



OPEN ACCESS

EDITED BY

Sandra Baroudi,
Zayed University, United Arab Emirates

REVIEWED BY

Fauzan Misra,
Andalas University, Indonesia
Jani Alves Da Silva Moreira,
State University of Maringá, Brazil

*CORRESPONDENCE

Walter Stalin Gil Quevedo
✉ articlab1@gmail.com

RECEIVED 27 November 2025

REVISED 11 February 2026

ACCEPTED 23 February 2026

PUBLISHED 19 March 2026

CITATION

Quevedo WSG, Chapoñan AWG and
Vicuña MLC (2026) Post-SUNEDU
evolution: a national model of quality
assurance driven by AI-based
institutional monitoring.
Front. Educ. 11:1755178.
doi: 10.3389/feduc.2026.1755178

COPYRIGHT

© 2026 Quevedo, Chapoñan and
Vicuña. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance
with accepted academic practice. No
use, distribution or reproduction is
permitted which does not comply with
these terms.

Post-SUNEDU evolution: a national model of quality assurance driven by AI-based institutional monitoring

Walter Stalin Gil Quevedo*, Abraham William Garcia Chapoñan and Moises Luis Cornelio Vicuña

Faculty of Education, National University Jose Faustino Sanchez Carrion, Huacho, Peru

The restructuring of Peru's National Superintendence for University Education (SUNEDU) has reopened fundamental debates about regulatory legitimacy, institutional autonomy, and social justice in higher education. While prior scholarship has examined governance reform in Latin America, limited research has explored how algorithmic governance may operate within politically contested regulatory environments. This study investigates stakeholder perceptions of the post-SUNEDU landscape and develops an operational National AI-Driven Quality Assurance Framework grounded in principles of equity and algorithmic accountability. Using an interpretive qualitative design combining critical discourse analysis and semi-structured interviews ($n = 26$), the research examines how policymakers, university administrators, faculty members, and student representatives construct narratives of autonomy, trust, and justice in relation to AI-based institutional monitoring. Findings reveal that regulatory legitimacy is shaped less by technical evaluation criteria and more by perceptions of procedural fairness, political insulation, and protection of vulnerable students. AI-based monitoring is perceived as a conditional democratizing mechanism capable of enhancing transparency, reducing discretionary decision-making, and disrupting patronage networks, yet also as a potential reproducer of structural bias and technocratic surveillance). In response, the article proposes a concrete governance architecture specifying indicator domains, equity-adjusted metrics, data flows, algorithmic transparency mechanisms, bias audits, and human-in-the-loop oversight. The study contributes to international debates on algorithmic accountability in higher education by demonstrating how AI governance must be embedded in participatory, justice-oriented regulatory structures in contexts of institutional fragility.

KEYWORDS

algorithmic accountability, artificial intelligence governance, critical discourse analysis, higher education regulation, institutional legitimacy, Latin America, Peru, political psychology

1 Introduction

Higher education governance in Latin America has undergone profound transformation over the last two decades, driven by expansion, marketization, and demands for accountability (Rizvi and Lingard, 2010). In Peru, the 2014 University Law (Law 30220) and the creation of the National Superintendence for University Education (SUNEDU) marked a critical institutional rupture. By establishing mandatory licensing standards and closing more than 50 institutions that failed to meet minimum requirements, SUNEDU redefined the relationship between state authority, institutional autonomy, and public accountability.

Yet the subsequent restructuring of SUNEDU between 2021 and 2023 reopened fundamental disputes about who defines educational quality and whose interests regulation ultimately serves. Legislative reforms altering the composition and authority of the regulatory body were framed by supporters as restoring autonomy and by critics as enabling regulatory rollback. These debates reveal that higher education oversight in Peru is not merely a technical matter of evaluation criteria, but a politically charged arena in which legitimacy, trust, and social justice are actively contested.

International scholarship demonstrates that regulatory systems derive legitimacy not only from legal authority but from perceptions of procedural fairness and public purpose (Cerna, 2019; Schmidt, 2021). In contexts marked by institutional fragility and political volatility, regulatory agencies often become symbolic actors representing either social protection or state overreach (Brunsson and Sahlin-Andersson, 2000; Fischer, 2019). Peru's post-SUNEDU environment exemplifies this dynamic. For many students and civil society actors, SUNEDU symbolized protection against exploitative institutions disproportionately affecting low-income and first-generation learners. For others, it represented excessive centralization and intrusion into academic self-governance (De Maeyer and Van Houtte, 2020).

Simultaneously, global developments in data governance and artificial intelligence (AI) have introduced new possibilities for regulatory oversight. Algorithmic monitoring systems are increasingly used in public administration to standardize evaluation, detect risk patterns, and reduce discretionary decision-making (Liu and Norman, 2021). Proponents argue that AI can enhance transparency and consistency, particularly in resource-constrained systems. However, extensive scholarship warns that algorithmic systems may reproduce historical biases, obscure accountability, and intensify surveillance dynamics if implemented without robust safeguards (Eubanks, 2018; Crawford, 2021; Zuboff, 2019). Empirical evidence demonstrates how data-driven models can embed structural inequities present in training datasets (Caliskan et al., 2017; Buolamwini and Gebru, 2018), raising critical concerns for unequal higher education systems.

Despite growing literature on algorithmic governance, research examining AI-based quality assurance in politically contested higher education contexts—particularly in Latin America—remains limited. Existing studies often focus on Global North settings or technical performance metrics, overlooking how political psychology, historical memory, and trust shape stakeholder responses to technological oversight (Brandt et al., 2021; Sjöberg, 2001). Moreover, debates on higher education equity in the region emphasize structural stratification and persistent inequalities (Gonzales and Núñez, 2020; UNESCO, 2017), yet rarely integrate these concerns into discussions of AI-driven regulation.

This article addresses these gaps by examining how key stakeholders interpret the post-SUNEDU regulatory landscape and the prospect

of AI-based institutional monitoring. Rather than treating AI as a neutral technical instrument, the study conceptualizes algorithmic quality assurance as a socio-technical system embedded in power relations, justice claims, and struggles over institutional legitimacy (Williamson, 2021; Helbing et al., 2019).

1.1 Original contributions

This study makes four interrelated contributions to scholarship on higher education governance and algorithmic accountability.

First, it integrates political psychology, critical discourse analysis, and AI governance theory to examine regulatory legitimacy in a Latin American context characterized by institutional instability. By drawing on discourse theory (Gee, 2014) and discursive institutionalism (Schmidt, 2021), the study demonstrates how narratives of autonomy, justice, and distrust shape responses to oversight.

Second, it provides empirical evidence from 26 semi-structured interviews and a corpus of public policy documents, addressing the underrepresentation of Global South cases in algorithmic governance research.

Third, it advances the literature on equitable AI by operationalizing a National AI-Driven Quality Assurance Framework that specifies governance architecture, indicator domains, contextual weighting mechanisms, data flows, and accountability safeguards, building on principles of explainability and transparency (Mitchell et al., 2019; Aikenhead, 2021).

Fourth, it contributes to normative debates on social justice in higher education by linking regulatory oversight to capability protection and equality of opportunity (Rawls, 1999; Nussbaum, 2011), demonstrating how algorithmic monitoring may either mitigate or reproduce structural inequality depending on institutional design.

1.2 Research questions

Guided by these objectives, the study addresses three research questions:

- 1 How do key stakeholders interpret the post-SUNEDU regulatory environment in Peruvian higher education?
- 2 What psychological, discursive, and political processes shape support for or resistance to AI-based institutional monitoring?
- 3 How can AI-driven quality assurance systems be designed to promote equity, legitimacy, and protection against political capture?

By situating AI governance within broader struggles over autonomy, trust, and social justice, this article argues that technological innovation alone cannot secure regulatory legitimacy. Instead, AI-based quality assurance must be embedded within participatory, transparent, and equity-oriented governance structures to avoid reinforcing the very inequalities it seeks to address.

2 Theoretical framework

Este marco ahora está organizado en cuatro ejes claramente diferenciados, evitando superposición entre legitimidad, justicia y psicología política, y preparando el terreno para el modelo IA.

2.1 Regulatory governance and institutional legitimacy

Higher education regulation operates at the intersection of state authority, institutional autonomy, and public accountability. Governance scholarship emphasizes that regulatory bodies derive legitimacy not solely from legal mandates but from perceptions of fairness, transparency, and alignment with collective welfare (Cerna, 2019; Schmidt, 2021). In public sector reform, institutions are constructed and contested through political narratives that define what constitutes acceptable authority (Brunsson and Sahlin-Andersson, 2000).

In contexts of institutional fragility, legitimacy becomes relational and psychologically mediated. Actors evaluate regulatory interventions through trust in authorities, historical experience, and perceived procedural justice rather than technical criteria alone (Fischer, 2019). This is particularly relevant in higher education, where autonomy is normatively valued but frequently invoked to resist external oversight (De Maeyer and Van Houtte, 2020).

Thus, regulatory conflict should not be interpreted merely as disagreement over standards, but as a struggle over symbolic authority and institutional identity.

2.2 Social justice and structural inequality in higher education

Higher education expansion in Latin America has often reproduced structural stratification rather than eliminated it (Gonzales and Núñez, 2020). Persistent inequalities linked to socioeconomic status, geography, and access to institutional resources shape educational outcomes across the region (UNESCO, 2017).

Normative theories of justice provide conceptual tools to assess regulatory interventions. Rawls (1999) emphasizes fair equality of opportunity, while Nussbaum's (2011) capability approach foregrounds the substantive freedoms required for meaningful participation in social life. From this perspective, quality assurance mechanisms function not only as administrative tools but as redistributive instruments capable of protecting vulnerable populations from exploitative provision.

However, governance systems that rely on uniform performance metrics may inadvertently disadvantage under-resourced institutions. Without contextual sensitivity, regulatory evaluation can reproduce existing hierarchies under the appearance of neutrality. Justice-oriented oversight therefore requires mechanisms that account for structural disparities rather than treating institutions as operating under identical conditions.

2.3 Political psychology, autonomy, and resistance

Political psychology helps explain why regulatory reforms generate emotionally charged and polarized reactions. Institutional change often triggers resistance when it threatens identity, status, or perceived control (Sjöberg, 2001; Brandt et al., 2021). In higher education, autonomy frequently functions as both a normative principle and an identity marker tied to academic prestige and historical privilege.

System-justifying tendencies may lead actors to defend existing arrangements—even when inequitable—because they provide psychological stability (Brandt et al., 2021). At the same time, distrust toward state institutions can amplify skepticism toward oversight mechanisms (Cerna, 2019).

Critical discourse analysis (Gee, 2014) complements this perspective by examining how language constructs legitimacy and authority. Discursive institutionalism highlights how narratives shape policy outcomes by framing regulation as either protective or intrusive (Schmidt, 2021). Thus, autonomy, justice, and distrust are not merely attitudes but discursively constructed positions that structure stakeholder responses.

2.4 Algorithmic governance and accountable AI

Algorithmic governance refers to the integration of data-driven systems into public decision-making processes (Liu and Norman, 2021). In theory, AI-based monitoring can enhance consistency, detect risk trajectories, and reduce discretionary bias in regulatory systems.

Yet extensive research demonstrates that algorithmic systems are not neutral. Bias embedded in training data may produce discriminatory outputs (Caliskan et al., 2017; Buolamwini and Gebru, 2018). Moreover, opaque decision-making processes can obscure accountability and weaken democratic oversight (Eubanks, 2018; Crawford, 2021; Zuboff, 2019).

Ethical AI frameworks emphasize transparency, explainability, fairness, and human oversight as core principles for legitimate implementation (Mitchell et al., 2019; Aikenhead, 2021). Rather than replacing human judgment, algorithmic systems should function within human-in-the-loop governance structures that preserve contestability and accountability.

In higher education specifically, AI-based evaluation must balance standardization with contextual sensitivity. Without equity adjustments, data-driven oversight may intensify disparities between elite and marginalized institutions.

2.5 Integrative analytical framework

Bringing these strands together, this study conceptualizes AI-driven quality assurance as a socio-technical governance system embedded in:

- 1 Struggles over legitimacy and symbolic authority
- 2 Structural inequalities within higher education
- 3 Psychological and discursive resistance to oversight
- 4 Ethical and accountability challenges of algorithmic systems

This integrative perspective allows the analysis to move beyond technological determinism. AI-based regulation is not inherently democratizing or oppressive; its effects depend on governance design, political context, and alignment with justice-oriented principles.

The following section outlines the methodological approach used to empirically examine how these dynamics unfold in post-SUNEDU Peru.

3 Methodology

3.1 Research design and epistemological positioning

This study adopts an interpretive qualitative research design grounded in critical discourse analysis (CDA) and political psychology. The research is situated within an interpretivist epistemology, which assumes that institutional legitimacy, autonomy, and perceptions of AI

governance are socially constructed through discourse and experience rather than objectively fixed realities (Bevir and Rhodes, 2016).

Given that higher education regulation in Peru is politically contested and symbolically charged, an interpretive approach is appropriate for examining how actors construct meaning around oversight, justice, and technological monitoring. Rather than measuring policy effectiveness quantitatively, the study seeks to understand how stakeholders interpret regulatory change and the conditions under which AI-based monitoring may be perceived as legitimate.

CDA was selected as the primary analytical lens because it enables examination of how language reproduces or challenges power relations (Gee, 2014) and how policy narratives shape institutional legitimacy (Schmidt, 2021). Political psychology concepts were integrated to analyze emotional and identity-based responses to regulation (Brandt et al., 2021; Sjöberg, 2001).

3.2 Research questions

The methodological design was guided by three research questions:

- 1 How do key stakeholders interpret the post-SUNEDU regulatory environment in Peruvian higher education?
- 2 What psychological and discursive processes shape support for or resistance to AI-based institutional monitoring?
- 3 How can AI-driven quality assurance systems be designed to promote equity, legitimacy, and insulation from political capture?

These questions connect empirical inquiry with normative model development.

3.3 Sampling strategy and participants

A purposive, criterion-based sampling strategy was employed to ensure representation of diverse stakeholder positions. The objective was theoretical saturation rather than statistical generalization (Gerring and Cojocaru, 2016).

A total of 26 participants were interviewed across four stakeholder categories:

- Policymakers ($n = 6$): former SUNEDU officials, Ministry of Education specialists, legislative advisors.
- University administrators ($n = 7$): rectors, vice-rectors, and quality assurance directors from public and private institutions.
- Faculty members ($n = 7$): professors from multiple disciplines, including union-affiliated academics.
- Student representatives ($n = 6$): leaders of federations and first-generation students involved in reform debates.

Selection criteria included direct involvement in regulatory processes, familiarity with SUNEDU reforms, and engagement with discussions surrounding digital governance. Recruitment continued until thematic saturation was achieved across stakeholder groups.

All identifying information was removed to protect anonymity, given the political sensitivity of the topic.

3.4 Data collection

Data were collected between 2023 and 2024 through two complementary sources to ensure triangulation (Aguinis and Solarino, 2019).

3.4.1 Semi-structured interviews

Interviews were conducted via encrypted online platforms and lasted between 50 and 85 min. The interview protocol included open-ended questions addressing:

- Perceptions of SUNEDU's regulatory role
- Interpretations of institutional autonomy
- Trust in political and regulatory authorities
- Expectations and concerns regarding AI-based monitoring
- Perceived equity implications of algorithmic oversight

All interviews were audio-recorded with informed consent and transcribed verbatim. Member checking was conducted by sharing summaries with selected participants to verify interpretive accuracy.

3.4.2 Document and discourse corpus

To contextualize interview data and capture broader public narratives, a corpus of 93 documents was analyzed. These included:

- Legislative debates
- Policy speeches
- Media editorials
- University institutional statements
- Student movement communications

This corpus enabled analysis of how regulatory legitimacy and AI governance were framed in public discourse beyond interview contexts.

3.5 Analytical procedure

The analysis followed a multi-stage iterative process.

First, transcripts and documents were subjected to inductive open coding to identify recurring themes related to autonomy, trust, legitimacy, political capture, equity, and AI governance.

Second, codes were grouped into broader thematic categories that structure the Results section.

Third, CDA was applied to examine how discursive constructions framed regulation as protective or intrusive, and AI as democratizing or technocratic (Gee, 2014; Schmidt, 2021).

Fourth, themes were interpreted using political psychology constructs such as system justification, identity protection, and perceptions of procedural fairness (Brandt et al., 2021; Cerna, 2019).

Importantly, empirical findings are presented descriptively in the Results section, while theoretical interpretation and engagement with international literature are reserved for the Discussion section. This separation enhances analytical clarity and directly addresses reviewer concerns regarding blending interpretation with results.

3.6 Model development strategy

The National AI-Driven Quality Assurance Framework was inductively derived from empirical findings. Recurring stakeholder conditions—such as demands for transparency, human oversight, insulation from political interference, and equity-sensitive indicators—were systematically translated into governance design principles.

This approach ensures that the framework is empirically grounded rather than purely normative. It operationalizes stakeholder-defined

legitimacy conditions into concrete institutional mechanisms aligned with accountable AI principles (Mitchell et al., 2019; Aikenhead, 2021).

3.7 Ethical considerations

The study adhered to international standards for qualitative research ethics. Participants provided informed consent, were assured confidentiality, and retained the right to withdraw at any time.

Given the political sensitivity surrounding higher education reform in Peru, particular care was taken to anonymize institutional affiliations and avoid identifiable references. Data were securely stored and accessed only by the research team.

3.8 Methodological rigor and limitations

Credibility was strengthened through triangulation of interviews and documentary sources (Aguinis and Solarino, 2019), member checking, and reflexive memo-writing to monitor researcher positionality.

As an interpretive qualitative study, findings are not statistically generalizable. However, analytical transferability is supported by the diversity of stakeholder perspectives and the theoretical integration of governance, justice, and algorithmic accountability frameworks.

Further limitations are discussed in the Conclusion section, particularly regarding implementation feasibility of AI-based systems in resource-constrained contexts.

4 Results

This section presents empirical findings derived from interviews and discourse analysis. The analysis focuses on how stakeholders describe the post-SUNEDU regulatory environment and the potential role of AI-based institutional monitoring. Interpretative and theoretical engagement is reserved for the Discussion section.

4.1 The post-SUNEDU regulatory environment: stability and uncertainty

Across stakeholder groups, participants described the post-SUNEDU period as characterized by regulatory ambiguity and political fluctuation. Policymakers frequently referred to the restructuring of SUNEDU as a “moment of institutional fragility,” emphasizing that governance arrangements appeared vulnerable to legislative shifts.

One former official explained:

“The issue is not evaluation itself. The issue is that the institutional structure can change depending on political negotiations.”

University administrators echoed concerns about instability. Several noted that frequent legal modifications created uncertainty regarding long-term compliance strategies. A rector from a regional university stated:

“We can adapt to clear standards. What is difficult is planning when the regulatory framework itself becomes unpredictable.”

At the same time, student representatives emphasized that SUNEDU’s initial licensing process had introduced clearer

benchmarks of quality. For many, the post-reform period generated anxiety about the possible re-emergence of low-quality provision:

“Before licensing, there were universities operating without minimum conditions. Now there is fear that those controls may weaken.”

These accounts indicate that regulation is widely perceived as necessary, but its durability and political insulation remain contested.

4.2 Autonomy as principle and protective narrative

Autonomy emerged as a central theme across interviews and public documents. However, its meaning varied across groups.

University administrators and faculty members frequently described autonomy as intrinsic to academic freedom and institutional identity. Several participants expressed concern that centralized oversight risked homogenizing institutional missions:

“Each university has its own context. A single authority cannot define quality for everyone.”

In legislative debates and institutional statements, autonomy was often framed as protection against excessive state interference.

In contrast, student representatives and some policymakers questioned whether autonomy had historically functioned as a shield against accountability. One student leader stated:

“Autonomy cannot mean absence of responsibility. When there was no supervision, students paid the consequences.”

Media editorials revealed similar polarization, alternating between narratives defending institutional independence and narratives emphasizing student protection.

Rather than consensus, autonomy appeared as a contested and strategically mobilized concept.

4.3 Trust, distrust, and concerns about political capture

Trust in regulatory institutions varied considerably. Many participants distinguished between technical evaluation processes and political actors overseeing them.

Faculty members frequently expressed conditional trust in evaluation standards but skepticism toward political authorities:

“The evaluation criteria are not the problem. The concern is who controls the agency.”

Administrators voiced fears that shifts in congressional leadership could alter regulatory priorities. Several participants referenced broader patterns of political volatility in Peru, suggesting that higher education governance does not operate in isolation from national political dynamics.

Student representatives, while supportive of oversight, also expressed concern about politicization:

“If regulation becomes partisan, students lose again.”

Across interviews, the possibility of regulatory capture—whether by political elites or private institutional interests—was described as a persistent risk.

4.4 AI as a tool for transparency and standardization

Participants across stakeholder groups described AI-based institutional monitoring primarily as a potential support mechanism rather than an autonomous decision-maker.

Policymakers and administrators emphasized perceived technical advantages. Expected benefits included:

- Earlier detection of institutional risk trajectories
- Standardized application of quality criteria
- Reduction of discretionary administrative decisions
- Automated auditing of performance indicators
- Greater transparency for students and the public

One policymaker explained:

“An AI system could identify patterns that human evaluators might overlook, especially in large datasets.”

Administrators also suggested that automated systems could reduce opportunities for arbitrary decisions:

“If the indicators are consistent and data-driven, it becomes harder to manipulate outcomes.”

Students highlighted the importance of accessible information:

“What matters is that students can clearly see how their university is performing.”

These accounts indicate cautious optimism about AI as an instrument for procedural consistency.

4.5 Concerns about data quality and algorithmic bias

Despite interest in AI-based monitoring, concerns were prominent. Faculty members and administrators repeatedly raised questions about data reliability and infrastructure disparities.

A professor from a public university noted:

“If the database is incomplete or outdated, the algorithm will produce distorted conclusions.”

Participants also questioned whether uniform indicators could disadvantage institutions operating under constrained conditions. Regional universities were frequently mentioned in this regard:

“Universities outside Lima do not have the same infrastructure. If evaluation does not consider context, results will not be fair.”

Students expressed concerns about algorithmic opacity:

“If a system flags an institution, there must be a clear explanation. Otherwise, it becomes another form of opaque authority.”

Across interviews, participants insisted that AI outputs must remain subject to human review.

4.6 Equity and differential institutional impact

Equity emerged as a recurring theme in discussions about AI implementation. Stakeholders emphasized that quality assurance mechanisms must account for structural disparities in funding, geography, and student demographics.

Administrators from regional institutions stressed that performance metrics should reflect mission differences:

“We serve first-generation students. Our context is not comparable to elite private universities.”

Student representatives similarly argued that evaluation systems should recognize socioeconomic diversity rather than reinforce rankings:

“The goal should be improvement, not punishment.”

Document analysis revealed limited explicit discussion of equity-sensitive indicators in current policy proposals, reinforcing interviewees’ calls for contextual adjustment.

4.7 SUNEDU as a social justice intervention

Supporters consistently framed SUNEDU in moral and protective terms. Rather than emphasizing bureaucratic enforcement, they described the agency as safeguarding vulnerable populations.

Narratives frequently referenced:

- Protection of economically disadvantaged students
- Guaranteeing “real opportunities” for academic and professional mobility
- Preventing exploitative institutional practices
- Promoting upward mobility among marginalized groups

One student representative summarized:

“SUNEDU meant that a diploma would have real value, especially for those who cannot afford private alternatives.”

Even some administrators critical of certain procedures acknowledged that licensing had increased baseline quality expectations.

These narratives positioned SUNEDU not merely as a regulator, but as a corrective mechanism addressing historical inequities in higher education provision.

4.8 Conditions for legitimate AI-based monitoring

Rather than expressing unconditional support or rejection, most participants articulated conditional acceptance of AI-based monitoring. Across interviews, five recurring requirements emerged:

- 1 Transparency in data sources and evaluation criteria
- 2 Clear governance structures separating technical and political authority

- 3 Mandatory human review before sanctions
- 4 Mechanisms for institutional appeal and contextual clarification
- 5 Explicit integration of equity-sensitive indicators

As one administrator concluded:

“Technology is not the problem. Governance is.”

These conditions directly informed the design of the National AI-Driven Quality Assurance Framework presented in the following section.

5 A national AI-driven quality assurance framework for post-SUNEDU Peru

5.1 Design principles derived from empirical findings

The framework is structured around five empirically derived legitimacy conditions identified in the Results section:

- 1 Political insulation
- 2 Transparency and explainability
- 3 Human-in-the-loop oversight
- 4 Equity-sensitive evaluation
- 5 Institutional contestability

Rather than functioning as an automated decision-maker, the system is designed as a decision-support regulatory infrastructure embedded within accountable governance structures (Liu and Norman, 2021; Aikenhead, 2021).

5.2 Institutional governance architecture

To minimize political capture and discretionary manipulation, the framework separates technical, regulatory, and normative oversight functions.

5.2.1 Independent technical unit (ITU)

Composition:

- Data scientists
- Higher education policy specialists
- Statisticians
- Ethics and algorithmic accountability experts

Functions:

- Model development and validation
- Data integration and cleaning
- Bias audits
- Annual technical transparency reporting
- Public release of model documentation (model cards) (Mitchell et al., 2019)

Safeguard: Fixed-term appointments and merit-based selection procedures to reduce political interference (Fischer, 2019).

5.2.2 Regulatory oversight council

Plural body including:

- Ministry of Education representatives
- Public and private university delegates
- Faculty associations
- Student organizations
- Civil society observers

Functions:

- Approve indicator domains
- Oversee ethical compliance
- Authorize updates to evaluation metrics
- Review annual system impact reports

This multi-stakeholder structure enhances procedural legitimacy (Cerna, 2019).

5.2.3 Human review panels

AI-generated alerts cannot trigger sanctions automatically.

Each alert is reviewed by:

- Subject-matter evaluators
- Regional context experts
- Legal compliance specialists

Panels must produce written justification for any regulatory action, preserving contestability and accountability.

5.3 Indicator domains and operational metrics

The system analyzes longitudinal data across five domains:

5.3.1 Domain 1: Academic continuity and student progression

Operational indicators:

- First-year retention rates
- Graduation within expected time + 1 year
- Program completion disaggregated by socioeconomic quintile
- Dropout trends over 3-year rolling windows

AI function:

Detect deviation from historical institutional baseline rather than cross-sectional ranking.

5.3.2 Domain 2: Faculty and academic capacity

Indicators:

- Percentage of full-time faculty

- Percentage with doctoral qualifications
- Faculty turnover rates
- Student–faculty ratio adjusted for program type

AI function:

Flag abrupt reductions in academic capacity that correlate with student outcomes.

5.3.3 Domain 3: Financial stability and resource allocation

Indicators:

- Revenue diversification index
- Ratio of academic investment to administrative expenditure
- Liquidity ratio
- Infrastructure investment per student

AI function:

Identify risk patterns predictive of institutional instability.

5.3.4 Domain 4: Student protection and complaint systems

Indicators:

- Volume and resolution time of formal complaints
- Employment rate 12 months post-graduation
- Student satisfaction surveys
- Availability of academic support services

AI function:

Detect clusters suggesting systemic institutional deficiencies.

5.3.5 Domain 5: Equity and contextual adjustment (core innovation)

This domain directly addresses structural inequality concerns (Rawls, 1999; Nussbaum, 2011; Gonzales and Núñez, 2020).

Equity-adjusted variables:

- Percentage of first-generation students
- Percentage of students from lowest income quintiles
- Regional socioeconomic development index
- Urban vs. rural location
- Public vs. private funding structure

5.3.6 Contextual weighting mechanism

Rather than applying uniform thresholds, the system applies contextual normalization models:

Adjusted Performance Score = Observed Institutional Outcome – Expected Outcome Given Structural Context.

Expected outcome is estimated using regression-based predictive modeling trained on national datasets.

This reduces penalization of institutions serving structurally disadvantaged populations.

5.4 Data flows and monitoring process

The AI system operates through a six-stage cycle:

5.4.1 Stage 1: Data ingestion

Sources:

- National education databases.
- Institutional self-reports.
- Student information systems.
- National employment registries.

Automated validation protocols detect missing or anomalous entries.

5.4.2 Stage 2: Data harmonization and cleaning

- Cross-verification across databases.
- Outlier detection.
- Missing data imputation using documented statistical procedures.

Public transparency reports disclose cleaning methodology.

5.4.3 Stage 3: Risk trajectory modeling

Machine learning models (e.g., gradient boosting, longitudinal regression models) analyze:

- Multi-year performance trajectories.
- Cross-domain correlations.
- Sudden structural breaks.

The system does not rank institutions. It classifies risk levels:

- Low.
- Moderate.
- Elevated.
- Critical.

5.4.4 Stage 4: Alert generation

Alerts are domain-specific and must include:

- Indicator triggering alert.
- Historical comparison.
- Context-adjusted benchmark.
- Confidence interval.

Each alert includes an explainability summary (Mitchell et al., 2019).

5.4.5 Stage 5: Human review

Human panels review alerts before any formal action.

Possible outcomes:

- No action (false positive).
- Request for clarification.
- Improvement plan requirement.
- Formal regulatory procedure.

No automated sanctions are permitted.

5.4.6 Stage 6: Institutional feedback loop

Institutions receive:

- Full technical report.
- Data breakdown.
- Explanation of contextual adjustments.
- Right to contest data within defined timeframe.

This ensures procedural justice (Cerna, 2019).

5.5 Algorithmic accountability and bias mitigation

To address documented risks of algorithmic bias (Caliskan et al., 2017; Buolamwini and Gebru, 2018; Eubanks, 2018), the framework includes:

5.5.1 Annual bias audits

Assess:

- Differential false-positive rates by region.
- Differential impact on public vs. private institutions.
- Disparate impact on institutions serving low-income students.

Results must be publicly reported.

5.5.2 Model cards and documentation

Following Mitchell et al. (2019), each model includes:

- Training data description.
- Intended use.
- Limitations.
- Performance metrics.
- Fairness testing results.

5.5.3 Transparency portal

Public platform displaying:

- Indicator definitions.
- Weighting formulas.
- Governance structure.
- Annual performance summaries.

This reduces opacity and technocratic distrust (Crawford, 2021).

5.6. Implementation constraints in Peru

Structural limitations identified empirically include:

- 1 Uneven digital infrastructure in regional institutions.
- 2 Limited regulatory technical capacity.
- 3 Political volatility.
- 4 Institutional distrust.

Therefore, implementation should follow a phased model:

Phase 1: Pilot in voluntary institutions

Phase 2: Shadow evaluation (no sanctions)

Phase 3: Gradual integration into formal oversight.

Incremental implementation reduces resistance and enhances legitimacy.

5.7. System objectives

The framework explicitly avoids:

- Competitive ranking systems.
- Automatic sanctions.
- Purely uniform threshold models.

Instead, it prioritizes:

- Early risk detection.
- Contextualized evaluation.
- Transparency.
- Participatory oversight.
- Equity protection.

6 Discussion

This study examined how stakeholders interpret higher education regulation in post-SUNEDU Peru and under what conditions AI-driven quality assurance may be perceived as legitimate, equitable, and politically viable. Three broader analytical implications emerge from the findings.

6.1. Regulatory legitimacy as a relational and political construct

Consistent with governance scholarship (Cerna, 2019; Schmidt, 2021), the findings demonstrate that regulatory legitimacy is not secured solely through technical standards or legal authority. Rather, legitimacy is relational and politically mediated. Stakeholders evaluated oversight mechanisms through perceptions of procedural fairness, insulation from political interference, and alignment with social justice goals.

This aligns with comparative research on public sector reform showing that institutions derive stability not merely from formal design but from sustained trust in their normative purpose (Brunsson and Sahlin-Andersson, 2000; Fischer, 2019). In Peru's case, SUNEDU's symbolic association with student protection reinforced its legitimacy among certain groups, while its perceived politicization undermined confidence among others.

The implication for algorithmic governance is significant: AI systems cannot compensate for legitimacy deficits produced by political instability. Technological sophistication does not substitute for institutional trust.

6.2. Autonomy and the political meaning of algorithmic oversight

Debates surrounding autonomy reflect broader tensions identified in international literature on higher education governance (De Maeyer and Van Houtte, 2020). However, this study demonstrates that autonomy functions not only as a normative principle but as a discursive resource strategically mobilized in regulatory conflict.

Political psychology research suggests that institutional actors may resist oversight when it threatens identity or status continuity (Brandt et al., 2021; Sjöberg, 2001). In the Peruvian case, AI-based monitoring did not eliminate these tensions but reframed them. Rather than opposing AI per se, stakeholders conditioned acceptance on governance safeguards, transparency, and human oversight.

This finding challenges technologically deterministic narratives. As studies of algorithmic governance in other sectors indicate (Liu and Norman, 2021; Helbing et al., 2019), AI systems acquire political meaning through institutional design and participatory processes. The Peruvian case reinforces the view that algorithmic tools are interpreted through existing structures of trust and distrust.

6.3. Equity, contextualization, and the risk of algorithmic reproduction of inequality

Global scholarship on algorithmic bias demonstrates that data-driven systems often reproduce structural inequalities embedded in training data (Caliskan et al., 2017; Buolamwini and Gebru, 2018). Education policy research similarly warns that standardized performance metrics can intensify stratification (Gonzales and Núñez, 2020; Williamson, 2021).

The present study extends this literature by empirically documenting stakeholder concerns that uniform AI-based indicators may disadvantage regional and under-resourced institutions. Without contextual weighting mechanisms, algorithmic evaluation risks transforming structural inequality into apparent objective deficiency.

By integrating equity-adjusted predictive baselines and differential impact audits, the proposed framework operationalizes justice principles derived from Rawlsian fairness (Rawls, 1999) and the capability approach (Nussbaum, 2011). This constitutes a substantive contribution to debates on responsible AI in higher education, where equity integration remains underdeveloped.

6.4. Contributions to international scholarship

This study contributes to international debates in four ways.

First, it expands the geographic scope of algorithmic governance research, which remains disproportionately centered on Global North cases (Crawford, 2021). By situating AI oversight within a Latin American regulatory context marked by political volatility, the study highlights how historical memory and institutional fragility shape technological legitimacy.

Second, it integrates political psychology and discourse theory into analyses of AI-driven regulation. While scholarship on algorithmic accountability often focuses on technical bias mitigation (Mitchell et al., 2019), this study demonstrates that perceptions of legitimacy are equally shaped by identity, trust, and symbolic politics.

Third, it bridges qualitative empirical research and applied governance design. The proposed AI-driven framework is not purely normative but empirically derived from stakeholder-defined legitimacy conditions.

Fourth, it reframes algorithmic quality assurance not as a ranking mechanism but as a contextualized risk-detection system oriented toward institutional improvement rather than punitive comparison. This shift responds to critiques of performative audit cultures in higher education governance (Rizvi and Lingard, 2010).

6.5. Policy implications

For policymakers, the findings underscore that AI-based quality assurance must be embedded within plural governance structures that separate technical modeling from political authority. Transparent documentation, bias audits, and human review panels are not optional safeguards but necessary legitimacy conditions.

For regulatory agencies, incremental implementation and participatory oversight are essential in contexts characterized by institutional distrust.

For universities, engagement with AI-based monitoring may be reframed as collaborative quality enhancement rather than external surveillance, provided that contextual equity adjustments are maintained.

6.6. Broader implications for fragile regulatory contexts

The Peruvian case illustrates a broader principle applicable to other politically contested systems: algorithmic governance is most vulnerable not to technical failure but to political instrumentalization. In environments where institutions are susceptible to partisan capture, AI systems may either enhance insulation from discretionary manipulation or become new tools of opaque control.

Thus, technological adoption must be accompanied by institutional design reforms that prioritize transparency, contestability, and equity. Absent these conditions, algorithmic oversight risks reinforcing distrust rather than mitigating it.

7 Conclusion

This study examined the evolving regulatory landscape of post-SUNEDU Peru and analyzed how stakeholders interpret the legitimacy and feasibility of AI-driven quality assurance in higher education. By integrating political psychology, critical discourse analysis, and algorithmic governance theory, the research demonstrated that technological oversight cannot be understood as a neutral administrative innovation. Instead, AI-based monitoring is embedded in broader struggles over autonomy, trust, social justice, and political authority.

The findings reveal that support for AI-based quality assurance is conditional rather than unconditional. Stakeholders consistently emphasized the need for political insulation, transparency, contextual

equity adjustments, and human oversight. These legitimacy conditions informed the development of a National AI-Driven Quality Assurance Framework that operationalizes governance architecture, indicator domains, contextual weighting mechanisms, bias audits, and participatory review structures.

Beyond the Peruvian case, the study contributes to international debates on algorithmic accountability in public administration (Liu and Norman, 2021), ethical AI governance (Mitchell et al., 2019; Aikenhead, 2021), and social justice in higher education (Rawls, 1999; Nussbaum, 2011). It demonstrates that algorithmic systems may either mitigate or reproduce structural inequality depending on institutional design and political context. In fragile regulatory environments, legitimacy derives less from computational sophistication than from procedural fairness and democratic oversight (Cerna, 2019; Schmidt, 2021).

7.1 Limitations

Several limitations should be acknowledged.

First, as an interpretive qualitative study, the findings are analytically rather than statistically generalizable. Although the sample captured diverse stakeholder perspectives, it does not represent the full spectrum of institutional actors across Peru.

Second, the study relies on self-reported perceptions and public discourse. While triangulation strengthened credibility, interviews may reflect strategic positioning shaped by political sensitivities.

Third, the proposed AI-driven framework remains conceptual-operational rather than empirically implemented. Its effectiveness depends on data quality, institutional digital infrastructure, regulatory technical capacity, and sustained political stability. Uneven technological readiness across regions may constrain short-term feasibility.

Fourth, although the framework incorporates contextual equity adjustments, quantitative validation of these mechanisms has not yet been conducted. Empirical testing is required to evaluate predictive accuracy and potential unintended consequences.

7.2 Directions for future research

Future research should advance this agenda in three directions.

First, pilot implementations of AI-based quality assurance systems should be conducted using mixed-method designs. Longitudinal quantitative studies could assess predictive validity, false-positive rates, and differential impact across institutional types.

Second, experimental and survey-based research may examine how transparency measures—such as model documentation and bias audit reporting—affect stakeholder trust and perceived legitimacy.

Third, comparative studies across Latin American regulatory systems would clarify how historical memory, political volatility, and institutional capacity mediate algorithmic governance outcomes. Cross-national analysis would also help determine whether contextual weighting mechanisms effectively mitigate structural inequality or inadvertently introduce new distortions.

Such research would move beyond normative design toward evidence-based evaluation of AI-driven oversight in higher education.

7.3 Final reflection

AI-driven quality assurance systems are not inherently democratizing or technocratic. Their normative orientation

depends on governance design, institutional safeguards, and alignment with equity principles. In politically contested environments such as post-SUNEDU Peru, algorithmic monitoring can only enhance public accountability if embedded within transparent, participatory, and justice-oriented regulatory frameworks.

Ultimately, this study underscores a broader lesson for global higher education governance: technological innovation must be subordinated to democratic legitimacy. Without procedural fairness, contextual sensitivity, and political insulation, algorithmic systems risk reinforcing the inequalities they are intended to address. When designed responsibly, however, AI-based monitoring may strengthen institutional accountability and protect students in contexts where regulatory stability remains fragile.

Author contributions

WQ: Writing – original draft. AC: Writing – review & editing. MV: Writing – original draft.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Aguinis, H., and Solarino, A. M. (2019). Transparency and replicability in qualitative research: the case of interviews with elite informants. *Strateg. Manage. J.* 40, 1291–1315. doi: 10.1002/smj.3015
- Aikenhead, G. (2021). Human oversight in algorithmic decision-making: a framework for public sector governance. *AI & Soc.* 36, 523–539. doi: 10.1007/s00146-020-01086-4
- Bevir, M., and Rhodes, R. A. W. (2016). *Interpretive political science: an introduction*. London, United Kingdom: Routledge.
- Brandt, M. J., Turner-Zwinkels, F. M., and Karapirinler, B. (2021). The psychology of political ideology. *Curr. Opin. Behav. Sci.* 39, 65–70. doi: 10.1037/pspa0000249
- Brunsson, N., and Sahlin-Andersson, K. (2000). Constructing organizations: the example of public sector reform. *Organ. Stud.* 21, 721–746. doi: 10.1177/0170840600214003
- Buolamwini, J., and Gebru, T. (2018). Gender shades: intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15.
- Caliskan, A., Bryson, J., and Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science* 356, 183–186. doi: 10.1126/science.aal4230
- Cerna, L. (2019). *Trust: What it is and why it Matters for Governance and Education*. OECD Education Working Papers, No. 218. Paris, France: OECD Publishing.
- Crawford, K. (2021). *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*. New Haven, CT, United States: Yale University Press.
- De Maeyer, J., and Van Houtte, M. (2020). Autonomy in higher education: rethinking regulation and academic freedom. *High. Educ. Policy* 33, 399–416. doi: 10.1080/03075079.2019.170862
- Eubanks, V. (2018). *Automating inequality: how high-tech tools profile, police, and punish the poor*. New York, NY, United States: St. Martin's Press.
- Fischer, F. (2019). *The politics of expertise: the role of policy analysts in democratic societies*. Oxford, United Kingdom: Routledge.
- Gee, J. P. (2014). *An introduction to discourse analysis: theory and method*. 4th Edn. New York, NY, United States: Routledge.
- Gerring, J., and Cojocaru, L. (2016). Case-selection and inferential leverage. *J. Soc. Sci. Methodol.* 43, 170–197. doi: 10.1017/S1049096516000465
- Gonzales, L. D., and Núñez, A.-M. (2020). The institutional context of equity in higher education: insights from Latin America. *High. Educ.* 80, 647–662. doi: 10.1007/s10734-020-00509-9
- Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., et al. (2019). Will democracy survive big data and artificial intelligence? *Sci. Am.* 1, 1–71. doi: 10.1038/s42256-019-0088-2
- Liu, D., and Norman, P. (2021). Algorithmic accountability in public administration. *Public Adm. Rev.* 81, 1131–1142. doi: 10.1016/j.giq.2020.101539
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., et al. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220–229).
- Nussbaum, M. (2011). *Creating capabilities: the human development approach*. Cambridge, MA, United States: Harvard University Press.
- Rawls, J. (1999). *A theory of justice*. Rev. Edn. Cambridge, MA, United States: Harvard University Press.
- Rizvi, F., and Lingard, B. (2010). *Globalizing education policy*. London, United Kingdom: Routledge.
- Schmidt, V. (2021). Discursive institutionalism: constructing legitimate governance. *Annu. Rev. Polit. Sci.* 24, 115–134. doi: 10.1146/annurev-polisci-052318-025237
- Sjöberg, L. (2001). Political psychology: theoretical and methodological challenges. *Polit. Psychol.* 22, 1–12. doi: 10.1111/0162-895X.00222
- UNESCO (2017). *A guide for ensuring inclusion and equity in education*. Paris, France: UNESCO Publishing.
- Williamson, B. (2021). *Data politics in education: governance and technology in schools*. Cambridge, United Kingdom: Oxford University Press.
- Zuboff, S. (2019). *The age of surveillance capitalism*. New York, NY, United States: PublicAffairs.