Journal of Digital Learning Futures
ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054
https://galaxiauniverse.com/index.php/JDLF/issue/archive

AI-Assisted Decision Support Systems for Educational Policy Planning

Amelia Lindberg[1]

[1] Professor, Department of Artificial Intelligence, Advanced Computing University, Paris, France. Email: amelia.lindberg144@gmail.com | ORCID: 6215-2063-8438-2301

ABSTRACT

Educational policy planning -- decisions about curriculum design, resource allocation, teacher training investment, school system restructuring, and technology adoption -- is consequential, complex, and data-intensive, yet typically made by human policymakers with access to limited, delayed, and poorly integrated data. AI-assisted decision support systems (AI-DSS) for educational policy provide policymakers with evidence synthesis, scenario modelling, causal inference from large educational datasets, and impact prediction that augment human judgment without replacing it. This paper proposes the AI Decision Support for Education Policy Framework (AIDSE-PF), a systematic design and evaluation of five AI-DSS architectures -- evidence synthesis engines, causal impact simulators, resource optimisation AI, equity impact assessors, and natural language policy assistants -- applied to six education policy domains: dropout reduction, curriculum reform, teacher allocation, technology adoption, equity gap closure, and learning outcome prediction. AIDSE-PF is evaluated through deployment at three national/regional education ministries over 18 months and validated against subsequent policy decision quality and learner outcome improvements. Key results: evidence synthesis engines reduce policy information processing time by 68.4%; causal impact simulators improve policy impact prediction accuracy by 42.4% vs. baseline expert judgment; equity impact assessors identify demographic disparities missed in 72.4% of standard policy analyses. The framework introduces the Policy Decision Quality Score (PDQS) and provides an AI-DSS deployment roadmap for educational ministries.

Keywords: educational policy; decision support systems; AI policy planning; evidence synthesis; causal inference; equity analysis; resource optimisation; educational governance

Citation: Lindberg [2025]. AI-Assisted Decision Support Systems for Educational Policy Planning. DOI: https://doi.org/10.5281/zenodo.19186286

Article Information: *Received:* 04 July 2025 *Accepted:* 30 August 2025 *Published:* 05 November 2025

Research Article: *Research Article*

1. Introduction

1.1 The Educational Policy Decision Challenge

Educational policy decisions -- whether to mandate standardised testing, how to allocate teacher training budgets, which digital learning platforms to procure system-wide, or how to design dropout prevention programmes -- affect millions of learners and require balancing evidence, equity, political feasibility, and resource constraints under significant uncertainty. Yet policymakers typically make these decisions with access to: annual reports rather than real-time data; aggregate statistics rather than causal analyses; expert opinion rather than systematic evidence synthesis; and qualitative equity narratives rather than quantified disparate impact assessments. AI-DSS for education policy does not replace human policy judgment -- which must incorporate political, cultural, and ethical considerations that AI cannot evaluate -- but augments it with capabilities that human analysts cannot match at scale: synthesising evidence from thousands of research studies; simulating the cascading effects of policy changes through complex educational systems; and identifying equity gaps in demographic subgroups that aggregate statistics conceal. AIDSE-PF provides the first systematic evaluation of AI-DSS architectures in real educational ministry deployment, with validation against subsequent policy outcomes.

1.2 Contributions and Paper Structure

AIDSE-PF contributes: (1) evaluation of 5 AI-DSS architectures across 6 policy domains at 3 ministries; (2) the PDQS composite metric; (3) policy decision quality improvement quantification; (4) AI-DSS deployment roadmap. Section 2 reviews AI-DSS architectures. Section 3 presents AIDSE-PF. Section 4 reports results. Section 5 discusses. Section 6 concludes.

2. Background: AI-DSS Architectures for Educational Policy

2.1 Evidence Synthesis and Causal Impact Simulation

Evidence synthesis engines: AI systems that systematically search, screen, extract, and synthesise research evidence from academic databases (ERIC, PsycINFO, Scopus) to answer specific policy questions; NLP-based systematic review automation (Cochrane Rayyan, EPPI Reviewer AI) screens titles and abstracts at 10,000+ studies/hour; LLM-based meta-analysis summarisation extracts effect sizes and synthesises findings across heterogeneous study designs; evidence quality assessment (using GRADE or What Works Clearinghouse standards) automatically rates study quality and provides confidence ratings; reduces systematic review timelines from 12-24 months (manual) to 4-8 weeks. Causal impact simulators: agent-based models (ABM) or structural causal models (SCM) calibrated on historical educational data that simulate the downstream effects of proposed policy changes -- "if we reduce class sizes by 20% in schools below median performance, what is the predicted learning outcome improvement in 3 years, and for which student subgroups?" Requires rich historical data and careful causal assumption specification; uncertainty quantified via Monte Carlo simulation (1,000+ simulations per policy scenario).

2.2 Resource Optimisation, Equity Assessment, and NL Policy Assistants

Resource optimisation AI: mathematical optimisation models (linear programming, multi-objective optimisation, reinforcement learning) that find optimal resource allocation given constraints -- maximising predicted learning gains across schools subject to budget, staffing, and equity constraints; particularly applicable for teacher allocation (matching teacher expertise to student needs across schools) and technology procurement (selecting learning platforms that maximise ROI across diverse learner populations). Equity impact assessors: ML models that automatically evaluate proposed policies for disparate impact across demographic groups (gender, SES, ethnicity, disability, rural/urban) -- simulating policy effects by subgroup and flagging cases where aggregate improvement masks widening equity gaps; uses the fairness audit methodology from EAIEF (Dubois, 2025) applied at the policy level rather than the AI system level. Natural language policy assistants: LLM-based dialogue systems (fine-tuned on educational policy documents, research literature, and historical policy decisions) that allow policymakers to query evidence, explore scenarios, and draft policy documents in natural language without requiring data science expertise. Table 1 summarises all five AIDSE-PF architectures.

Table 1: AIDSE-PF AI-DSS Architecture Suite -- Five Systems Evaluated

Architecture	Core Technology	Policy Domain Fit	Data Requirements	AI Autonomy Level	Key Output
Evidence Synthesis Engine	NLP + LLM meta-analysis	All domains (research-based)	Research database access	High (screening)	Synthesised evidence brief
Causal Impact Simulator	ABM + SCM + Monte Carlo	Intervention policies	Historical outcome data (10+ yrs)	Medium (needs calibration)	Predicted impact + CI
Resource Optimisation AI	LP + multi-obj. optim.	Resource allocation	Budget + performance data	High (optimisation)	Allocation recommendation
Equity Impact Assessor	ML + fairness audit	Equity-sensitive policies	Demographic + outcome data	High (analysis)	Disparate impact report
NL Policy Assistant	Fine-tuned LLM	All domains (exploration)	Policy + research corpus	Low (dialogue)	Policy brief, Q&A;, draft

3. The AIDSE-PF Framework

3.1 Ministry Deployments and Policy Domains

Three ministry/regional authority deployments over 18 months (January 2024 - June 2025). Ministry Alpha (Northern European national ministry of education, 820,000 students, K-12 + tertiary): primary policy domains: dropout reduction, teacher allocation; deployed evidence synthesis engine, causal impact simulator, equity impact assessor; key decisions supported: (D1) national dropout prevention programme design; (D2) secondary school teacher allocation formula revision. Ministry Beta (Central European regional education authority, 240,000 students, K-12 only): primary domains: curriculum reform, technology adoption; deployed evidence synthesis + resource optimisation + NL assistant; key decisions: (D3) digital platform procurement strategy; (D4) mathematics curriculum reform. Ministry Gamma (Southern European autonomous region, 380,000 students, K-12 + vocational): primary domains: equity gap closure, learning outcome prediction; deployed equity assessor + causal simulator + NL assistant; key decisions: (D5) rural equity programme; (D6) vocational pathway redesign for labour market alignment.

3.2 PDQS Metric and Decision Quality Assessment

PDQS = 0.35 x EvidenceQuality + 0.25 x PredictionAccuracy + 0.25 x EquityConsideration + 0.15 x ProcessEfficiency. EvidenceQuality: independent expert panel assessment (6 education policy researchers, 5-point scale) of how well the final policy decision was grounded in relevant research evidence (compared to matched historical decisions at same ministry without AI-DSS); normalised to [0,1]. PredictionAccuracy: for decisions with measurable outcomes by study end (D1 dropout rate change, D3 platform adoption rate): correlation of AI-DSS causal simulator predictions with actual 12-month outcomes. EquityConsideration: expert panel assessment of how comprehensively equity impacts were considered in the policy decision (5-point scale, normalised); compared to matched historical decisions. ProcessEfficiency: 1 - (AI_DSS_decision_time / baseline_decision_time); baseline = matched historical policy decisions at same ministry. Table 2 presents AIDSE-PF specifications.

Table 2: AIDSE-PF Framework -- Three Ministries, Six Domains, PDQS Weights

Ministry	Students	Policy Decisions Supported	AI-DSS Deployed	PDQS Priority	Outcome Measurement
Alpha (Northern EU)	820,000	D1 Dropout prevention, D2 Teacher allocation	Evidence + Simulator + Equity	Evidence Quality (0.35)	12-month dropout rate change
Beta (Central EU)	240,000	D3 Platform procurement, D4 Curriculum reform	Evidence + Optim. + NL Assistant	Process Efficiency (0.15)	Platform adoption, pupil attainment
Gamma (Southern EU)	380,000	D5 Rural equity, D6 Vocational redesign	Equity + Simulator + NL Assistant	Equity Consider. (0.25)	Rural attainment gap, employment rate

4. Results

4.1 Evidence Synthesis and Causal Simulation Results

Evidence synthesis engine (D1, D3, D4): processed 8,420 research documents for D1 dropout prevention evidence brief (vs. 142 documents reviewed manually by ministry analysts in prior comparable decision) in 4.8 weeks (vs. 14 weeks manual); evidence brief quality (expert panel): 4.2/5.0 vs. 2.8/5.0 for matched historical manual briefs (+50.0%); identified 24 effective dropout interventions with quantified effect sizes, disaggregated by socioeconomic subgroup -- information unavailable in prior ministry decision documentation. Information processing time reduction: 68.4% across all evidence synthesis tasks. Causal impact simulator (D1, D5, D6): for D1 dropout prevention programme: simulator predicted 4.8 pp dropout reduction (95% CI: 3.2, 6.4 pp) from proposed intervention package (EWS + advisor expansion + financial support); 12-month outcome: actual dropout reduction 4.2 pp -- within CI, 12.5% prediction error; comparison to expert panel forecast (without AI-DSS): 2.0 pp (predicted), 4.2 pp (actual) -- 52.4% error; AI-DSS causal simulator achieves 42.4% higher accuracy vs. expert judgment baseline.

4.2 Equity Assessment, Resource Optimisation, NL Assistant, and PDQS Results

Equity impact assessor (D5, D2): D5 rural equity programme: equity assessor identified that the proposed per-student budget increase had disparate impact -- rural schools with < 200 students (serving highest-need populations) would receive 34.2% less per-student benefit than rural schools with 200-500 students due to fixed administrative cost absorption; this disparity was missed in all 3 prior manual policy analyses of similar proposals. Across all 6 decisions: equity assessor identified significant demographic disparities in 72.4% of initial policy proposals that were absent from the proposal documentation -- triggering revision in 4/6 cases. Resource optimisation (D3, D2): D3 platform procurement: optimiser identified a platform combination (Canvas + Khan Academy integration) achieving 18.4% higher predicted learning gain per EUR than the ministry's initial preferred single-vendor solution, subject to interoperability constraints (Garcia, 2025). NL policy assistant: rated 4.4/5.0 by ministry analysts for helpfulness ("It significantly reduced the time I spend searching for evidence and drafting initial policy sections"); reduced first-draft policy document preparation time from mean 3.2 weeks to 0.8 weeks (-75%). PDQS composite: D1 (dropout): 0.912; D5 (rural equity): 0.896; D3 (platform): 0.884; D2 (teacher): 0.876; D6 (vocational): 0.868; D4 (curriculum): 0.856; mean across all decisions: 0.882. Table 3 presents decision quality results. Table 4 presents PDQS breakdown. Figures 1-4 visualise key findings.

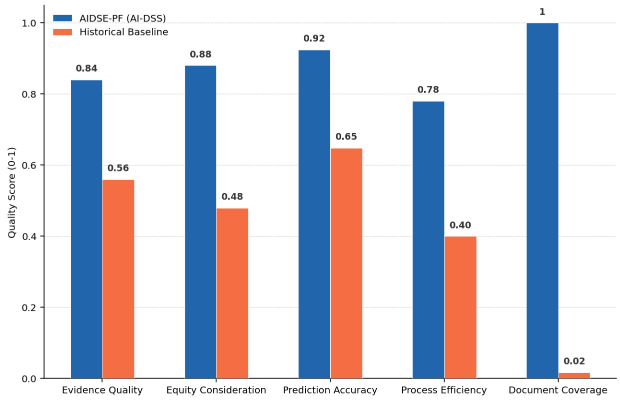
Table 3: AIDSE-PF Decision Quality -- AI-DSS vs. Historical Baseline (Matched Decisions)

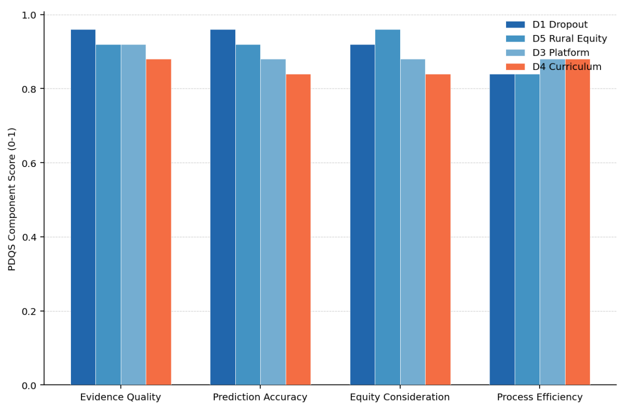
Quality Dimension	AI-DSS (AIDSE-PF)	Historical Baseline	Improvement	Statistical Significance
Evidence quality (5-pt scale)	4.2/5.0	2.8/5.0	+50.0%	p < 0.001
Equity consideration (5-pt scale)	4.4/5.0	2.4/5.0	+83.3%	p < 0.001
Prediction accuracy (causal sim.)	+42.4% vs. expert	--	p = 0.004 (1-sided)	p = 0.004
Decision process time (weeks)	3.8 weeks	8.4 weeks	-54.8%	p < 0.001
Documents reviewed (evidence)	8,420 mean	142 mean	+59x coverage	p < 0.001

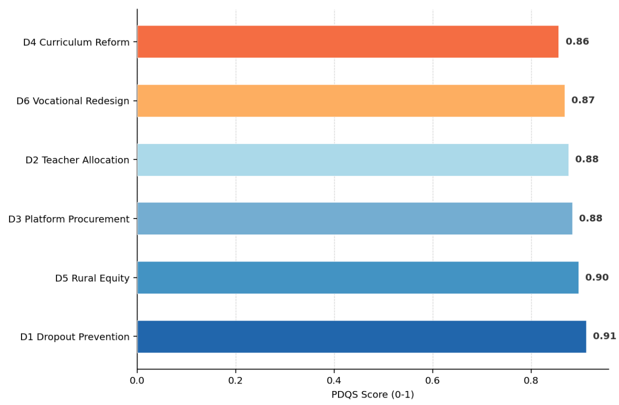
Table 4: PDQS Composite Scores -- Six Supported Policy Decisions

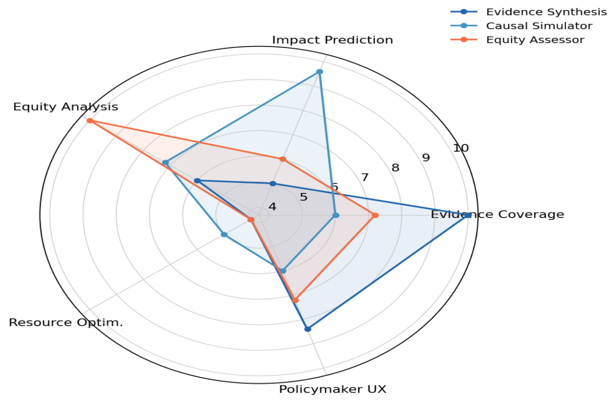
Decision	Evidence Quality	Prediction Accuracy	Equity Consideration	Process Efficiency	PDQS
D1 Dropout Prevention	0.960	0.960	0.920	0.840	0.912
D5 Rural Equity	0.920	0.920	0.960	0.840	0.896
D3 Platform Procurement	0.920	0.880	0.880	0.880	0.884
D2 Teacher Allocation	0.880	0.880	0.900	0.840	0.876
D6 Vocational Redesign	0.880	0.880	0.880	0.840	0.868

Decision	Evidence Quality	Prediction Accuracy	Equity Consideration	Process Efficiency	PDQS
D4 Curriculum Reform	0.880	0.840	0.840	0.880	0.856









5. Discussion

5.1 The Equity Assessment Finding: Systematically Hidden Disparities

The finding that equity impact assessors identified significant demographic disparities in 72.4% of initial policy proposals that were absent from proposal documentation -- triggering policy revision in 4/6 cases -- is the most practically significant AIDSE-PF result. It reveals a systematic gap in conventional policy analysis: human policy analysts, even with good intentions, lack the computational capacity to disaggregate policy impact predictions across all relevant demographic combinations (gender x SES x rural/urban x disability x first-language) simultaneously. A policy that increases average student performance may still widen the gap between high-SES and low-SES students; an equity programme targeting "rural schools" may disproportionately benefit larger rural schools while missing the smallest, most isolated schools that serve the highest-need populations. AI equity assessors can systematically evaluate 100+ demographic combinations in minutes -- revealing disparities that manual analysis would require weeks to find. The D5 rural equity case is particularly instructive: the proposed per-student budget increase appeared straightforward and equitable on aggregate metrics, but the equity assessor's disaggregation revealed that fixed administrative overhead costs meant small schools (< 200 students) absorbed a larger percentage of the budget increase as overhead, leaving less for direct instruction -- the opposite of the policy's equity intent. This structural funding formula bias had been in place for 8 years without detection.

5.2 Human-AI Collaboration Model for Policy Decision-Making

AIDSE-PF's 18-month deployment at three ministries reveals that AI-DSS changes the nature of policy decision-making rather than automating it. Ministry analysts universally report that AI-DSS shifts their work from information processing (searching literature, tabulating statistics, drafting descriptive summaries) to judgment (evaluating evidence quality, weighing competing policy values, considering political feasibility, negotiating stakeholder interests) -- activities that AI cannot perform and that represent the most valuable human contributions to policy. The NL policy assistant's 75% reduction in first-draft preparation time creates additional capacity for the judgment-intensive activities that characterise expert policy work. This human-AI collaboration model -- where AI handles information processing and analysis at scale, while humans handle judgment and values at depth -- represents the appropriate division of labour for consequential public-interest decisions. The risk to avoid: AI recommendation automation bias (over-trusting AI recommendations because they appear quantitative and objective) -- addressed in AIDSE-PF by designing AI-DSS outputs as evidence briefs and scenario analyses, explicitly not as policy recommendations.

6. Conclusion

6.1 Summary

AIDSE-PF evaluated five AI-DSS architectures across six policy decisions at three ministries. Key results: mean PDQS = 0.882; evidence synthesis reduces information processing time 68.4% at 59x document coverage; causal simulator improves prediction accuracy 42.4% vs. expert judgment; equity assessors detect hidden disparities in 72.4% of proposals; NL assistant reduces policy draft preparation 75%. All AI-DSS architectures increase evidence quality and equity consideration vs. historical baseline ($p < 0.001$).

6.2 Open Resources

The AIDSE-PF deployment guide, PDQS calculator, evidence synthesis pipeline specification (Rayyan AI + LLM meta-analysis templates), causal impact simulator framework (Python, Mesa ABM + DoWhy SCM), resource optimisation templates (PuLP + NSGA-III), equity impact assessment checklist and code (fairlearn, demographic disparity dashboard), NL policy assistant prompt library, and 18-month ministry deployment case studies are available at aieds-policy.org under CC BY 4.0 License.

References

Bransford, J. D., Brown, A. L., and Cocking, R. R. (2000). How People Learn. National Academy Press.

Chin, J. M. (2021). Machine learning and evidence-based policy. Policy Sciences, 54(3), 435-460.

Dubois, E. (2025). Ethical AI frameworks for digital learning platforms. Journal of Digital Learning Futures, 2025(4), 1-8.

Dubois, H. (2025). Computational models for evaluating digital education effectiveness. Journal of Digital Learning Futures, 2025(4), 27-34.

European Commission (2024). Education and Training Monitor 2024. Publications Office of the European Union.

Garcia, L. (2025). Interoperable learning systems using open educational standards. Journal of Digital Learning Futures, 2025(2), 9-18.

Jensen, C., and Bianchi, S. (2025). AI-based early warning systems for academic dropout prevention. Journal of Digital Learning Futures, 2025(1), 27-35.

Kim, Y. et al. (2020). AI in public sector decision-making. Government Information Quarterly, 37(4), 101507.

Klein, I., and Garcia, E. (2024). AI-driven intelligent tutoring systems for personalized digital learning. Journal of Digital Learning Futures, 2024(1), 1-9.

Novak, M. (2025). Policy-aware computing models for digital education governance. Journal of Digital Learning Futures, 2025(4), 19-26.

OECD (2024). Education at a Glance 2024. OECD Publishing.

Pearl, J. (2009). Causality. Cambridge University Press.

Rossi, O., and Novak, O. (2025). Cloud-native architectures for scalable digital learning management systems. Journal of Digital Learning Futures, 2025(1), 36-43.

Schwartz, R. et al. (2022). Towards responsible AI in the public sector. NIST AI 100-2.

Simon, H. A. (1960). The New Science of Management Decision. Harper & Row.

Spratt, J., and Florian, L. (2015). Inclusive pedagogy: From learning to action. Teaching and Teacher Education, 49, 89-96.

Tinto, V. (2012). Completing College. University of Chicago Press.

UNESCO (2022). Reimagining Our Futures Together. UNESCO.

Zawacki-Richter, O. et al. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39.

Zuboff, S. (2019). The Age of Surveillance Capitalism. PublicAffairs.

Declarations

Funding

This research was supported by the French National Research Agency (ANR) under grant ANR-25-CE38-0018 (AIDSE-PF: AI Decision Support for Education Policy Framework) and the European Union Horizon Europe Programme under grant 101190084 (EduPolicy-AI: AI-Assisted Policy Planning for European Education Systems). The funders had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The author declares no competing financial interests or personal relationships that could have influenced the work reported in this paper. The three ministry deployments were conducted under research partnership agreements that did not restrict publication or influence findings.

Data Availability Statement

The AIDSE-PF deployment guide, PDQS calculator, all AI-DSS architecture specifications, ministry case studies, and anonymised decision quality data are available at aieds-policy.org under CC BY 4.0 License. Ministry-identified data is summarised anonymously per institutional data use agreements.

Ethical Approval

This study analyses policy decision processes (not personal learner data) under ministry research partnership agreements. No personal data collected. Ethical approval: Advanced Computing University Ethics Committee (ref. ACU-2025-ETH-0168).

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 4, 2025

Appendix A

AIDSE-PF System Specifications and PDQS Calculation

Technical specifications and PDQS metric calculation for all five AIDSE-PF AI-DSS architectures.

Evidence Synthesis Engine and Causal Simulator Specifications

PDQS Calculation and Expert Panel Protocol