this chapter uses governance in a deliberately broad sense that comprises three components. The first is formal rules: laws, regulations, standards, and codified guidelines. The second is institutions: the ministries, agencies, quality assurance bodies, and organisational units that produce, monitor, and enforce those rules. The third is implementation practice: the everyday routines through which rules are interpreted, applied, stretched, or quietly ignored in schools and universities. Governance so understood operates at three levels: a supranational level of recommendations and regulations that reach beyond any single state, a national level of strategies and ministerial policy, and an institutional level of university and school policies, procurement choices, and classroom practice. This chapter deliberately treats all three levels together, on the assumption that the gap between principle and practice is produced across levels rather than within any single one.

One further conceptual commitment underpins the analysis. Responsible AI in education is not solely a problem of technical standardisation, although standards for privacy, security, and explainability are clearly necessary. Whether an adaptive platform respects learner autonomy appears to depend at least as much on the pedagogical values designed into it, on the capacity of the adopting institution to configure, monitor, and contest it, and on the power relations among vendors, ministries, teachers, and learners as on any measurable property of the

software. The same technical system may operate responsibly in one institutional configuration and irresponsibly in another. This chapter therefore treats responsible AI as a property of socio-technical arrangements rather than of artefacts, an orientation that keeps questions of value and power visible alongside questions of engineering.

These foundations leave one question conspicuously open. If the principles are broadly shared while their interpretation and enforcement remain contested, then the analytical weight falls on the machinery through which principles acquire institutional form: the international recommendations, regional regulations, and national strategies that have accumulated around AI and education over the past decade. How far that machinery has actually converged, and where its coherence ends, is a question about the global governance architecture itself.

3. The Global Governance Architecture

Responsible AI in education is no longer shaped by professional ethics debates alone. Between 2019 and 2024, supranational instruments, national strategies and institutional policies accumulated into what this chapter treats as a three-level governance architecture, and each level claims a distinct form of authority. Supranational bodies set principles and, increasingly, binding obligations; national governments translate these into strategies and regulation; schools and universities write the rules that actually reach classrooms. The distinction matters because the three levels do not simply nest into one another. A principle adopted in Paris or Brussels may arrive in a lecture hall considerably changed, if it arrives at all, and the mechanisms that carry it downward differ sharply from one country to another. This section maps the supranational, national and institutional levels in turn, and then asks what the resulting architecture agrees on and where its apparent consensus begins to thin.

The supranational level is anchored by the UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted by the General Conference in November 2021 as the first global standard-setting instrument on the ethics of AI (UNESCO, 2021). Education occupies a dual position in the text. It is one of the policy areas in which member states are asked to act, and it is at the same time the mechanism through which the wider ethical programme is meant to work, since the Recommendation calls for open and accessible education, digital skills and AI ethics training, and media and information literacy as routes to public understanding of AI and data (UNESCO, 2021). The document also arrives with implementation machinery, including a readiness assessment methodology intended to help states gauge their capacity to apply its provisions, which distinguishes it from earlier and purely declaratory ethics texts.

The release of generative AI tools pushed UNESCO toward something more operational. Its Guidance for Generative AI in Education and Research, published in 2023 as the first global guidance of its kind, argues for a human-centred approach in which generative systems support rather than replace human capabilities (Miao & Holmes, 2023). The Guidance asks governments to regulate providers, to protect learner data and to invest in teacher preparation,

and it proposes thirteen as a minimum age for the independent use of conversational generative AI in classrooms (Miao & Holmes, 2023). The specificity of that last provision indicates how quickly the supranational conversation moved from abstract principle to concrete pedagogical rule.

The OECD occupies an adjacent position with a different constituency. Its Recommendation of the Council on Artificial Intelligence, adopted in 2019 as the first intergovernmental standard on AI, sets out value-based principles for trustworthy AI, and its revision of May 2024 responded to the emergence of general-purpose and generative systems with particular attention to information integrity, misuse and transparency (OECD, 2024). The Recommendation is not an education instrument as such, yet its principles supply the reference vocabulary for many of the national strategies discussed below, and its adherents, now 47 in number, include the European Union alongside states well beyond the OECD membership (OECD, 2024).

A further hardening of the supranational layer came from the Council of Europe. Its Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, adopted on 17 May 2024 and opened for signature on 5 September 2024, is the first legally binding international treaty on AI (Council of Europe, 2024). Its initial signatories included the European Union, the United Kingdom, the United States and Israel, which suggests that a treaty negotiated in Strasbourg may come to structure obligations well outside Europe (Council of Europe, 2024). For education, the Convention matters less through sector-specific provisions than through its requirement that activities within the life cycle of AI systems remain consistent with human rights, democracy and the rule of law, a formulation broad enough to cover public education systems.

The clearest move from principle to enforceable law is the EU Artificial Intelligence Act, Regulation (EU) 2024/1689, which classifies AI systems according to the risk they pose (European Union, 2024). Education and vocational training appear in Annex III of the Regulation, the list of high-risk systems referred to in Article 6(2), through four use cases: systems that determine access or admission to educational institutions, systems that evaluate learning outcomes, systems that assess the level of education an individual will receive, and systems that monitor and detect prohibited behaviour of students during tests (European Union, 2024). Providers and deployers of such systems face conformity assessment, documentation and human oversight obligations before and after deployment. The Act follows a staged timetable: it entered into force on 1 August 2024, its prohibitions and AI literacy provisions have applied since 2 February 2025, rules for general-purpose AI models followed on 2 August 2025, and the bulk of the high-risk regime, including the education use cases, is scheduled to apply from 2 August 2026 (European Commission, n.d.). Education has thereby become one of the few sectors in which responsible AI carries the threat of legal sanction rather than reputational cost alone.

National strategies form the second level, and their sheer volume is itself a finding. The OECD.AI Policy Observatory, which has collected data on state activity through a structured government survey since 2019, gives access to more than 900 national AI policies and initiatives (OECD.AI, 2026). Education tends to appear in these strategies in two roles at once. It is treated as a supply channel, the place where AI talent and workforce skills are produced, and as an application domain in which intelligent tutoring, learning analytics and administrative automation promise efficiency gains. This pattern has been documented systematically: in a thematic analysis of 24 national AI strategies, Schiff (2022) found that education figures overwhelmingly as an instrument for producing an AI-ready workforce, while the use of AI within education, and its ethical implications, remain largely absent from the same documents. Strategy documents of this kind also occupy an intermediate position in the architecture: they are domestic in origin yet frequently borrow their normative vocabulary from the supranational instruments described above, which is one reason the three levels can appear more coordinated than they are. The balance between the two roles varies considerably across strategies, as does the attention given to safeguards, and this variation is an early sign of the divergence examined below. What national strategies rarely do, however, is specify who inside an education system carries responsibility when an AI system fails a learner.

The institutional level is the youngest and the least settled. Universities and school systems began issuing their own generative AI policies almost immediately after conversational tools became publicly available, yet the early evidence suggests a reactive and uneven process. Moorhouse et al. (2023) reviewed the publicly available guidelines of the world's top 50 universities, as ranked by Times Higher Education, and found that fewer than half had developed and published guidance on the use of generative AI in assessment at the time of their review, and that the documents which did exist concentrated on academic integrity. Institutional policy is nonetheless where the architecture touches actual teaching, since it is departmental rules, assessment briefs and syllabus statements (rather than treaties) that tell a student what responsible use means in a given course. Early institutional responses also reveal a translation problem, since a university policy must convert broad commitments to integrity and transparency into decisions about specific assignments, disclosure statements and permitted tools. The thinness of this layer relative to the density above it may be the architecture's most consequential imbalance.

Viewed together, the three levels display a striking convergence of vocabulary. Human-centredness, transparency, accountability and fairness recur from the UNESCO Recommendation to institutional assessment rules, and this repetition echoes an earlier finding: in an analysis of 84 AI ethics documents, Jobin et al. (2019) identified global convergence around five principles alongside substantive divergence in how those principles are interpreted, justified and implemented. The governance of AI in education appears to reproduce exactly this pattern. The instruments agree that AI in education should be human-centred; they diverge, or fall silent, on who verifies compliance, on what fairness requires of an admission algorithm

or a proctoring system in a particular examination culture, and on whether responsibility rests with the ministry, the institution, the teacher or the vendor. Convergence, in other words, holds at the level of principle and begins to fracture at the level of application. The fracture lines do not run randomly. They follow administrative traditions, assessment cultures and understandings of the teacher's role that differ across societies, which suggests that the next question is not primarily legal but cultural: why does the same principle land so differently in different educational systems?

4. The Cross-Cultural Lens

The convergence documented in the preceding section rests on a narrower foundation than its universal vocabulary suggests. In their review of 84 AI ethics guidelines, Jobin et al. (2019) found that the United States (21 documents) and the United Kingdom (13) together produced more than a third of the corpus, while Africa, South and Central America, and Central Asia were barely represented at all. The authors read this distribution as evidence of a power imbalance in which economically dominant countries shape the global debate, with the attendant risk that local knowledge and cultural pluralism are pushed to the margins. This chapter takes that finding as its point of departure, because it complicates an assumption that runs quietly through much of the governance architecture: that ethical principles formulated in one part of the world can simply be applied everywhere else. The assumption has a name, ethical universalism, and it deserves scrutiny rather than acceptance. Principles such as fairness or transparency may command near universal assent at the level of the word, yet what counts as fair treatment of student data, or adequate transparency toward a parent, is likely to be answered differently in different institutional and cultural settings. The geography of guideline production therefore matters, because it determines whose answers become the default.

Two related bodies of work push this argument further. Couldry and Mejias (2019) describe the contemporary data economy as a form of data colonialism, in which human life itself is appropriated as raw material through data extraction, an arrangement they present as continuous with the extractive logic of historical colonialism rather than merely analogous to it. Working from a related position, Mohamed et al. (2020) propose decolonial theory as a form of sociotechnical foresight for AI, identify practices such as the testing of systems on populations in the Global South and the outsourcing of low-paid data labour as contemporary sites of coloniality, and call for the meaningful inclusion of marginalised communities in the design and governance of these systems. One does not need to accept every element of these arguments to take their central lesson seriously. AI systems, and the rules written for them, are not neutral instruments that happen to land in different cultures; they carry the values, interests, and blind spots of the places where they are made. A governance discussion that ignores culture is therefore likely to reproduce existing hierarchies rather than correct them.

The claim that cultural context shapes responses to AI is not only theoretical; it has empirical support, although that support requires careful handling. The largest dataset comes from the

Moral Machine experiment, which collected roughly 40 million judgements about autonomous vehicle dilemmas from participants in 233 countries and territories (Awad et al., 2018). The study found broadly shared preferences, for instance for sparing human lives, alongside three distinguishable clusters of countries whose moral preferences diverged in ways that correlated with cultural and institutional characteristics. Survey evidence points in a similar direction. In a study of 10,005 respondents across eight countries on six continents, Kelley et al. (2021) reported that public sentiment toward AI (organised in their analysis around excitement, usefulness, worry, and futuristic framing) varied considerably across national settings, with respondents in countries such as India and Nigeria expressing more enthusiasm than those in high-income countries, where ambivalence was more common. Two cautions apply to such findings. First, frameworks that reduce culture to national scores, of which Hofstede's (2001) dimensional model remains the most widely used, have been criticised for treating nations as internally homogeneous and culture as a fixed causal essence (McSweeney, 2002), and this chapter shares that concern, since cross-cultural analysis slides easily into stereotype. Second, attitudinal variation does not translate mechanically into policy variation; it indicates that publics bring different expectations to AI, not that any country holds a single view. Read with these caveats, the evidence still supports a modest but consequential conclusion: uniform governance instruments will meet non-uniform publics.

Institutional divergence is easier to document. Comparative work on digital regulation identifies three distinct models: a market-driven American approach that privileges innovation and private ordering, a state-driven Chinese approach that ties AI development to public authority and social control, and a rights-driven European approach that subordinates both market and state to individual and collective rights (Bradford, 2023). Peer-reviewed comparison of the European Union and the United States reaches compatible conclusions, showing that the two pursue a good AI society from different starting assumptions about the relationship between citizens, markets, and the state (Roberts et al., 2021). These models are ideal types rather than descriptions; most jurisdictions combine elements of all three, and regulatory positions shift over time. The point for this chapter is narrower. The divergence is not noise around an emerging universal standard but the expression of durable institutional and political traditions, which suggests that the implementation gap identified earlier is not a temporary coordination failure. It reflects, at least in part, the fact that societies want different things from AI governance.

These differences reach the classroom through several channels. Regulatory models shape what schools may do with student data: the same learning analytics dashboard may be routine in one jurisdiction and legally precarious in another, and expectations about parental consent, data retention, and commercial reuse of educational records vary with the surrounding privacy regime. Curricula differ in how, and whether, they address AI at all, and the degree of professional autonomy that teachers hold conditions whether responsible AI guidance is enacted in practice or remains a document. Generative AI adds a further, less visible channel:

language. The development of large language models has concentrated on English, and their training corpora overrepresent the perspectives of those most present online (Bender et al., 2021). When Lai et al. (2023) evaluated ChatGPT across seven tasks and 37 languages spanning high, medium, low, and extremely low resource levels, they found its performance markedly uneven and, for many tasks and languages, below that of specialised models, a result that casts doubt on the assumption that English-centred tools serve all linguistic communities equally well. UNESCO's guidance on generative AI in education identifies the corresponding risk to linguistic and cultural diversity (Miao & Holmes, 2023). For a student working in a lower-resource language, the system that mediates a growing share of learning is likely to be weaker, and culturally thinner, than the system available to an English-speaking peer. Equity in AI-supported education is therefore partly a question of language infrastructure, a dimension that principle-level documents rarely address.

Viewed from the Global South, the cross-cultural question becomes a capacity question. The asymmetry in who writes ethics guidelines (Jobin et al., 2019) sits on top of deeper asymmetries in computing infrastructure, research funding, data resources, and regulatory expertise, so that many countries encounter responsible AI frameworks primarily as importers rather than authors. Policy transfer under these conditions has limits. A rights-based instrument modelled on the European approach presupposes data protection authorities, enforcement mechanisms, and administrative capacity that cannot be assumed everywhere, and a framework transplanted without them may deliver the vocabulary of responsible AI without its substance. The decolonial critique presses the point further: if frameworks travel only from North to South, even well-intentioned transfer risks repeating the pattern in which the South supplies data and test populations while the North supplies norms (Couldry & Mejias, 2019; Mohamed et al., 2020). This chapter does not conclude that international frameworks are useless; supranational instruments remain among the few channels through which smaller and less resourced education systems can influence global rules. It concludes, more cautiously, that transfer without translation is unlikely to produce responsible outcomes.

The argument so far has moved at the level of regions and regulatory models, and that is where cross-cultural analysis is weakest, because principles, strategies, and capacities collide not in typologies but inside particular education systems. Observing such a collision requires a country that cuts across the categories used in this section: one subject to European regulatory pull, exposed to capacity constraints more typical of the Global South, and engaged in building its own national AI policy for education. Türkiye occupies precisely this position, and its experience gives the tensions traced above an institutional shape that regional comparison cannot provide.

5. Türkiye as a Bridging Case

Türkiye's structural position is worth stating precisely: a middle-income economy with OECD membership, a candidate for European Union accession since 1999, and a country whose

institutions have absorbed regulatory templates from the Global North while sharing capacity-building concerns with much of the Global South. The intent of this positioning is analytic rather than evaluative. Türkiye is presented neither as a model to be emulated nor as a cautionary tale; it is examined because the tensions described in the preceding sections, between globally circulating governance frameworks and locally grounded implementation conditions, can be observed here within one governance system rather than across several. A bridging case of this kind may reveal dynamics that comparisons between clearly Northern and clearly Southern contexts tend to obscure.

At the national level, the policy framework rests on the National Artificial Intelligence Strategy 2021-2025, prepared jointly by the Presidency's Digital Transformation Office and the Ministry of Industry and Technology and adopted through Presidential Circular No. 2021/18 (Republic of Türkiye Digital Transformation Office & Ministry of Industry and Technology, 2021). The strategy organizes its measures under six priorities, and the development of human capital (graduate training, workforce programs, and public awareness efforts) appears among the most visible of these. An updated action plan for 2024-2025, approved by the strategy's steering committee in line with the Twelfth Development Plan, shifts attention toward generative AI, the development of Turkish large language models, and the expansion of expert human resources (Republic of Türkiye Digital Transformation Office, 2024). Education therefore appears in these documents in a dual role: it is a sector to be governed responsibly, and it is also the main mechanism through which the specialists the strategy requires are expected to emerge.

The higher education layer illustrates how global ethical vocabularies are translated into national instruments. In 2024, the Council of Higher Education circulated to all universities an ethical guide on the use of generative AI in scientific research and publication, built around principles such as transparency, honesty, privacy, and accountability (Council of Higher Education, 2024). These principles echo the international frameworks discussed earlier, yet the guide responds to a distinctly national situation, since Turkish universities are expected to carry much of the strategy's human capital agenda (İçen, 2022) while simultaneously regulating the very tools their researchers are encouraged to adopt. Institutional capacity for this dual task appears unevenly distributed. A descriptive analysis of 41 university AI units in Türkiye found that these units operate mostly as research centers within public universities and that their vision statements refer most frequently to training qualified human resources (Serpil & Kesim, 2024). This pattern suggests an ecosystem oriented more toward workforce production than toward governance research, which matters for a country importing governance frameworks it did not design.

The school system has followed a similar path with its own instruments. The Ministry of National Education put the Artificial Intelligence in Education Policy Document and Action Plan (2025-2029) into force in June 2025, a roadmap organized around four objectives, fifteen

policies, and forty actions that extend from curriculum revision to teacher training and AI-supported assessment (Ministry of National Education, 2025). The ministry subsequently issued an ethics guide for AI applications in education, which establishes ethics boards at provincial and school levels and introduces an online ethical declaration system for teachers (Ministry of National Education, 2026). Publishing the policy document in English signals, at least in part, an intention to participate in international policy exchange rather than to treat AI governance as a purely domestic concern. Read together, the national strategy, the higher education guide, and the school-level instruments form a reasonably coherent vertical chain, a degree of alignment that cannot be assumed in advance in middle-income systems.

What makes Türkiye a bridging case, rather than simply another national example, is the direction of its regulatory gravity. The European Union's Artificial Intelligence Act (Regulation 2024/1689) exerts influence well beyond the Union's borders, and this influence is visible in the Turkish legislative agenda: a private member's bill submitted to the Grand National Assembly in June 2024 proposes a risk-based framework for AI systems that parallels the Act's structure, although it remained in committee at the time of writing (Grand National Assembly of Türkiye, 2024). The national strategy itself names harmonization with international regulatory standards among its aims. At the same time, textual alignment does not guarantee implementation capacity. Qualitative work with Turkish academics points to substantial expectations placed on universities for workforce development (İçen, 2022), and the institutional mapping cited above indicates that governance expertise occupies a smaller place in the ecosystem than technical training does. The productive tension in the Turkish case is therefore not between accepting and rejecting Global North frameworks, a framing that recurs in the critical literature, but between the relatively rapid adoption of regulatory texts and the slower accumulation of the enforcement structures and pedagogical capacity those texts presuppose.

This configuration is unlikely to be unique to Türkiye. Candidate countries and other middle-income states that orbit large regulatory blocs occupy a comparable structural position: they import governance templates whose cultural and institutional assumptions were formed elsewhere, and they must localize those templates with fewer specialized resources. The Turkish experience suggests that culturally responsive AI governance involves more than adjusting ethical content to local values; it also involves sequencing, that is, deciding which imported commitments a system can credibly implement first and which must wait for capacity to develop. A synthesis built on this observation would ask what responsible AI governance in education looks like when it is designed from the position of such bridging systems rather than merely extended to them, and it is to this question that the final section turns.

6. Toward Culturally Responsive AI Governance in Education

The preceding sections yield a diagnosis that can be stated briefly. The governance of AI in education rests on principles that are widely shared in name, embedded in an architecture

whose three levels are unevenly developed, and transferred across cultural and institutional settings that differ in what they expect from education, how they regulate data, and what capacity they can devote to enforcement. If this diagnosis is broadly correct, the design question that follows is how governance arrangements might be built to take cultural difference seriously without dissolving into relativism. This chapter proposes no named model and no acronym; it proposes a small set of commitments that can be read as design principles for the framework sketched here, each of which follows from the analysis above.

The first commitment is translation rather than adoption. The Turkish case indicates that regulatory texts travel faster than the institutional conditions they presuppose, and the cross-cultural evidence suggests that publics bring different expectations to the same systems. A ministry or university that imports a governance instrument therefore takes on a task of active translation: identifying which assumptions of the source framework hold locally, which must be rebuilt with local material, and which should be rejected as poor fits for the local pedagogical culture. Translation in this sense is not dilution. A framework that has been argued through against local assessment cultures and administrative traditions is likely to bind more strongly, not less, than one adopted verbatim.

The second commitment is capacity equalisation. The analysis of the Global South position showed that frameworks travel from where capacity is concentrated to where it is scarce, and that transfer without capacity delivers vocabulary without substance. Culturally responsive governance therefore treats investment in regulatory expertise, data infrastructure, and teacher preparation not as a complement to principle formulation but as part of governance itself. The sequencing observation from the Turkish case belongs here: systems with limited capacity may govern more responsibly by implementing fewer commitments credibly than by adopting many commitments nominally.

A third commitment concerns participation. The critique of ethical universalism does not entail that every culture requires its own ethics of AI; it entails that those affected by governance arrangements should have a hand in shaping them. In education this means more than consulting ministries. Teachers hold the professional knowledge on which any workable classroom rule depends, students hold the experiential knowledge of what these systems do to learning, and both groups appear only marginally in the instruments reviewed in this chapter. Participatory governance also operates between countries: supranational fora remain among the few channels through which less resourced education systems can influence rules they will otherwise merely receive.

The fourth commitment concerns language and content plurality. The evidence on multilingual performance indicates that generative systems serve linguistic communities unevenly, which turns language infrastructure into a governance issue rather than a market outcome. Education systems that depend on tools optimised for English delegate part of their curriculum, by default, to the linguistic and cultural priorities of others. Governance that takes this seriously will treat

support for local language resources, and scrutiny of the cultural assumptions embedded in imported tools, as standing obligations.

The last commitment, monitoring with teeth, needs the least introduction: every review cited in this chapter converges on the observation that enforcement is the weakest link in the architecture. Principles without verification mechanisms drift toward decoration, and the institutional level, where verification would have to happen, is precisely the thinnest layer of the architecture. Culturally responsive governance therefore requires that each level specify who checks compliance, at what intervals, and with what consequences, and that these mechanisms be designed for the administrative culture in which they must operate rather than copied from elsewhere.

These commitments carry different implications for different actors. For policymakers, the practical lesson is that importing a well-regarded framework is the beginning of a governance process, not its conclusion; the work lies in translation, sequencing, and the deliberate construction of enforcement capacity. For institutional leaders, the lesson is that the institutional layer is where the architecture currently fails, which makes university and school policies, procurement decisions, and assessment rules the most consequential governance documents most learners will ever encounter. For educators, the lesson is that responsible AI is not a compliance regime imposed from above but a professional practice: the judgment of a teacher who understands what an adaptive system assumes about learning remains among the few mechanisms that can catch governance failures where they occur.

This synthesis has limits that should be named. It rests on documents and published studies rather than on new empirical work, and the policy field it describes is moving quickly enough that some of its details will age within years. In this respect the chapter shares the limitation it diagnoses in the wider field: it works with texts, and the distance between texts and classroom practice can be established only by empirical study. The commitments proposed here are derived from the analysis but have not been tested against implementation data, and testing them would require precisely the kind of cross-cultural empirical research that the chapter has argued is scarce. These limits define a research agenda as much as they qualify a conclusion.

7. Conclusion

The chapter set out to examine the governance of responsible AI in education through a cross-cultural lens, and its argument can be restated in one movement. The principles of responsible AI converge in vocabulary while diverging in interpretation and enforcement; the governance architecture built on those principles is dense at the top and thin where it meets classrooms; and the divergence between settings follows durable cultural and institutional lines, which means it will not be closed by better drafting alone. The experience of bridging countries such as Türkiye suggests that the decisive work of governance happens in translation: in the slow

adjustment between imported commitments and local capacity, values, and pedagogical traditions.

Read in this way, responsible AI in education is better understood as a relationship to be maintained among unequal partners than as a standard to be met once, and this is where the concerns of this volume converge. Digital solidarity, understood as the commitment to ensure that the benefits and burdens of digital transformation are shared across communities rather than concentrated within them, describes fairly precisely what culturally responsive governance attempts to institutionalise. A governance regime that equalises capacity, admits many voices, and protects linguistic plurality is solidarity given administrative form.

Future research can carry this agenda forward along three lines: comparative empirical studies of how identical governance instruments are implemented in different cultural settings, closer examination of bridging countries whose position between regulatory traditions makes translation visible, and evaluation research on whether culturally adapted governance arrangements in fact produce more responsible outcomes for learners. The principles are, by now, largely written; whether they come to matter will be decided in the places where they must be translated, resourced, and enforced.

References

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The Moral Machine experiment. *Nature*, *563*(7729), 59-64. https://doi.org/10.1038/s41586-018-0637-6

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Bradford, A. (2023). *Digital empires: The global battle to regulate technology*. Oxford University Press.

Council of Europe. (2024). *Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

Council of Higher Education. (2024). *Ethical guide on the use of generative artificial intelligence in scientific research and publication activities of higher education institutions* [Yükseköğretim kurumları bilimsel araştırma ve yayın faaliyetlerinde üretken yapay zekâ kullanımına dair etik rehber]. https://eski.yok.gov.tr/Sayfalar/Haberler/2024/yuksekogretim-kurumlari-bilimsel-arastirma-ve-yayin-faaliyetlerinde-uretken-yapay-zeka-kullanimina-dair-etik-rehber.aspx

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer. https://doi.org/10.1007/978-3-030-30371-6

European Commission. (n.d.). *Timeline for the implementation of the EU AI Act*. AI Act Service Desk. https://ai-act-service-desk.ec.europa.eu/en/ai-act/timeline/timeline-implementation-eu-ai-act

European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Grand National Assembly of Türkiye. (2024). *Yapay zeka kanun teklifi* [Artificial intelligence law proposal] (Esas No. 2/2234). https://www.tbmm.gov.tr/Yasama/KanunTeklifi/e21539a0-888a-4500-81be-01904a918c53

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines, 30*(1), 99-120. https://doi.org/10.1007/s11023-020-09517-8

High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Sage.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education, 32*(3), 504-526. https://doi.org/10.1007/s40593-021-00239-1

İçen, M. (2022). The future of education utilizing artificial intelligence in Turkey. *Humanities and Social Sciences Communications, 9*, Article 268. https://doi.org/10.1057/s41599-022-01284-4

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

Kelley, P. G., Yang, Y., Heldreth, C., Moessner, C., Sedley, A., Kramm, A., Newman, D. T., & Woodruff, A. (2021). Exciting, useful, worrying, futuristic: Public perception of artificial intelligence in 8 countries. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 627-637). Association for Computing Machinery. https://doi.org/10.1145/3461702.3462605

Lai, V. D., Ngo, N., Pouran Ben Veyseh, A., Man, H., Dernoncourt, F., Bui, T., & Nguyen, T. H. (2023). ChatGPT beyond English: Towards a comprehensive evaluation of large language models in multilingual learning. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 13171-13189). Association for Computational Linguistics. https://doi.org/10.18653/v1/2023.findings-emnlp.878

McSweeney, B. (2002). Hofstede's model of national cultural differences and their consequences: A triumph of faith - a failure of analysis. *Human Relations, 55*(1), 89-118. https://doi.org/10.1177/0018726702551004

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709

Ministry of National Education. (2025). *Artificial intelligence in education policy document and action plan (2025-2029)* [Eğitimde yapay zekâ politika belgesi ve eylem planı (2025-2029)]. https://yegitek.meb.gov.tr/meb_iys_dosyalar/2025_07/19150909_egitimde_yz_politika_belgesi_18072025_en.pdf

Ministry of National Education. (2026). *Ethics guide for artificial intelligence applications in education* [Eğitimde yapay zekâ uygulamaları etik kılavuzu]. https://ttkb.meb.gov.tr/meb_iys_dosyalar/2026_01/09152326_egitimdeyapayzekauygulamalari_etik_kilavuzu.pdf

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

Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open, 5*, Article 100151. https://doi.org/10.1016/j.caeo.2023.100151

OECD. (2024). *Recommendation of the Council on Artificial Intelligence* (adopted 2019, amended May 2024). https://oecd.ai/en/ai-principles

OECD.AI. (2026). *The OECD Artificial Intelligence Policy Observatory*. Retrieved July 18, 2026, from https://oecd.ai/en/

Republic of Türkiye Digital Transformation Office. (2024). *National artificial intelligence strategy 2024-2025 action plan* [Ulusal yapay zekâ stratejisi 2024-2025 eylem planı]. https://www.sanayi.gov.tr/assets/pdf/UlusalYapayZekaStratejisi2024-2025EylemPlani.pdf

Republic of Türkiye Digital Transformation Office & Ministry of Industry and Technology. (2021). *National artificial intelligence strategy 2021-2025* [Ulusal yapay zekâ stratejisi 2021-2025]. https://cbddo.gov.tr/SharedFolderServer/Genel/File/TRNationalAIStrategy2021-2025.pdf

Roberts, H., Cowls, J., Hine, E., Mazzi, F., Tsamados, A., Taddeo, M., & Floridi, L. (2021). Achieving a 'good AI society': Comparing the aims and progress of the EU and the US. *Science and Engineering Ethics, 27*(6), Article 68. https://doi.org/10.1007/s11948-021-00340-7

Schiff, D. (2022). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education, 32*(3), 527-563. https://doi.org/10.1007/s40593-021-00270-2

Serpil, H., & Kesim, E. (2024). Artificial intelligence units in Turkish universities: A descriptive analysis. *Open Praxis, 16*(4). https://doi.org/10.55982/openpraxis.16.4.725 [başlık dizgi öncesi birincil kaynaktan teyit edilecek]

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

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