

INCLUDE-AI and G.I.U.L.I.A.: A Multi-Level European and Territorial Framework for Inclusive, Trustworthy, and Gender-Aware Artificial Intelligence in Education

Matteo Ciaschi
0009-0009-5119-3563

National Research Council (CNR), Rome, Italy
matteo.ciaschi@cnr.it

Daniel Dan
0000-0002-7251-7899

School of Applied Data Science, Modul University Vienna, Austria
daniel.dan@modul.ac.at

Abstract—Artificial Intelligence (AI) is increasingly having a strong impact on education, spanning personalization, automation, and data-driven decision-making. Although promising and disruptive, its implementation also entails structural risks, including algorithmic bias, unequal access to digital resources, and a lack of governance frameworks. OECD (2023) reports that over 60% of schools in high-income countries use AI-driven educational tools, while a UNESCO (2023) survey of 450 institutions identifies the absence of national regulations as a primary reason for the lack of institutional guidance. This results in uneven and fragmented adoption of AI tools and policies.

This paper provides a dual contribution. First, it introduces INCLUDE-AI (INCLUsive Artificial Intelligence for Democratic and Accessible Education), a Europe-oriented multi-level governance framework designed to operationalize inclusive and trustworthy AI across educational systems. Second, it presents G.I.U.L.I.A. (Gender-Inclusive Use of Learning Intelligence & AI), a territorial implementation model developed in the Bologna educational ecosystem as a context-sensitive extension of the broader framework. The study adopts a cross-national and participatory design perspective involving Italy and Austria, combining ethnographic observation with applied socio-technical analysis. To demonstrate how European-level strategies can be translated into localized, gender-aware educational practices, the paper proposes a socio-technical architecture, a governance taxonomy, and a set of measurable indicators. Evidence from an exploratory mixed-methods pilot involving 312 students and 78 educators indicates a persistent gender-based gap in self-assessed AI literacy and conditional educator trust, suggesting that inclusion treated as a design principle rather than a corrective measure is a precondition for equitable AI adoption in education.

Index Terms—Artificial Intelligence in Education, Inclusive Education, AI Ethics, Digital Governance, Learning Analytics, Socio-Technical Systems, Educational Innovation, Complex Systems Management

I. INTRODUCTION

ATTEMPTS to introduce Artificial Intelligence into education began in 1960 with PLATO (Programmed Logic for Automatic Teaching Operations), but the major breakthrough came with ChatGPT in 2022. The integration of Natural Language Processing is among today's major transformations

in knowledge distribution, with formal and informal learning now using adaptive environments, predictive analytics, and automated assessment. This fast diffusion of generative AI poses significant ethical and regulatory challenges, including data management, algorithmic bias, and systemic inequalities [1]. On the opportunity side, personalized learning has been shown to produce statistically significant gains in student performance, particularly in mathematics [2].

However, notwithstanding the technological progress, critical evidence challenges this promise, pointing to systemic risks. Because algorithmic systems use data at their core, implement assumptions, and are applied in institutional contexts, they are not unbiased. Commercial facial analysis systems based on machine learning can exhibit substantial disparities in accuracy across demographic groups [3]. This points out to biases that are automatically embedded at the core of AI systems through data. In educational settings, these risks are further amplified by structural inequalities in access to digital infrastructures. The World Bank (2021) [4] estimates that a majority of students in many countries lack access to digital devices or internet connectivity. This results in a strong, uneven participation and a broadened lack of opportunities for them.

At the same time, governance has not kept pace with technological development. The EU has responded with major initiatives—the AI Act (2024), which classifies educational applications as high-risk given their impact on fundamental rights, and the Digital Education Action Plan (2021–2027) [5], which sets out a vision for inclusive, accessible digital education. Yet the translation of these frameworks into operational practice remains uneven and fragmented, particularly at the regional level: schools frequently lack the technical infrastructure, competencies, and governance structures needed to engage critically with AI, while the opacity of many systems further reduces educators' trust and capacity to evaluate algorithmic recommendations [6].

This paper contributes to the emerging field of Artificial Intelligence in education in three main ways. First, it devel-

ops INCLUDE-AI, a multi-level governance framework that integrates infrastructure, data governance, algorithmic design, and pedagogy within a unified socio-technical perspective. Second, it introduces G.I.U.L.I.A. as a territorially grounded implementation model, offering an applied, context-sensitive extension of European-level strategies, with a specific focus on gender inclusion. Third, the paper advances the operationalization of inclusion-by-design by proposing a set of measurable dimensions and indicators to evaluate fairness, accessibility, and participation in AI-enabled educational systems.

Unlike existing frameworks, this study explicitly operationalizes inclusion through measurable indicators and integrates gender as a structural dimension of AI governance.

The remainder of the paper is organized as follows. Section II reviews related work and positions INCLUDE-AI against existing frameworks. Section III presents the INCLUDE-AI architecture, governance taxonomy, and indicators. Section IV introduces G.I.U.L.I.A. and maps it onto INCLUDE-AI. Section V describes the mixed-methods design and empirical findings. Sections VI and VII discuss implications and limitations, and Section VIII concludes.

II. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

The literature on Artificial Intelligence in education increasingly recognizes both its potential and its limitations. While AI enables new forms of personalization, allowing systems to adapt to individual learning paths, several studies highlight that these systems are often built on data structures that reflect existing social and institutional biases. Students from disadvantaged backgrounds use less structured and less complete data due to factors such as unstable internet access, lower digital literacy, and limited engagement with digital learning platforms [7]. A recent meta-systematic review highlights that the field of AI in education remains dominated by performance-oriented studies, while ethical and governance dimensions are still underdeveloped [8].

From an ethical perspective, the debate has increasingly focused on the principles of transparency, fairness, and accountability. A comprehensive framework for ethical AI has emphasized the importance of human-centered design and responsible innovation [9]. However, ethical principles alone are insufficient if they are not embedded within the technical and organizational structures of AI systems [10]. This observation is particularly relevant in educational contexts, where the interaction between technology and pedagogy is deeply intertwined.

AI literacy is widely seen as a critical enabler of meaningful engagement with AI systems [11], encompassing not only technical skills but also the capacity to understand, critically evaluate, and ethically reflect on algorithmic systems. Crucially, this competence is unevenly distributed, with disparities associated with gender, socio-economic background, and geographic context [12]. Where students and educators cannot critically assess the outputs of adaptive systems, existing

inequalities are reproduced rather than reduced, making inclusive and context-sensitive AI-literacy programs a precondition for equitable integration of AI in education [13].

In parallel, research on digital infrastructures highlights the importance of robust, interoperable systems to support AI applications. In education, interoperable digital infrastructures and shared platforms enable data to flow across systems and support large-scale storage, processing, and analysis [14]. Yet, the relationship between infrastructure and pedagogy remains underexplored, with most studies treating these domains separately. This fragmentation limits the ability to develop integrated solutions that address the multifaceted nature of AI in education. For example, recent FedCSIS contributions present capable AI-driven educational tools—an LLM-based feedback generator for an intelligent tutoring system [15] and an AI-supported generator of programming exercises [16]—yet such tools are typically designed and evaluated as isolated artifacts, without an explicit data-governance, fairness, or inclusion layer wrapped around them. A device-provisioning decision taken independently of teacher training and data governance can likewise leave a well-designed adaptive tool unused where connectivity or competence is lacking. INCLUDE-AI is positioned precisely at this gap: rather than proposing another tool or another set of abstract principles, it specifies how infrastructure, data, algorithms, and pedagogy must be governed jointly, and pairs each dimension with a measurable indicator—an emphasis that the compared frameworks [6], [9], [13] address only partially (Table I).

The theoretical foundation of this work builds upon the concept of socio-technical systems, which underlines the interdependence between technological components and social practices. Within this perspective, educational institutions are systems in which AI technologies, data governance mechanisms, pedagogical practices, and institutional values are intertwined. Taken together, these perspectives suggest that AI in education cannot be adequately addressed through isolated technical or ethical approaches. It requires an integrated socio-technical framework in which governance, infrastructure, and pedagogy are treated as interdependent dimensions.

Table I provides a dimension-based comparison between INCLUDE-AI and the most relevant existing frameworks in this area. Specifically, the table contrasts INCLUDE-AI with Floridi et al.'s general ethical AI framework [9], Holmes et al.'s education-specific treatment of AI ethics [6], and UNESCO's AI Competency Framework for Teachers [13]. This comparison is not intended to equate these frameworks methodologically, but rather to examine whether they explicitly address core dimensions relevant to equitable AI integration in education, namely infrastructure, data governance, algorithmic fairness, pedagogy, gender inclusion, and measurable indicators.

III. THE INCLUDE-AI FRAMEWORK

The central contribution of this paper is a multilevel framework that operationalizes inclusive and ethical AI within

TABLE I
COMPARISON OF INCLUDE-AI WITH EXISTING FRAMEWORKS

Dimension	INCLUDE-AI	Floridi	Holmes	UNESCO
Infrastructure	✓	–	–	–
Data	✓	✓	✓	Partial
Governance	✓	✓	✓	Partial
Algorithmic	✓	–	✓	✓
Fairness	✓	–	✓	✓
Pedagogy	✓	–	Partial	Partial
Gender Inclusion	✓	–	–	Partial
Measurable Indicators	✓	–	–	Partial

educational ecosystems, moving beyond abstract ethical discussion toward a design-oriented approach to AI governance at the European level. Its premise is that inclusion and trustworthiness cannot be achieved through isolated interventions but require systemic integration across multiple layers of the educational ecosystem.

At the infrastructural level, the framework emphasizes the importance of robust, scalable digital environments that support data-intensive AI applications. This includes cloud-based architectures, interoperable data systems, and secure network infrastructures aligned with European strategies on digital sovereignty. Infrastructure is not treated as a neutral enabler, but as a determinant of access and participation, directly influencing the distribution of opportunities within educational systems.

The data governance dimension addresses issues related to data quality, bias mitigation, and regulatory compliance. Particular attention is given to the alignment with European legal frameworks, including the General Data Protection Regulation and the emerging AI Act. The framework aims to adopt data models that ensure transparency, accountability, and ethical principles.

At the algorithmic level, INCLUDE-AI supports the development of explainable and fairness-aware models. The opacity of many AI systems has been identified as a major barrier to trust and adoption, particularly among educators, for whom understandability and explainability are central to confidence in AI-enabled educational tools [17]. By promoting explainability and auditability, the framework seeks to enhance users' capacity to critically engage with algorithmic outputs.

The pedagogical level integrates principles of inclusive education into the design and implementation of AI systems. This involves adapting technologies to diverse learning needs and rethinking educators' roles as active agents in building AI-enhanced learning environments. This level also addresses the risk of cognitive offloading. In such cases, students may delegate critical thinking and analytical reasoning to AI systems, thereby undermining the development of their cognitive skills. The framework advocates pedagogical approaches that position AI as a tool to augment, not replace, human cognitive engagement, ensuring that learners retain agency over their learning processes.

The four levels of the INCLUDE-AI framework are not

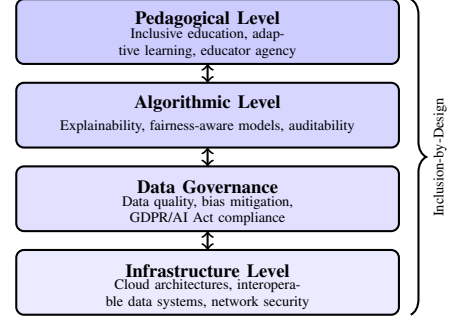


Fig. 1. INCLUDE-AI Framework: Four-layer socio-technical architecture with bidirectional interdependencies.

conceived as independent layers, but as interdependent dimensions that interact dynamically. For example, the quality and representativeness of training data (data governance) directly affect the fairness of algorithmic outputs (at the algorithmic level). This, in turn, shapes the pedagogical experiences of diverse learner populations (pedagogical level). Similarly, the availability and robustness of digital infrastructure (at the infrastructure level) constrain the range of AI applications that can be deployed in a given educational context. Figure 1 illustrates the four-layer architecture. The vertical ordering should be read as a socio-technical stratification (from material substrate to human practice) rather than as a strict service-dependency stack: the bidirectional arrows denote mutual constraint between adjacent strata, not a one-directional, top-down dependency. The model is thus better understood as a constrained interaction graph rendered in layered form for readability, and the layers are not assumed to be independently substitutable.

The governance taxonomy maps core dimensions across levels, actors, and mechanisms. This structure facilitates the transition from normative principles to actionable governance practices, supporting measurable, context-sensitive implementation. Table II presents this taxonomy.

A key innovation of INCLUDE-AI is the development of measurable indicators. This allows for the empirical assessment of inclusion, addressing a major limitation of existing frameworks that remain primarily normative. Table III presents the core inclusion-by-design indicators that guide the framework's operationalization.

IV. G.I.U.L.I.A.: TERRITORIAL IMPLEMENTATION AND GENDER-INCLUSIVE GOVERNANCE

While INCLUDE-AI provides a European-level architecture, its effective implementation requires translation into context-specific practices. The G.I.U.L.I.A. framework represents this translation, embodying a territorial¹ and applied dimension developed within the Bologna (Emilia-Romagna Region) educational ecosystem.

¹We use *territorial* in the sense of European cohesion and regional policy, i.e. a sub-national (regional/local) level of implementation; the term is used interchangeably with *regional* throughout.

TABLE II
INCLUDE-AI GOVERNANCE TAXONOMY

Dimension	Level	Actors	Mechanisms	Indicators
Infrastructure	Macro/Meso	Governments, Institutions	Cloud systems, interoperability	Access rate, system availability
Data Governance	Macro/Meso	Regulators, Schools	GDPR compliance, data stewardship	Data transparency, bias audits
Algorithmic	Meso/Micro	Developers, EdTech providers	Explainability, fairness-aware models	Perceived fairness, auditability
Pedagogical	Micro/Meso	Teachers, Students	Co-design, inclusive curricula	Engagement, inclusion index
Governance	Cross-level	Policymakers, Stakeholders	Multi-stakeholder accountability	Trust level, participation rate

TABLE III
INCLUDE-AI: INCLUSION-BY-DESIGN INDICATORS

Dimension	Indicator	Measurement Approach
Access	Availability of AI-enabled learning tools	Infrastructure audit; student survey
Fairness	Perceived algorithmic fairness	Likert-scale educator/student survey
Literacy	Gender-disaggregated AI literacy levels	Self-assessment questionnaire
Transparency	Transparency perception among educators	Semi-structured interviews; survey
Participation	Stakeholder involvement in AI governance	Co-design workshop attendance; qualitative analysis

The conceptual contribution of this project lies in its explicit focus on gender inclusion as a structural component of AI governance. Existing approaches often treat gender disparities as secondary effects of technological systems; by contrast, G.I.U.L.I.A. frames gender as a central dimension influencing access to resources, participation in decision-making, and interaction with AI systems.

In operational terms, “gender-aware” refers to four concrete mechanisms enacted in the pilot rather than to a general orientation. First, *gender-disaggregated baseline measurement*: AI-literacy and confidence were measured separately for male and female students (Table VIb) and used as a design input, so that observed gaps shaped subsequent activities rather than merely being reported. Second, *balanced participation*: co-design workshops applied a participation target ensuring that female students and teachers were represented in proportion to, or above, their share in each setting, and facilitation explicitly elicited their input. Third, *stereotype review of materials*: worked examples, prompts, and role models used in the AI-supported activities were screened to avoid gendered stereotypes (e.g. default male framing of technical roles) and revised accordingly. Fourth, *gender-sensitive monitoring*: the inclusion indicators in Table III retain gender as a disaggregation axis, so equity effects are tracked over time rather than assumed. These mechanisms are deliberately lightweight and procedural, which is what makes them transferable to institutions with

TABLE IV
MAPPING G.I.U.L.I.A. LEVELS OF ACTION ONTO INCLUDE-AI LAYERS

G.I.U.L.I.A. level	INCLUDE-AI layer(s) enacted	Concrete mechanism in pilot
Individual	Pedagogical, Algorithmic (end-user view)	Critical AI-literacy modules; reading and contesting algorithmic outputs
Pedagogical	Pedagogical, Data Governance	Inclusive curricula; consent-aware data use; bias review of teaching materials
Organizational	Infrastructure, Data Governance, Governance (cross-level)	Procurement and interoperability decisions; accountability roles; transparency protocols

limited resources.

The framework adopts a multi-level perspective. At the individual level, it promotes the development of critical AI literacy among educators and students, enabling them to understand and question algorithmic processes. At the pedagogical level, it supports the design of inclusive learning environments that account for diverse cognitive and social dynamics. At the organizational level, it introduces governance mechanisms to enhance transparency, accountability, and participation in AI adoption decision-making.

It is important to clarify the relationship between the two frameworks, since their organizing axes differ by design. INCLUDE-AI is structured by socio-technical *layers* (infrastructure, data governance, algorithmic, pedagogical), i.e. *what* must be governed; G.I.U.L.I.A. is structured by *levels of action* (individual, pedagogical, organizational), i.e. *who* acts and at which scale. G.I.U.L.I.A.’s levels are therefore not an alternative taxonomy but the operational projection of INCLUDE-AI’s layers onto the actors that enact them in a given territory: each level of action takes responsibility for one or more layers and translates the corresponding requirements into concrete practices. Table IV makes this mapping explicit. This separation of concerns is what allows a single European-level architecture to be implemented by heterogeneous local actors without prescribing a uniform technical stack.

Bologna offers fertile ground for this experimentation: a longstanding tradition of educational innovation, major research institutions, and a dense network of schools and civic associations support multi-stakeholder collaboration. The pilot activities involved secondary schools and vocational training institutions across the Bologna and Ravenna metropolitan

areas, spanning urban and semi-urban settings to capture variation in institutional capacity and digital readiness.

The pilot implementation followed a phased approach. In the first phase, baseline assessments were conducted to map existing levels of AI literacy, digital infrastructure availability, and institutional governance practices across participating schools. In the second phase, co-design workshops brought together educators, students, school administrators, and technology providers to collaboratively develop inclusive pedagogical strategies and governance protocols for AI adoption. In the third phase, the proposed strategies were implemented in classroom settings and monitored through surveys, interviews, and ethnographic observation. This iterative cycle of assessment, co-design, and implementation embodies the research-action methodology underpinning the project, ensuring the framework remains grounded in real-world constraints and responsive to stakeholder feedback.

As summarized in the crosswalk of Table IV, the model translates INCLUDE-AI’s European-level governance into the Bologna ecosystem through three levels of action (individual, pedagogical, organizational), with gender inclusion as a cross-cutting dimension spanning all three.

V. METHODOLOGY

The research adopts a mixed-method and participatory approach, combining qualitative and quantitative techniques within a comparative framework. The choice of Italy and Austria as case studies reflects their differing levels of digital maturity and policy development, thus providing a meaningful basis for cross-contextual analysis.

Data collection integrates multiple sources and operates across three analytical levels—macro, meso, and micro—ensuring both robustness and real-world (contextual) validity.

A. Macro-Level Analysis

At the macro-level, consolidated international datasets contextualize AI in education globally. OECD data (2021; 2024) [14], [18] indicate that 60–65% of AI-based educational systems rely on behavioral and interaction data, with adaptive technologies improving outcomes by up to 30% in well-resourced settings. UNESCO reports (2023–2025) [19], [20] reveal a governance gap: fewer than 10% of countries have comprehensive AI-in-education regulation and fewer than 35% have formal governance guidelines, while some 2.6 billion people remain offline. World Bank data (2021) [4] confirm that around 40% of students lack reliable access to digital learning tools, with markedly lower rates in low-income contexts.

B. Meso-Level Analysis

At the meso-level, the research develops a comparative analysis between Italy and Austria, integrating data from the European Commission’s Digital Education Action Plan (2022–2025) [5] and national statistics. Table V presents a comparative assessment of digital readiness between the two countries.

TABLE V
COMPARATIVE DIGITAL READINESS: ITALY VS. AUSTRIA

Indicator	Austria	Italy
Schools with digital platforms (%)	85	68
Teacher digital training (%)	64	42
Female STEM participation (%)	34	27
AI experimentation level	Medium-High	Medium

Here, the *AI experimentation level* is an ordinal summary of three observable signals: the presence of a national AI-in-education strategy or guidelines, the share of schools running structured AI pilots (as opposed to ad-hoc individual use), and the availability of dedicated teacher-training provision. A *medium* level denotes emerging, largely bottom-up experimentation without a consolidated national strategy, whereas *medium-high* denotes a coordinated national strategy with institutionalized pilots. By this measure both countries exhibit medium to medium-high levels of AI experimentation; the gaps in digital platforms (85% vs. 68%) and teacher digital training (64% vs. 42%) reflect Austria’s more structured national investment in educational technology and teacher competence.

Gender disparities in STEM participation persist in both contexts (34% in Austria, 27% in Italy). This is relevant to AI education because lower female STEM participation correlates with reduced exposure to computational thinking, which affects self-assessed AI literacy and confidence—providing empirical grounding for the gender-aware design of G.I.U.L.I.A.

C. Micro-Level Analysis and Empirical Findings

At the micro-level, the study incorporates primary data collected through the G.I.U.L.I.A. pilot implementation within the Bologna educational ecosystem. The empirical component includes surveys administered to students (N=312) and teachers (N=78), ethnographic observation in educational settings, semi-structured interviews, and co-design workshops involving local stakeholders.

The qualitative component included 4 co-design workshops involving educators, students, and institutional stakeholders (N=46 participants total), as well as 12 semi-structured interviews with teachers, school leaders, and policy actors. The co-design workshops followed a participatory design protocol structured in three phases: (i) problem framing, (ii) collaborative ideation, and (iii) validation of governance and pedagogical solutions. Each workshop lasted approximately two hours and used scenario- and persona-based prompts (e.g. an adaptive tutoring tool producing an opaque recommendation) to surface concrete trust and fairness concerns; their outputs fed directly into draft governance protocols and the inclusion indicators. The student and teacher surveys used closed Likert-scale items adapted from established AI-literacy and technology-acceptance instruments, complemented by demographic items enabling the gender-disaggregated analysis. Semi-structured interviews were conducted using a protocol focusing on trust, perceived risks, usability, and gover-

nance expectations related to AI systems. All qualitative data were analyzed using thematic coding, following an inductive-deductive approach.

The quantitative findings are summarized in Table VI.

Gender-disaggregated analysis reveals significant differences in self-perceived AI competence. While 26% of male students report high confidence in using AI systems, only 14% of female students do so. Conversely, a higher proportion of female students (42%) report low confidence compared to their male counterparts (26%). Within this exploratory sample, these findings indicate a gender-based confidence gap, consistent with existing literature on digital self-efficacy and STEM participation. As the data are self-reported and the sample is local, the gap should be read as a measured association rather than as a generalizable effect.

Thematic analysis of the qualitative data identified three dominant dimensions that reinforce and contextualize the quantitative findings:

Trust. Educators consistently expressed that trust in AI systems is contingent on understanding how decisions are produced. As one secondary school teacher stated: “I would use AI more if I understood how decisions are made.” This conditional trust aligns with the quantitative finding that 61% of educators report trust dependent on improved transparency.

Agency. Participants emphasized the importance of maintaining human control and critical engagement with AI systems. A workshop participant noted: “Students should not just use AI, but understand and question it.” This theme underscores the need for AI literacy as a prerequisite for meaningful participation.

Opacity. The perceived lack of transparency in AI systems emerged as a recurring concern across both interviews and workshops. As one teacher expressed: “AI feels like a black box—we don’t know what happens behind the interface.” This perception is directly connected to the 63% of educators who identified a lack of transparency as a primary risk.

These themes underscore the centrality of transparency and user empowerment, and directly informed the design of the INCLUDE-AI governance taxonomy and inclusion indicators; the participatory, co-design dimension also strengthens the study’s real-world (contextual) validity.

Overall, the integrated dataset reveals three structural tensions that underpin the rationale of this study: (i) a performance–equity gap, whereby the benefits of AI are unevenly distributed; (ii) an adoption–governance gap, reflecting the misalignment between technological diffusion and regulatory frameworks; and (iii) an access–capability gap, between infrastructure availability and actual competencies and trust. Exploratory in scale, these insights nonetheless directly informed the design of the dual INCLUDE-AI and G.I.U.L.I.A. frameworks as multi-level, inclusion-oriented governance models.

VI. DISCUSSION

To keep claims of different kinds distinct, we read the results in three registers, as suggested in the review process: empirical

findings from the pilot, which are exploratory and local; the conceptual proposals embodied in the two frameworks; and policy recommendations, which follow analytically from the analysis and remain to be tested. Read in this way, the exploratory findings are consistent with—rather than proof of—the argument that integrating governance, inclusion, and socio-technical design enables more transparent and context-aware adoption of AI in education. At the conceptual level, the analysis highlights a recurring limitation of current approaches to AI in education: the tendency to treat ethical considerations as external constraints rather than integral components of system design.

Theoretically, the work extends socio-technical systems theory by treating inclusion and gender as structural rather than incidental dimensions; practically, the Bologna case suggests that regional ecosystems can serve as laboratories for validating and refining European strategies, underlining the value of place-based innovation in responsible AI.

The pilot data give each of the three tensions a concrete face. The performance–equity gap appears in the contrast between students’ positive perceptions of AI (74%) and the gender disparities in self-assessed literacy: recognition of AI’s benefits is not matched by an even capacity to engage with it critically. The adoption–governance gap shows in conditional educator trust (61%), indicating that current deployment does not yet meet stakeholders’ governance expectations. The access–capability gap surfaces in the Italy–Austria comparison, where differences in infrastructure and teacher training track different levels of AI readiness.

These tensions are not unique to the Italian or Austrian context but reflect broader challenges in the European educational landscape. The INCLUDE-AI framework, by addressing these tensions through an integrated socio-technical approach, provides a model that is potentially transferable to other national and regional contexts, subject to appropriate contextual adaptation.

A legitimate objection is that many schools and institutions lack the technical infrastructure, staff competencies, and governance structures needed to adopt AI tools at all, which could place a framework of this kind out of reach. The design responds to this in three ways. First, the framework is meant to be entered incrementally rather than wholesale: an institution with limited resources can begin at the data-governance and pedagogical layers—which are largely organizational and procedural—before committing to infrastructure investment. Second, the governance taxonomy (Table II) doubles as a readiness self-assessment, letting an institution locate its current capacity per dimension and prioritize the smallest viable next step rather than a full programme. Third, by assigning the infrastructure and interoperability burden primarily to the macro/meso levels (regional authorities, shared platforms), the model deliberately shifts cost away from individual under-resourced schools toward pooled provision. The gender-aware mechanisms described in Section IV were chosen on the same principle: they are procedural and low-cost rather than technology-dependent. We do not claim this removes the ca-

TABLE VI

G.I.U.L.I.A. PILOT SURVEY RESULTS: (A) STUDENT PERCEPTIONS AND SELF-ASSESSED AI LITERACY (N=312); (B) SELF-ASSESSED AI LITERACY BY GENDER (N=312; MALE=168, FEMALE=144); (C) EDUCATOR TRUST AND PERCEIVED RISKS (N=78).

Dimension	Category	%				Dimension	Category	%
Overall perception	Positive	74	AI literacy	Male %	Female %	Trust level	High	22
	Neutral	21					Conditional	61
	Negative	5					Low	17
AI literacy (self)	High	18				Perceived risks	Data privacy	71
	Medium	46	High	26	14		Lack of transparency	63
	Low	36	Medium	48	44		Algorithmic bias	48
(a) Students			Low	26	42	(c) Educators		
			(b) Literacy by gender					

capacity gap—closing it ultimately requires sustained investment and policy support—but the framework is structured so that partial, resource-constrained adoption is still meaningful.

Within the EDUC-AI-TION perspective, conversational AI systems (chatbots and generative assistants) are a critical application domain, raising specific challenges of (i) epistemic trust in AI-generated responses, (ii) bias amplification, as language models may reinforce stereotypes, and (iii) student dependency and cognitive offloading. By integrating explainability, co-design, and inclusion indicators, INCLUDE-AI offers a governance structure for deploying such tools transparently and responsibly.

A. Implications for Practitioners

The findings carry concrete implications for three groups of actors. For educators, the results underscore AI literacy not merely as a technical skill but as a critical competence; teacher-training programs should integrate modules on algorithmic transparency, data governance, and ethical AI use, so that educators can act as informed mediators between AI systems and learners. The finding that only 22% of teachers report high trust suggests that current professional development does not yet address the governance and transparency concerns that shape adoption decisions. For school administrators, the governance taxonomy (Table II) functions as a readiness tool: by mapping dimensions, actors, and mechanisms, leaders can locate gaps in institutional capacity, while the inclusion indicators (Table III) support monitoring of equity effects over time.

For EdTech developers, the emphasis on explainability and fairness-aware design should inform tools that expose transparent decision processes and accommodate diverse learners. The recurrent theme of opacity indicates that current educational AI systems fall short of users' transparency expectations, implying both an ethical and a market imperative for more interpretable, auditable tools.

VII. LIMITATIONS AND FUTURE WORK

This study has several limitations. First, the empirical component is limited in scale: the G.I.U.L.I.A. pilot is confined to the Bologna/Ravenna context, which favours real-world (contextual) validity and depth but constrains generalizability, so transferability to other—especially non-European—contexts needs further validation. Second, the research-action design

prioritizes iterative development over large-scale quantitative testing, and the absence of longitudinal data limits claims about long-term impact. Third, the operationalization of inclusion remains partially exploratory: the indicators require psychometric refinement and cross-institutional comparison. Fourth, explainability and auditability are moving targets as AI systems evolve, and the alignment between European policy, local practice, and infrastructure is an ongoing institutional process beyond the scope of this study.

These limitations define a research agenda. The priorities are: (i) large-scale, cross-country validation beyond a single region (e.g. Nordic, Southern-European, and non-European contexts); (ii) longitudinal designs tracking AI literacy, trust, and pedagogical practice over several years to test whether the governance mechanisms produce sustained equity gains; (iii) refinement of the inclusion indicators (Table III) into validated composite indices, and extension of the four gender-aware mechanisms—as a reusable template—to other equity axes such as socio-economic status, migrant or cultural background, and disability, ultimately through intersectional rather than single-axis analysis; and (iv) deeper study of human-AI interaction (trust, agency, decision-making), where AI shapes cognitive and emotional dimensions of learning in ways not yet well understood [21], alongside practical compliance tools for aligning institutional practice with the European AI Act.

VIII. CONCLUSION

This paper reframes Artificial Intelligence in education not as a purely technological innovation, but as a socio-technical transformation that reshapes access, participation, and governance within educational ecosystems. Because AI systems are embedded in structures of power, data, and institutional design, they call for governance models that are correspondingly multi-level and inclusive.

The dual framework proposed—INCLUDE-AI at the European level and G.I.U.L.I.A. at the regional level—addresses a persistent gap between normative principles and implementation by integrating infrastructure, data governance, algorithmic transparency, and pedagogy into a single architecture and pairing each dimension with a measurable indicator. The exploratory evidence indicates that AI can enhance learning outcomes while also amplifying inequalities unless inclusion is embedded by design rather than added as a corrective measure;

inclusion-by-design is thus the paper's central proposition. The G.I.U.L.I.A. pilot further illustrates the value of regional ecosystems as settings in which European strategies can be translated, tested, and refined, while the explicit treatment of gender as a structural dimension moves the discussion beyond generic notions of fairness toward context-sensitive, measurable inclusion.

The work aligns with emerging research positioning AI in education as a force shaping not only performance but also cognition, agency, and emotional engagement [21]. From a policy standpoint, the governance taxonomy (Table II) offers institutional leaders a structured mapping to align practice with evolving European regulation, including the AI Act; from a theoretical standpoint, the paper embeds inclusion, governance, and gender into socio-technical systems theory. The same architecture extends to conversational AI tools—chatbots and generative assistants—whose governance challenges (epistemic trust, bias amplification, and cognitive offloading) are increasingly salient within the EDUC-AI-TION community.

Ultimately, the trajectory of AI in education will depend less on technological sophistication alone than on the capacity to design systems that are transparent, accountable, and equitable—an outcome that, as recent work stresses, requires multi-stakeholder alignment among educators, students, and institutions [22] and integrated approaches combining ethical, technical, and institutional dimensions [23]. Rather than a finished paradigm, INCLUDE-AI and G.I.U.L.I.A. are offered as an operational starting point for the empirical validation and policy-oriented implementation of inclusive AI in education.

ACKNOWLEDGMENTS

This work was encouraged by Fondazione Del Monte di Bologna e Ravenna. The authors also acknowledge the contribution of local educational institutions and stakeholders involved in the G.I.U.L.I.A. pilot activities. Claude Sonnet (Anthropic) was used as an editorial assistant to enhance readability and ensure the absence of typographical errors in the final manuscript. All intellectual content, analysis, and argumentation remain the sole responsibility of the authors.

REFERENCES

- [1] I. M. García-López and L. Trujillo-Lián, "Ethical and regulatory challenges of generative AI in education: A systematic review," *Frontiers in Education*, vol. 10, 2025, doi: 10.3389/educ.2025.1565938.
- [2] J. F. Pane, E. D. Steiner, M. D. Baird, and L. S. Hamilton, "Informing progress: Insights on personalized learning implementation and effects," RAND Corporation, Tech. Rep. RR-2042-BMGF, 2017, doi: 10.7249/RR2042.
- [3] J. Buolamwini and T. Gebru, "Gender shades: Intersectional accuracy disparities in commercial gender classification," in *Proc. 1st Conf. Fairness, Accountability and Transparency (FAT*)*, ser. Proceedings of Machine Learning Research, vol. 81, 2018, pp. 77–91.
- [4] World Bank, "Remote learning during COVID-19: Lessons from today, principles for tomorrow," World Bank Group, Tech. Rep., 2021.
- [5] European Commission, *Digital Education Action Plan (2021–2027)*. European Union, 2020, cOM(2020) 624 final.
- [6] W. Holmes, K. Porayska-Pomsta, and S. Nemorin, "The ethics of AI in education," in *Handbook of Artificial Intelligence in Education*, 1st ed., B. du Boulay, A. Mitrovic, and K. Yacef, Eds. Cheltenham, UK: Edward Elgar Publishing, 2023, pp. 571–604, doi: 10.4337/9781800375413.00038.
- [7] R. F. Kizilcec and H. Lee, "Algorithmic fairness in education," in *The Ethics of Artificial Intelligence in Education*, W. Holmes and K. Porayska-Pomsta, Eds. Routledge, 2022, doi: 10.4324/9780429329067-10.
- [8] M. Bond, H. Khosravi, M. De Laat *et al.*, "A meta-systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour," *International Journal of Educational Technology in Higher Education*, vol. 21, no. 4, 2024, doi: 10.1186/s41239-023-00436-z.
- [9] L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, B. Schafer, P. Valcke, and E. Vayena, "AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations," *Minds and Machines*, vol. 28, no. 4, pp. 689–707, 2018, doi: 10.1007/s11023-018-9482-5.
- [10] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, 2016, doi: 10.1177/2053951716679679.
- [11] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao, "Conceptualizing AI literacy: An exploratory review," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100041, 2021, doi: 10.1016/j.caeai.2021.100041.
- [12] K. J. Davin, S. Kissau, H. Haudeck, B. Ade-Thurrow, D. Gómez-Pereira, P. T. Oon, and A. Dornburg, "Global and educational disparities in AI integration: A study of L2 teacher training and usage patterns," *System*, vol. 136, p. 103901, 2026, doi: 10.1016/j.system.2025.103901.
- [13] UNESCO, *AI Competency Framework for Teachers*. Paris: UNESCO, 2024.
- [14] OECD, *Digital Education Outlook 2023: Towards an Effective Digital Education Ecosystem*. OECD Publishing, 2023, doi: 10.1787/c74f03de-en.
- [15] E. Rudolph, H. Seer, C. Mothes, and J. Albrecht, "Automated feedback generation in an intelligent tutoring system for counselor education," in *Proc. 19th Conf. Computer Science and Intelligence Systems (FedCSIS)*, ser. Annals of Computer Science and Information Systems, vol. 39, 2024, pp. 501–512, doi: 10.15439/2024F1649.
- [16] T. C. Freitas, M. J. V. Pereira, A. C. Neto, and P. R. Henriques, "Goliath, a programming exercises generator supported by AI," in *Proc. 19th Conf. Computer Science and Intelligence Systems (FedCSIS)*, ser. Annals of Computer Science and Information Systems, vol. 39, 2024, pp. 331–342, doi: 10.15439/2024F8479.
- [17] Y. Feldman-Maggor, M. Cukurova, C. Kent, and G. Alexandron, "The impact of explainable AI on teachers' trust and acceptance of AI EdTech recommendations: The power of domain-specific explanations," *International Journal of Artificial Intelligence in Education*, 2025, doi: 10.1007/s40593-025-00486-6.
- [18] OECD, *OECD Digital Education Outlook 2021*. OECD Publishing, 2021, doi: 10.1787/589b283f-en.
- [19] UNESCO, *Guidance for Generative AI in Education and Research*. UNESCO Publishing, 2023.
- [20] —, *AI and the Future of Education: Disruptions, Dilemmas and Directions*. UNESCO Publishing, 2025.
- [21] L. Faverio, J. A. Pérez-Ortiz, T. Käser, and N. Oliver, "AI in education beyond learning outcomes: Cognition, agency, emotion, and ethics," *arXiv preprint arXiv:2602.04598*, 2026, doi: 10.48550/arXiv.2602.04598.
- [22] A. J. Karran, P. Charland, J.-T. Martineau, A. Ortiz de Guinea Lopez de Arana, A.-M. Lesage, S. Sénécal, and P.-M. Léger, "Multi-stakeholder perspective on responsible artificial intelligence and acceptability in education," *arXiv preprint arXiv:2402.15027*, 2024, doi: 10.48550/arXiv.2402.15027.
- [23] P. Mahajan, "What is ethical: AIHED driving humans or human-driven AIHED? A conceptual framework enabling the ethos of AI-driven higher education," *arXiv preprint arXiv:2503.04751*, 2025, doi: 10.48550/arXiv.2503.04751.