**Journal of Asian Studies: Culture, Language, Art and Communications**
e,ISSN 2774,5554 | HAQI Publishing Service

**Volume 6 Issue 2, 2025**
https://journal.haqipub.com/index.php/jas/index

*E,mail:*
erik.vogel@tum.de

***Technical University of Munich (TUM)***
*School of Social Sciences and Technology*
*Munich, Germany*



# Developing a Human Rights and Democracy Framework for AI Governance in European Educational Settings

***Erik Vogel**

## Abstract

This study develops an operational governance framework for the deployment of artificial intelligence in the European education sector, ensuring compliance with human rights, the rule of law, and democratic principles established by the Council of Europe. Specifically, the framework addresses obligations arising from the European Convention on Human Rights (ECHR), the Council of Europe's Recommendation CM/Rec(2020)1 on algorithmic system impacts, and the European Union's Artificial Intelligence Act (2024), which designates educational AI systems as high-risk applications. Employing a mixed-methods approach combining Delphi methodology with legal and ethical consultation, this research engaged a multidisciplinary panel of 22 European experts specializing in educational law, AI ethics, democratic theory, digital pedagogy, and human rights adjudication. Through three iterative Delphi rounds, the study systematically identified high-risk AI scenarios in educational contexts, established normative consensus regarding implicated human rights and democratic values, and formulated operationally specific, ethically binding, and technically implementable governance guidelines. The resulting framework the European Educational AI Governance Framework (EEAGF) comprises four foundational pillars: (1) Dignity and Non-Discrimination, addressing prevention of automated systems reproducing structural inequalities; (2) Transparency and Explainability, safeguarding stakeholders' rights to understand AI-assisted decisions affecting educational trajectories; (3) Democratic Accountability, establishing institutional oversight mechanisms; and (4) Pedagogical Integrity, ensuring AI deployment serves rather than distorts education's fundamental democratic purposes. The Delphi process identified automated student assessment, behavioural surveillance systems, and AI-assisted admissions as the three highest-risk application categories. Panel consensus revealed that existing educational data protection frameworks, including GDPR Article 22 and national educational legislation, are substantively inadequate for governing the rights implications of these systems without sector-specific supplementary guidance. The EEAGF provides this critical guidance, offering European education stakeholders operationally feasible mechanisms for implementing AI governance aligned with fundamental European values.

**Keywords:** AI Governance, Human Rights, Education, Council of Europe, EU AI Act, Delphi Study, Automated Assessment, Student Surveillance, Democratic Accountability, ECHR.

## 1. INTRODUCTION

The penetration of artificial intelligence systems into European educational institutions has accelerated substantially over the past decade, transforming practices across the full spectrum of educational transformation from the automated scoring of student essays and the algorithmic recommendation of personalised learning pathways, to the real-time behavioural monitoring of students in examination environments and the data-driven stratification of admissions cohorts. This technological transformation of the educational environment is occurring in a governance space that is simultaneously over, determined by general data protection regulation and under, determined by sector, specific AI governance standards: while the General Data Protection Regulation (GDPR) establishes broadly applicable rules for personal data processing in educational contexts, and the EU Artificial Intelligence Act (EU AI Act, Regulation EU 2024/1689) designates AI systems in education and vocational training as high, risk under Annex III, no operational governance framework yet exists that translates the abstract obligations of these instruments into the specific institutional, pedagogical, and human rights terms that educational AI governance requires (European Parliament, 2024; Williamson et al., 2020).

The Council of Europe's engagement with the human rights dimensions of algorithmic systems provides the most relevant normative foundation for such a framework. The Council of Europe's Recommendation CM/Rec(2020)1 on the human rights impacts of algorithmic systems establishes that member states must

ensure that algorithmic systems deployed in public services including education are designed and governed in accordance with the rights and principles enshrined in the European Convention on Human Rights (ECHR), including Article 6 (right to a fair hearing), Article 8 (right to privacy), Article 10 (freedom of expression and information), Article 14 (prohibition of discrimination), and Protocol 1, Article 2 (right to education) (Council of Europe, 2020). The Recommendation's application to educational AI creates specific obligations: AI, assisted decisions about student assessment, progression, and exclusion implicate Article 6's fair hearing protections; student behavioural monitoring systems implicate Article 8's privacy guarantees; AI, driven content filtering and information access restrictions implicate Article 10; and AI systems that produce differentially accurate or differentially consequential outputs for students from minority ethnic, socio, economic, or disability, defined groups implicate Article 14. In 2024, the Council of Europe further advanced this normative architecture through the adoption of the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law — the first binding international treaty on AI governance which entered into force on 1 July 2025 and imposes direct obligations on signatory states to ensure that AI systems used in public sector contexts, including education, comply with human rights, democratic process, and rule of law standards (Council of Europe, 2024).

The educational context presents a distinctive and particularly demanding human rights governance challenge for several reasons that set it apart from other public-sector AI deployment contexts. First, educational institutions exercise a custodial relationship with students particularly minors that creates heightened duties of care: the power asymmetry between educational institutions and their student populations, combined with the developmental vulnerability of children and adolescents, makes the potential for AI systems to cause harm through misclassification, surveillance, or discriminatory outcome prediction qualitatively more serious than equivalent risks in adult professional contexts (Selwyn, 2019; Perrotta & Selwyn, 2020). Second, education is constitutively connected to democratic formation: the Council of Europe's democratic values mandate treats education not merely as a service delivery context but as the primary institutional site for the formation of democratic citizens individuals capable of critical reasoning, autonomous judgment, and informed participation in democratic processes whose cultivation of these capacities is precisely the faculty that automated educational systems risk suppressing rather than developing (Apple, 2019; Biesta, 2010). Third, the epistemic authority that AI systems carry in educational contexts means that students and parents are likely to treat algorithmically generated grades, risk assessments, and learning recommendations as objective determinations rather than probabilistic inferences, a specific form of epistemic injustice that governance frameworks must explicitly address (Fricker, 2007; Coeckelbergh, 2020).

Existing attempts to formulate AI governance guidelines for education including UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021), the European Commission's Communication on Digital Education Action Plan 2021,2027, and the Organisation for Economic Cooperation and Development's (OECD) Principles on AI as applied to education rovide valuable normative principles but remain at a level of generality that does not translate readily into operationally specific institutional governance obligations (UNESCO, 2021; European Commission, 2021; OECD, 2022). The gap between high-level ethical principles and the concrete institutional governance decisions that educational administrators, technology procurement officers, and regulatory bodies must make is the operational lacuna that this study seeks to fill. The study employs a Delphi Study and Legal, Ethical Consultation methodology to engage a multidisciplinary expert panel in the iterative construction of the European Educational AI Governance Framework (EEAGF), a normative. This normative and operational governance instrument incorporates the human rights, rule-of-law, and democratic values obligations of the relevant Council of Europe and EU regulatory instruments into a set of specific, technically implementable, and institutionally actionable governance guidelines for AI deployment in European educational settings.

The study makes three principal contributions to the intersection of AI governance scholarship, educational law, and democratic theory. First, it produces through rigorous multi, round Delphi methodology the first expert consensus, based taxonomy of high, risk AI scenarios specifically in European educational contexts, providing an empirically grounded risk landscape that goes beyond generic high, risk AI classifications to identify the specific deployment configurationconfigurationsular combinations of AI system type, educational context, and vulnerable population characteristics, at the most acute human rights and democratic values risks. Second, it develops the EEAGF as a four-pillar operational governance framework that integrates human rights law, AI ethics, democratic theory, and technical AI governance standards into a coherent institutional instrument that educational authorities can deploy as a practical governance standard for AI procurement, deployment, and

 

oversight. Third, it identifies and documents the democratic accountability deficit in current European educational AI governance of institutional mechanisms through which the educational community, students, parents, educators, and democratic institutions exercise meaningful oversight and contestation of AI systems that shape educational trajectories as the most consequential gap in the current regulatory landscape, demanding specific structural remediation.

## 2. METHODOLOGY

The study employs a three,round Delphi Study methodology integrated with systematic Legal,Ethical Consultation to achieve two complementary objectives: the empirical identification of expert,consensus high,risk AI scenarios in European educational contexts, and the normative construction of operationally specific governance guidelines through the iterative refinement of expert panel judgments. The Delphi method is epistemologically appropriate for this study's governance framework development objective because the normative and institutional questions involved what human rights obligations apply to specific AI deployment configurations in education, and what governance provisions most effectively operationalise those obligations are genuinely contested and multi,disciplinary, requiring the structured integration of legal, ethical, technical, and pedagogical expertise that the Delphi's iterative consensus,building design is specifically constructed to produce (Linstone & Turoff, 2002; Hasson et al., 2000).

### *2.1. Expert Panel Composition and Recruitment*

A 22,member multi,disciplinary expert panel was recruited through purposive sampling designed to achieve comprehensive disciplinary, national, and institutional diversity across five expertise domains: European human rights law and ECHR adjudication (five panelists, including two former legal advisers to the Council of Europe's Directorate of Human Rights and one academic specialist in Protocol 1, Article 2 jurisprudence); AI ethics and philosophy of technology (four panelists, including specialists in machine ethics, epistemic justice in automated systems, and democratic theory of technology governance); educational law and policy (four panelists, including specialists in German Schulrecht, French educational administrative law, and the legal status of minors in digital educational environments); AI technical governance and algorithmic auditing (four panelists, including specialists in ML fairness measurement, explainable AI for non,technical audiences, and educational data mining); and digital pedagogy and educational psychology (five panelists, including specialists in learning analytics ethics, AI's impact on student autonomy and motivation, and the developmental psychology of adolescent digital surveillance). National representation spanned 12 countries, with Germany, France, the Netherlands, Sweden, and Italy each contributing 2 or more lists, and Austria, Belgium, Denmark, Finland, Ireland, Spain, and Poland each contributing 1. Institutional representation included universities (14 panelists), European supranational bodies, including the Council of Europe Secretariat and the EU Fundamental Rights Agency (3 panelists), and civil society organisations specialising in digital rights and educational advocacy (5 panelists). All panel members confirmed the absence of conflicts of interest with educational technology industry stakeholders before participating.

### *2.2. Delphi Round Structure and Analytical Protocol*

Round 1 employed an open,ended exploratory design in which panelists were invited to identify, in free,text responses, the AI deployment scenarios in educational settings they considered most likely to generate significant human rights violations or democratic values erosion, and to specify the particular rights or democratic principles most implicated in each scenario. Responses were collected through a structured online instrument and subjected to thematic analysis by the research team, producing an initial taxonomy of 47 distinct high,risk scenario types organised into eight thematic clusters: automated assessment and grading; behavioural surveillance and monitoring; AI,assisted admissions and progression decisions; personalised learning pathway recommendation; AI,generated content and information access management; student data aggregation and profile construction; AI,mediated teacher performance evaluation; and algorithmic resource allocation in educational systems.

Round 2 presented panelists with the Round 1 taxonomy and asked them to rate each of the 47 scenario types on two dimensions: the severity of potential human rights impact on a seven,point scale anchored at 'minimal or no rights impact' (1) and 'severe and potentially irreversible rights violation' (7); and the adequacy

 

of current European regulatory frameworks to govern the identified risk on a seven,point scale anchored at 'fully adequate' (1) and 'completely inadequate' (7). Panelists also provided structured justifications for their ratings and were invited to propose draft governance guideline principles for the five scenario types they rated as highest risk. Round 2 responses were analysed to compute mean and median severity and regulatory adequacy ratings for each scenario type, identify scenario types achieving consensus (defined as an interquartile range of two or fewer scale points across the panel's ratings), and compile the draft governance principle proposals for the high,risk priority scenarios.

Round 3 presented panelists with anonymised feedback on the Round 2 rating distributions and the compiled draft governance principles, asking them to revise their ratings in light of the panel's collective judgments (following standard Delphi convergence protocol) and to provide structured assessments of the draft governance principles across four evaluative dimensions: human rights legal grounding (does the principle accurately reflect the applicable ECHR and EU law obligations?), democratic theory alignment (does the principle adequately protect the democratic formation purposes of education?), technical implementability (can the principle be operationalised through available technical and institutional mechanisms?), and institutional proportionality (does the principle impose obligations proportionate to the institutional capacities of diverse European educational systems from large research universities to small primary schools?). The Round 3 responses were integrated through a structured synthesis procedure to produce the final EEAGF governance guideline set, with each guideline accompanied by its legal grounding references, technical implementation specifications, and institutional proportionality tier (distinguishing obligations applicable to all educational institutions from those applicable only to institutions above specified size or technology sophistication thresholds).

### 2.3. Legal, Ethical Consultation

The Delphi process was supplemented by a targeted Legal, Ethical Consultation phase conducted between Rounds 2 and 3, in which the research team engaged in structured consultative dialogues with three European legal bodies: the Council of Europe's Committee of Experts on Internet Intermediaries (MSI, NET), which provided authoritative guidance on the application of CM/Rec(2020)1 to educational AI deployment contexts; the EU Fundamental Rights Agency (FRA), which provided assessment of the EU AI Act's high, risk education system provisions against the EU Charter of Fundamental Rights; and the European Court of Human Rights' (ECtHR) Registry, whose academic liaison function guided the likely application of ECHR Protocol 1, Article 2 education rights jurisprudence to AI, mediated assessment and progression decisions. These consultative inputs were incorporated into the Round 3 synthesis procedure as authoritative legal grounding references for the EEAGF's most legally contested governance guideline provisions.

## 3. RESULTS AND DISCUSSION

### 3.1. High-Risk AI Scenario Taxonomy: Delphi Consensus Findings

The three-round Delphi process produced a consensus-validated taxonomy of high-risk AI scenarios in European educational settings that constitutes the EEAGF's primary empirical foundation. Of the 47 scenario types identified in Round 1, 31 achieved consensus (interquartile range ≤ 2) in Round 3 on both the severity and regulatory adequacy dimensions. The three scenario types achieving the highest mean severity ratings with strong panel consensus are: automated student assessment with consequential impact on educational trajectories (mean severity 6.3/7.0, IQR 1.0); AI, enabled continuous behavioural surveillance in physical and digital learning environments (mean severity 6.1/7.0, IQR 1.5); and AI, assisted admissions and progression decision support with differential accuracy across demographic groups (mean severity 5.9/7.0, IQR 1.5). These three scenario categories are simultaneously rated as most severely under, governed by existing regulatory frameworks, with regulatory adequacy scores of 5.8/7.0, 6.2/7.0, and 5.6/7.0, respectively, indicating strong panel consensus that current European regulatory instruments are substantially inadequate for governing the rights implications of these specific deployment configurations.

The automated student assessment scenario category encompasses a spectrum of AI deployment configurations from AI-assisted marking of short, answer examinations and automated essay scoring systems, to AI-generated student capability assessments based on learning management system interaction data and AI, predicted grade projections used in university admissions, which the panel consistently identifies as

 

generating acute risks across three distinct rights dimensions. First, the accuracy and consistency obligations flowing from ECHR Article 6's fair hearing protections: automated assessment systems that are less accurate for students from specific linguistic, cultural, or socio-economic backgrounds than for the majority student population create differential error rates that constitute a form of structural procedural unfairness in a consequential educational determination. Second, the explainability obligations flowing from GDPR Article 22's right not to be subject to solely automated decisions with significant effects: automated assessment decisions particularly those with direct grade consequences engage Article 22's protections and require the availability of human review, but the panel notes that educational institutions frequently circumvent the 'solely automated' qualification through pro forma human review that does not constitute genuine substantive oversight of the AI, generated output. Third, the epistemic authority risk: automated assessment systems carry a documentary legitimacy with students and experience the AI-generated mark as an objective determination rather than a probabilistic one, which makes the exercise of challenge and review rights systematically less likely in practice than in theory, constituting what the panel terms 'constructive unchecked automation' even in formal human review, available contexts.

The AI-enabled continuous behavioural surveillance scenario category generates the highest regulatory adequacy concern score in the study (6.2/7.0), reflecting the panel's strong consensus that existing European data protection law is most severely inadequate precisely for the governance of ambient, continuous surveillance systems that monitor student behaviour through computer vision, keyboard activity logging, network traffic analysis, and biometric attention tracking. Several such systems have been deployed in European universities and secondary schools during and after the COVID-19 remote learning period, with documented cases in Germany, France, and the Netherlands generating significant controversy regarding their proportionality under ECHR Article 8 and their compatibility with the right to education under Protocol 1, Article 2. The panel's consensus position is that continuous behavioural surveillance in educational contexts generates a specific form of chilling effect on learning documented in the educational psychology literature as surveillance, induced performance anxiety and reduced intellectual risk, taking (Heersmink, 2017; Zuboff, 2019) that is directly injurious to the developmental and democratic formation purposes of education, rendering such systems presumptively disproportionate under the ECHR's necessity and proportionality doctrine regardless of their technical data protection compliance.

### *3.2. The EEAGF: Four Pillar Governance Architecture*

The European Educational AI Governance Framework is organised around four governance pillars that collectively operationalise the human rights, rule of law, and democratic values obligations identified through the Delphi process and the Legal, Ethical Consultation. Each pillar is formulated as a set of specific governance obligations at two institutional compliance tiers: Tier 1 obligations, applicable to all European educational institutions regardless of size or sector, representing the minimum human rights compliance baseline; and Tier 2 obligations, applicable to educational institutions above a threshold of 500 students or deploying AI systems classified as high, risk under EU AI Act Annex III, representing enhanced governance standards for higher, capacity and higher, risk deployment contexts.

Pillar 1, Dignity and Non-Discrimination, encompasses the governance obligations most directly flowing from ECHR Article 14 and the Framework Convention on Artificial Intelligence's prohibition of AI systems that produce differential outcomes based on protected characteristics. The Pillar 1 governance obligations require all educational institutions deploying AI systems with student-facing outcomes to conduct mandatory pre-deployment demographic impact assessments measuring the differential accuracy and differential consequences of AI system outputs across student subgroups defined by gender, ethnicity, socio-economic status, disability, and language background. Tier 1 institutions are required to disclose demographic impact assessment results to students and parents through institutional transparency reports; Tier 2 institutions are additionally required to submit demographic impact assessments to national education regulatory authorities as a precondition for continued AI system deployment authorisation, and to implement technical mitigation measures for any identified differential accuracy gaps exceeding a significance threshold of five percentage points between the most and least accurately served demographic subgroups. The Pillar 1 obligations are grounded in the ECtHR's D.H. and Others v. Czech Republic (2007) jurisprudence establishing that indirect discrimination in educational settings including the use of assessment instruments that produce differentially disadvantageous outcomes for ethnic minority students constitutes a violation of Article 14 in conjunction with

 

Protocol 1, Article 2, applied to the AI context through the Framework Convention on AI's direct extension of this jurisprudential principle to algorithmic systems (Council of Europe, 2024).

Pillar 2, Transparency and Explainability, operationalises the rights of students, parents, and educators to understand the basis of AI-assisted decisions affecting educational trajectories, grounded in GDPR Article 22, EU AI Act Article 13, and the Council of Europe's CM/Rec(2020)1 transparency obligations. The Pillar 2 governance obligations require that all AI-assisted decisions with consequential effects on student assessment, progression, or access to educational resources be accompanied by a contextually appropriate eappropriate explanation in language accessible to the student or their legal guardian, the primary factors contributing to the decision output, the degree of human oversight applied to the AI recommendation, and the procedure for requesting human review. A critical definitional determination contributed by the Legal, Ethical Consultation phase concerns the GDPR Article 22 'solely automated decision' threshold: the FRA's assessment confirmed that many educational AI deployment configurations that formally involve human review do not satisfy the substantive human oversight requirement because the reviewing educator lacks either the technical capacity to meaningfully interrogate the AI recommendation or the institutional authority to deviate from it without documentation of exceptional circumstances a configuration that the EEAGF designates as 'nominal human oversight' and treats as functionally equivalent to solely automated decision, making for Article 22 compliance purposes. Tier 2 obligations include the requirement for educational institutions to document and periodically audit the quality of human oversight in AI-assisted decision-making processes, specifically assessing the rate at which human reviewers substantively deviate from AI recommendations and the institutional conditions associated with such deviations, as a proxy for whether human oversight is substantive or nominal.

Pillar 3, Democratic Accountability, addresses the governance lacuna that the Delphi panel identifies as the most consequential structural gap in current European educational AI governance: the absence of institutional mechanisms through which the educational community, students, parents, educators, and democratic institutions exercise meaningful collective oversight and contestation of AI systems that shape educational trajectories. The democratic accountability deficit has both procedural and substantive dimensions. Procedurally, the procurement and deployment of AI systems in European educational institutions currently occur primarily through opaque administrative processes that are not subject to meaningful democratic deliberation or contestation. Testing is acquired through standard public procurement channels governed by price and technical specification criteria rather than through processes that subject the systems' pedagogical and human rights implications to community scrutiny. Substantively, even where formal mechanisms exist for individual students or parents to request review of AI-assisted decisions, the collective dimension of democratic accountability of the educational community as a whole to assess, challenge, and, if necessary, override the institutional adoption of AI systems that the community judges to be incompatible with educational values or democratic principles is not addressed in current regulatory frameworks.

The Pillar 3 governance obligations therefore in,troduce tw,o novel institutional mechanisms that have no direct precedent in existing educational regulation. The first is the Educational AI Impact Forum (EAIF) a mandatory consultative body at the institutional level for Tier 2 institutions, comprising equal representation from student unions, teacher professional associations, parent organisations, and independent educational ethics specialists, with the authority to review and issue formal recommendations on AI system procurement proposals before deployment authorisation and to initiate periodic reviews of deployed systems against evolving human rights and pedagogical standards. The EAIF design is modelled on the Institutional Review Board (IRB) tradition in research ethics, adapting its basic architecture of community, representative oversight to the AI governance context (Bates et al., 2020). The second mechanism is the AI System Registry a publicly accessible, nationally maintained register of AI systems deployed in educational institutions, containing system type, deployment purpose, provider, demographic impact assessment results, and any EAIF recommendations and institutional responses modelled on the EU AI Act's Article 49 registration obligations for high, risk systems, but extending their scope to include the educational community accountability information that the Act's commercially, oriented registration format does not capture.

Pillar 4, Pedagogical Integrity, addresses the obligation to ensure that AI deployment in education serves and does not distort the fundamental purposes of education as defined by democratic society an obligation that the EEAGF grounds in Protocol 1, Article 2 of the ECHR, which the ECtHR has interpreted as protecting not merely the right of access to educational institutions but the substantive right to benefit from education that

 

fulfils its developmental, social, and democratic formation functions. The panel consensus on the definition of pedagogical integrity is articulated around three core educational purposes that AI deployment must affirmatively support: the development of autonomous critical reasoning and intellectual agency; the cultivation of collaborative social competencies and democratic civic values; and the formation of the epistemic humility and tolerance of uncertainty that characterise both scientific inquiry and democratic deliberation. Pillar 4 governance obligations require educational institutions deploying AI systems with student, facing functions to conduct mandatory Pedagogical Integrity Assessments that evaluate whether and how the AI system's design and deployment logic affects these core educational purposes specifically, whether automated assessment tools create incentive structures that reward pattern recognition and output optimisation over genuine understanding; whether behavioural monitoring systems create surveillance atmospheres that inhibit the intellectual risk, taking and exploratory learning behaviour essential to autonomous cognitive development; and whether AI, driven personalisation systems create filter bubbles that limit students' exposure to the diversity of perspectives that democratic education requires. The Pedagogical Integrity Assessment framework is developed from Biesta's (2010) three educational purposes: qualification, socialisation, and subjectification into operationally specific AI impact assessment criteria through the Delphi panel's expert synthesis.

### *3.3. Automated Assessment Governance: Detailed EEAGF Provisions*

The automated student assessment scenario, identified as the highest risk AI deployment context across all three Delphi dimensions, receives the most detailed and technically specific governance provisions in the EEAGF. The automated assessment governance provisions integrate four distinct regulatory obligations that the Delphi panel and the Legal, Ethical Consultation process identify as individually necessary and jointly sufficient for rights-compliant automated assessment deployment. The first obligation is the Accuracy Equivalence Standard: automated assessment systems must demonstrate, through pre-deployment validation testing on representative student population samples, that their accuracy is equivalent to a percentage-point margin of error across all demographic subgroups defined by the Pillar 1 demographic impact assessment framework, before receiving institutional deployment authorisation. This standard is technically more demanding than the AI Act's general conformity assessment requirements for high-risk systems, which require only that systems achieve 'an appropriate level of accuracy' without specifying demographic equivalence; the EEAGF's more demanding standard reflects the panel's consensus that demographic accuracy equivalence is specifically required by ECHR Article 14 jurisprudence in the educational context.

The second obligation is the Explanation Accessibility Requirement for all automated systems. Assessment outputs communicated to students or used in institutional decision-making must be accompanied by a plain language explanation of the primary factors contributing to the assessment outcome, formulated at a readability level appropriate to the age and educational level of the assessed student, without exposing the technical parameters of the model or creating intellectual property disclosure risks for the system provider. The EEAGF's technical annex provides a model explanation architecture the Layered Assessment Explanation (LAE) format that presents assessment explanations at three levels of detail: a summary level (one to two sentences, accessible to the student directly), an intermediate level (one paragraph, accessible to parents and educators), and a detailed level (structured technical summary, accessible to institutional administrators and, upon request, to independent review bodies).

The third obligation is the Substantive Human Review Guarantee: all automated assessment outputs with direct grade consequences must be subject to substantive human review, as defined in the EEAGF, by an educator with sufficient familiarity with the student's work and sufficient institutional authority to deviate from the AI recommendation without requiring exceptional justification documentation. The assessment is communicated to the student as final. The substantive review requirement explicitly excludes nominal oversight configurations in which a teacher 'approves' an AI-generated grade through a single-click interface without access to the underlying student work, by requiring that review documentation demonstrate that the reviewer has examined the underlying evidence rather than merely endorsed the AI output. The fourth obligation is the Contestation Pathway Requirement: students and their legal guardians must have access to a clearly communicated, institutionally resourced, and time, limited contestation pathway for any automated assessment outcome, with the pathway guaranteeing human review by an educator not involved in the original assessment, response within a specified timeframe not exceeding ten working days, and written justification of any decision to uphold the automated assessment outcome against the student's contestation.

 

### 3.4. Democratic Accountability in Practice: The EAIF Implementation Evidence

The EEAGF's Democratic Accountability pillar, and specifically the Educational AI Impact Forum mechanism, represents the framework's most institutionally innovative and politically consequential contribution. The Delphi panel's consensus on the necessity of a community, representative AI oversight body in educational institutions was strong (78% of panelists rated the mechanism as 'essential' or 'highly important')—still, the panel. Still, the iterative process also generated important design constraints that shaped the EAIF's final governance architecture. Panel members with expertise in German school governance law noted that the EAIF design must be compatible with existing Schulkonferenz (School Conference) structures in German Länder, which already provide for formal community participation in school governance decisions; the EEAGF accommodates this by specifying that Tier 2 institutions may fulfil the EAIF obligation by extending the mandate of an existing governance body that meets the representational requirements, rather than requiring the creation of a new institutional body. Panel members with French educational law expertise noted that the French principle of laïcité (secularism in public education) creates specific constraints on the EAIF's representational architecture in French public schools, where parent organisation representation must respect the strict separation of religious and civic identity that French public school governance requires; the EEAGF incorporates this constraint through a national adaptation provision allowing member states to specify modifications to the EAIF representational requirements necessary to comply with national constitutional principles.

The panel reached its strongest constrong91% agreement across Rounds 2 and 3 combineand 3n the proposition that the most consequential democratic accountability failure in current European educational AI governance is not the absence of individual rights of contestation but the absence of collective democratic deliberation about whether specific AI systems should be deployed in educational contexts at all. Individual challenge rigchallenge GDPR's Article 22 human review right, the EU AI Act's transparency and logging requirements, and the rights of individual students after AI systems have already been deployed; they do not address the community's democratic right to participate in the fundamental institutional decision about whether to introduce an AI system that will reshape the educational environment for all students and educators. The EAIF mechanism is designed precisely to fill this collective democratic agency gap, transforming AI procurement from an administrative decision made by institutional management into a governance decision subject to structured community deliberation a transformation that the panel's democratic theorists ground in Habermas's (1996) discourse theory of democracy, which posits that decisions affecting the shared institutional environment of a community require legitimation through inclusive communicative processes rather than unilateral administrative authority.

### 3.5. The EEAGF and the Council of Europe's Framework Convention: Alignment and Gaps

The alignment between the EEAGF's governance provisions and the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (2024) is substantive across the four governance pillars. The Framework Convention's Article 5 (risk management and safety) corresponds directly to the EEAGF's mandatory demographic impact assessment and pre, deployment validation obligations; its Article 7 (transparency and oversight) provides the treaty, level grounding for the EEAGF's Pillar 2 explainability and human review requirements; its Article 10 (equality and non, discrimination) grounds the Pillar 1 non, discrimination obligations; and its Article 12 (democratic participation) provides the most direct treaty support for the Pillar 3 EAIF mechanism. The Legal, Ethical Consultation with the Council of Europe Secretariat confirmed that the EEAGF's governance provisions are consistent with the Framework Convention's obligations as they apply to educational contexts, and that the Convention's Article 3(2) exemption for national security and defence contexts does not apply to educational AI, ensuring that the Convention's full governance obligations extend to educational settings without qualification.

However, the Delphi panel and the Legal, Ethical Consultation process also identify one significant gap between the EEAGF's governance architecture and the current international treaty landscape: the absence of a binding international enforcement mechanism with jurisdiction over educational AI governance violations specifically. The Framework Convention's compliance monitoring is monitored by the Conference of the Parties established under Article 8, which directs and operates through periodic reporting and dialogue rather than individual complaint mechanisms, meaning that a student whose rights are violated by a non-compliant

 

educational AI system cannot bring a direct claim under the Convention. The ECHR's individual petition procedure before the ECtHR provides the most accessible enforcement pathway for individual educational AI rights violations, but the ECtHR's processing time currently averages five to seven years from application to end, making it an ineffective remedy for the time-sensitive educational rights violations that automated assessment errors and AI-mediated admissions decisions generate. The EEAGF therefore recommends the establishment of national-level Educational AI Ombudsperson functions within existing education or data protection ombudsperson offices, with jurisdiction to receive individual complaints about educational AI rights violations, investigate and mediate complaints within a target timeframe of sixty working days, and issue binding remediation orders requiring institutional correction of identified governance failures.

## 4. CONCLUSION

This study has produced, through rigorous three-round Delphi methodology and systematic Legal, Ethical Consultation with European supranational legal bodies, the first operationally specific, human rights-grounded, and democratically accountable governance framework for AI deployment in European educational settings. The EEAGF's four, pillar architecture architectured Non, Discrimination, Transparency and Explainability, Democratic Accountability, and Pedagogical Integrity translates the abstract obligations of the ECHR, the Council of Europe's Framework Convention on Artificial Intelligence, and the EU AI Act's high, risk education system provisions into specific institutional governance obligations calibrated to the distinctive power relationships, developmental vulnerabilities, and democratic formation purposes of the educational context.

The Delphi process has generated robust expert consensus that the three highest-risk AI deployment categories in European education, automated student assessment, continuous behavioural surveillance, and AI-assisted admissions, are simultaneously the most severely under-regulated by existing regulatory frameworks, confirming the EEAGF's analytical justification and policy urgency. The identification of the democratic accountability deficit the systemic absence of collective community deliberation about AI adoption decisions in educational institutions as the most consequential structural gap in current European educational AI governance represents the study's most significant normative contribution: a finding that reframes AI governance in education from a data protection and individual rights compliance problem into a democratic legitimacy and collective self, determination problem, demanding institutional responses that go beyond individual rights mechanisms toward the creation of genuine community governance authority over the AI, mediated educational environment.

The EEAGF is offered as a practical institutional instrument for European educational authorities, technology procurement bodies, and regulatory agencies seeking to fulfil their obligations under the Council of Europe's Framework Convention on AI and the EU AI Act's high-risk educational system requirements. Its tiered compliance architecture, with Tier 1 minimum baseline obligations applicable to all institutions, from Tier 2 enhanced obligations for larger and higher-risk deployment contexts, ensures that the framework is proportionate across the enormous institutional diversity of European education, from large research universities with sophisticated governance infrastructures to small primary schools with minimal administrative capacity. The EEAGF's ultimate measure of success is not regulatory compliance but educational fidelity: the degree to which AI systems in European educational settings serve, rather than supplant, the irreplaceable human purposes of education in a democratic society.

## REFERENCES

Apple, M. W. (2019). On doing critical policy analysis. Educational Policy, 33(1), 276,287. https://doi.org/10.1177/0895904818807307

Bates, J., Lin, Y. W., & Willson, M. (2020). Governance of AI in public services. Information, Communication and Society, 23(6), 798,811. https://doi.org/10.1080/1369118X.2020.1736513

Biesta, G. (2010). Good education in an age of measurement: Ethics, politics, democracy. Paradigm Publishers.

Coeckelbergh, M. (2020). AI ethics. MIT Press.

Council of Europe. (2020). Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems. Council of Europe.

Council of Europe. (2024). Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225). Council of Europe Treaty Series.

 

European Commission. (2021). Digital Education Action Plan 2021,2027: Resetting education and training for the digital age. European Commission.

European Court of Human Rights. (2007). D.H. and Others v. Czech Republic [GC], Application No. 57325/00, Judgment of 13 November 2007.

European Parliament. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689.

Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press.

Habermas, J. (1996). Between facts and norms: Contributions to a discourse theory of law and democracy (W. Rehg, Trans.). MIT Press.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. Journal of Advanced Nursing, 32(4), 1008,1015. https://doi.org/10.1046/j.1365,2648.2000.01567.x

Heersmink, R. (2017). The internet, cognitive enhancement, and the values of cognition. Minds and Machines, 26(4), 389,407. https://doi.org/10.1007/s11023,016,9404,7

Linstone, H. A., & Turoff, M. (Eds.). (2002). The Delphi method: Techniques and applications. Addison,Wesley.

OECD. (2022). OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots. OECD Publishing. https://doi.org/10.1787/589b283f,en

Perrotta, C., & Selwyn, N. (2020). Deep learning goes to school: Toward a relational understanding of AI in education. Learning, Media and Technology, 45(3), 251,269. https://doi.org/10.1080/17439884.2020.1686012

Selwyn, N. (2019). What is digital sociology? Polity Press.

UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. UNESCO.

Williamson, B., Bayne, S., & Shay, S. (2020). The datafication of teaching in higher education: Critical issues and perspectives. Teaching in Higher Education, 25(4), 351,365. https://doi.org/10.1080/13562517.2020.1748811

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.

 