THE UNIVERSITY OF SYDNEY

Automated Essay Scoring in Australian Schools: Key Issues and Recommendations

White Paper

November 2022

Sydney Social Sciences and Humanities
Advanced Research Centre

Education Innovations White Paper Series
ISSN 2653-6749

Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC)

Sydney Social Sciences and Humanities Advanced Research Centre is a collaborative research facility designed to advance innovative research partnerships and methodologies across the humanities and social sciences. We have a commitment to research excellence and knowledge sharing. We broker and advance research between researchers at the University of Sydney and their outside interlocuters via visiting fellowships, exchanges, and a range of in-house initiatives designed to build authentic research partnerships that will grow and change to meet the demands of the future.

More information is available at:

sydney.edu.au/arts/our-research/centres-institutes-and-groups/sydney-social-sciences-and-humanities-advanced-research-centre/about-sssharc.html

Education Innovations and Education Futures Studio

Education Innovations is a SSSHARC-funded initiative led by Professor Kalervo Gulson at the University of Sydney. Central to this initiative is the Education Futures Studio (EFS), where diverse stakeholders come together to experiment with contemporary technologies of governance in education. At the EFS, we try to unpack and contextualise socio-technical controversies in education. We advance collaborations across the education sector in order to explore the links between teaching, emerging technologies, policy, learning, and research.

More information is available at:

education-futures-studio.sydney.edu.au

Acknowledgements

This project was funded by the Australian Research Council (Grant ID: FT180100280) and Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC). We thank everyone who participated in the hybrid forum for sharing their ideas and diverse expertise, which informs this white paper. We also appreciate the support of the SSSHARC team, the Faculty of Arts and Social Sciences, and the University of Sydney.

Authors

- Professor Kalervo Gulson, Sydney School of Education and Social Work, University of Sydney.
- Professor Greg Thompson, Education Research, Queensland University of Technology.
- Dr Teresa Swist, Senior Research Officer, Education Futures Studio, University of Sydney.
- Associate Professor Kirsty Kitto, Connected Intelligence Centre, University of Technology Sydney
- Professor Leslie Rutkowski, Quantitative Research Methodology, Indiana University.
- Professor David Rutkowski, Educational Research Methodology, Indiana University.
- Dr Anna Hogan, Senior Research Fellow, School of Teacher Education and Leadership, Queensland University of Technology.
- Vincent Zhang, Co-founder of Tech for Social Good, a non-profit for young Australians interested in responsible technology.
- Dr Simon Knight, Senior Lecturer, Director of the Centre for Research on Education in a Digital Society, University of Technology Sydney.

Recommended citation

Gulson, K., Thompson, G., Swist, T., Kitto, K., Rutkowski, L., Rutkowski, D., Hogan, A., Zhang, V., Knight, S. (2022). *Automated Essay Scoring in Australian Schools: Key Issues and Recommendations*. White Paper, November 2022. Education Innovations White Paper Series ISSN 2653-6749. Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC), University of Sydney, Australia.

Automated Essay Scoring in Australian Schools: Key Issues and Recommendations

White Paper

November 2022

Sydney Social Sciences and Humanities
Advanced Research Centre

Contents

Executive summary

This white paper outlines critical issues associated with the use of Automated Essay Scoring (AES) technology in the Australian education system. The key insights presented in this paper emerged from a collaborative, multi-stakeholder workshop held in July 2022 that explored an automated essay-scoring trial and generated future possibilities aligned with participant interests and expertise. Drawing on the workshop and our expert understanding of the wider landscape, we propose recommendations that can be adopted by various stakeholders, schools, and educational systems.

There are compelling reasons for Australian schools and education departments to investigate the use of AES. AES could potentially alleviate aspects of teachers' workload at a time when teacher attrition is historically high and teacher recruitment historically low. At the same time, AES also has the potential to de-professionalise and deskill of teachers. Educationalists are acutely aware that quality feedback can help students improve their learning across multiple subjects and domains, however parents and many are reluctant to hand that responsibility over to AES.

In 2018, concerns among teachers, teachers' unions, principles, and parents became apparent when the federal Department of Education, Skills and Employment attempted to implement a form of AES in The National Assessment Program – Literacy and Numeracy (NAPLAN). These concerns primarily registered around three issues:

- de-professionalisation of teachers,
- inequitable infrastructure in Australian schools, and
- lack of transparency from examination authorities as to how marking decisions are made.

AES cannot be approached in one dimension. It is crucial to frame the use of AES in schools as an issue with interrelated ethical, social, technical, and political implications.

The use of AES in NAPLAN ultimately proved to be politically unpopular, leading to its suspension. However, the growing implementation of AES in schools across the globe means that the use of this technology is likely to re-emerge as a controversial issue in Australia. Without political leadership in this area, it is ultimately up to educational institutions and agencies, policymakers, and school communities to assess the benefits and pitfalls of AES and navigate the way forward. Our recommendations will assist the emergence of good governance in this area.

To begin, it is crucially important to identify whether AES will be used in high-stakes or low-stakes tests. High-stakes tests are defined as those with consequential outcomes for students or educators, such as the determination of progression of students or rankings of school institutions.

If AES is to be used in Australian schools, the following issues must be considered:

- the capacity of stakeholders, including principals, teachers, and parents, to understand how AES systems work
- the infrastructure required to support the use of AES
- the potential impacts of AES on assessment and workload practices which requires adequate professional development resources
- competing interests and values between schools, departments, and institutions associated with using AES
- how the use of AES relates to and integrates with broader policy frameworks.

The investigation of these issues requires information sharing, dialogue, and negotiation among diverse stakeholders, including teachers, parents, students, leaders, and policymakers.

In addition to this engagement, schools and other educational institutions must also discuss the implementation of AES tools with AES system developers and commercial vendors, so as to better understand the functions and limitations of the AES tool, as well as its implications for professional and assessment practices. Only then can decision-makers evaluate whether a specific AES system is worth the investment of funds and resources, including teacher workload, in both the medium and longer term.

Although it appears as yet another drag on teacher time, the participatory and collaborative development of AES guidance, policy, and regulation is crucial. It ensures that pluralistic views and shared values are reflected in any innovations or reforms across the education sector. To ensure a collaborative foundation, the introduction of AES must be informed by stakeholder expertise across multiple locations and decision-making levels, including classrooms, schools, organisations, and state, territory, and national jurisdictions. **For Australia, we recommend multi-scalar policy development informed by educators, policymakers, and representatives from educational technology companies engaging in cooperative learning and action.**

When is AES not recommended?
The use of AES is not recommended for high-stakes tests in schools in most instances.

Where high stakes are likely for students, teachers, principals, or school communities, how AES works and the ethical implications of its use must be clearly explained.

Summary infographic

The key issues, takeaways, critical questions, and recommendations from this white paper are outlined here. These insights can support diverse stakeholders to navigate the interrelated social, technical, ethical, and political dimensions of AES systems in high-stakes education contexts.

01

Issue 1: AES system complexity and contexts

Key takeaway

More learning about AES system complexity and contexts is needed.

Critical questions

Has the AES vendor been transparent about any immediate or longer-term social and ethical impacts?

Are you aware of publicly available learning tools that would help people to understand AES systems?

Recommendations

Make learning tools that enable diverse stakeholders to understand how AES systems work.

02

Issue 2: School infrastructure capacity to deploy AES

Key takeaway

The range of infrastructure and capacity-building to fully support AES across Australian schools is underexplored.

Critical questions

Is there adequate resourcing, internet reliability, and labour to maintain the digital infrastructure and introduced system?

Do all people in your school or organisation have access to the required technology and skills to implement AES?

Recommendations

Identify the digital infrastructure and skills required to support the use of AES across urban, regional, and remote schools.

03

Issue 3 : Impact of AES upon professional practice

Key takeaway

Assessment and workload practices are being displaced with AES, so there is a need to expand teachers' socio-technical expertise and to build co-designed systems.

Critical questions

Does your school or organisation provide opportunities to discuss the positive and negative impacts of new technology, such as AES, on professional practices?

Would you attend professional development that provided opportunities to learn about and experiment with automated technologies like AES?

Recommendations

Prioritise professional development and co-designed AES systems which value, and build upon, teachers' judgement and socio-technical expertise.

04

Issue 4 : Cross-sectoral interests and values associated with AES

Key takeaway

Opportunities to explore multiple stakeholder interests and values about AES are currently lacking.

Critical questions

Do you know the details of why, when, and where an AES system is introduced?

Who decides if and how an AES system is introduced into your jurisdiction, organisation or school?

Recommendations

Provide opportunities for sharing knowledge and decision-making about the use of AES between diverse stakeholders.

05

Issue 5 : Policy uncertainty regarding AES and emerging EdTech

Key takeaway

Existing policies are not keeping pace with rapid technological change, such as AES, in Australian schools.

Critical questions

Do you know what policies frame the introduction and use of AES or other education technologies in your jurisdiction, organisation, or school?

What avenues are there for appeals, or new approaches, to be made about AES decisions and systems?

Recommendations

Connect and integrate policies for the use of AES in high-stakes education contexts.

Glossary

ACARA: Australian Curriculum and Reporting Authority.

Artificial intelligence (AI): An autonomous, or semi-autonomous computer system that employs algorithms to learn from patterns in large data sets in order to improve predictive abilities. (See also machine learning).

Assessment: A process of gathering information and using observation to judge the progress of students. Tests as a measuring tool are one part of assessment.

Automated essay scoring (AES): A psychometric-based form of digitalised education testing that integrates algorithm models with essay datasets in order to score student writing according to specific features or criteria.

Black boxes: In automated systems that use artificial intelligence, a black box system allows someone to see the input or output but does not allow a view of what happens in between. If an AES system uses deep learning it is considered a black box as there is no way to know exactly how the system makes a decision and provides a score.

Deep learning (also known as unsupervised learning): A form of machine learning where a computer is enabled to predict and classify information without human input.

Education technology (EdTech): Technology used in a range of education areas, including administration and teaching and learning. EdTech is most commonly associated with commercial products.

High-stakes test: A test that carries "serious consequences for students or for educators". This may encompass the decision to pass or certify a particular individual or the ranking of institutions based on cohort results. In such a test, high scores "may bring public praise or financial rewards; low scores may bring public embarrassment or heavy sanctions"[1].

Hybrid forum: A form of consultation involving stakeholders with diverse expertise that focuses upon collective learning and experimentation in response to a particular socio-technical controversy.

Machine learning: A form of artificial intelligence that uses algorithms to make predictions from data.

NAPLAN: National Assessment Program – Literacy and Numeracy.

Natural language processing: The capacity of a computer trained to understand spoken and written human language.

Socio-technical controversy: Controversies that involve both social and technical dimensions. Examples include nuclear power, urban planning and the use of automated technologies like artificial intelligence.

Supervised learning: An approach to machine learning where a computer algorithm is trained by a human on input data that has been labelled for a particular output.

1 American Educational Research Association [AERA] (2000). Position statement on high-stakes testing. https://www.aera.net/About-AERA/AERA-Rules-Policies/Association-Policies/Position-Statement-on-High-Stakes-Testing

Global developments indicate that AES is likely to become attractive to education system leaders in Australia in coming years.

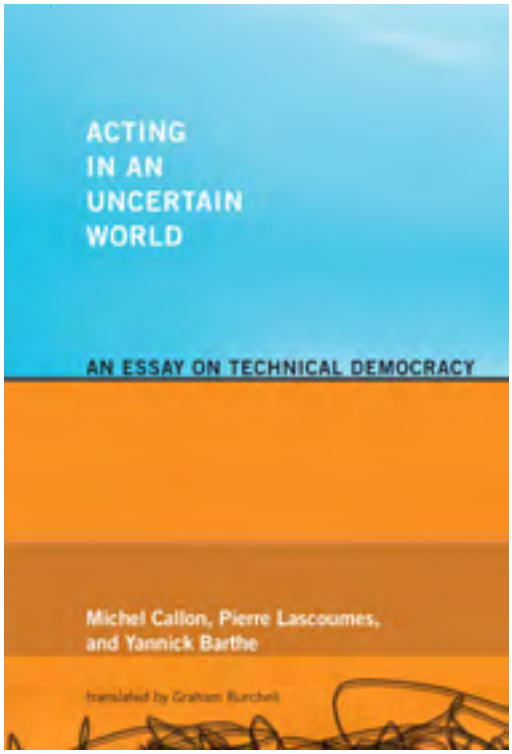
Controversy as method

On 26 July 2022, 19 academic and non-academic stakeholders, including psychometricians, policy scholars, teachers, union leaders, system leaders, and computer scientists, gathered at the University of Sydney to discuss the use of Automated Essay Scoring (AES) in education, especially in primary and secondary schooling. The combined expertise of this group spanned: digital assessment, innovation, teaching, psychometrics, policy, assessment, privatisation, learning analytics, data science, automated essay feedback, participatory methodologies, and emerging technologies (including artificial intelligence and machine learning).

The workshop adopted a technical democracy approach[2] and was held as a hybrid forum, a methodology that has proved useful in response to specific socio-technical controversies[3]. A hybrid forum uses collective learning and experimentation to allow for greater amplification of diverse perspectives and ideas than more traditional policy forums. One reason for adopting this approach was to explore the potential of controversy to inform future research, practice, and policy around technology in education.

Prior to the forum, participants were sent an information pack with links and resources, including an overview of the AES workflow, information about how machine learning works, and news articles about the controversial trial of "robot marking", which was opposed by sections of the education community, such as teachers' unions, as a "direct attack on teaching"[4]. The information pack deliberately amplified rather than assuaged diversity of opinion.

During the forum, participants collectively experimented with the black box of an AES system, including key features used in tests, such as marking rubrics for narrative and persuasive writing. These experiments opened up space to discuss issues that might otherwise have been seen as the territory of a particular professional or expert group. Participants discussed intersecting issues that emerged from the collective learning, rather than speaking as isolated stakeholders or experts. Exploring the complex space of professional and disciplinary knowledges and experiences, the hybrid forum provided a unique opportunity: to not only expand understanding of problems and concerns arising from the use of AES, but also to identify possible strategies for influencing policy development in the area. Key insights from the day inform this white paper. The paper itself was written by nine participants, but all forum participants were invited to review the key insights and test the critical questions prior to finalisation.



2 Callon, M., Lascoumes, P. & Barthe, Y. (2009) *Acting in an uncertain world: An essay on technical democracy.* Translated by Graham Burchell. MIT Press.

3 Thompson, G., Gulson, K.N., Swist , T. & Witzenberger, K. (2022). Responding to sociotechnical controversies in education: a modest proposal toward technical democracy. *Learning, Media and Technology.* doi: 10.1080/17439884.2022.2126495

4 New South Wales Teachers Federation [NSWTF] (2017). Opposition to NAPLAN Online and robot marking grows. https://news.nswtf.org.au/blog/news/2017/12/opposition-naplan-online-and-robot-marking-grows

Background to Automated Essay Scoring (AES)

Education technology is a rapidly growing industry of which AES is a key part. AES uses methods from machine learning to classify, score or rank essays. In 2018, the use of AES in NAPLAN caused significant controversy in the Australian school sector.

Over the past 20 years, the use of education technology (EdTech) in schools and education systems around the world has grown rapidly. EdTech includes products that support administrative tasks like monitoring attendance or behaviour, business applications such as word processing and spreadsheets, and products that aim to support student learning and track progress through schooling. Some products are supported by artificial intelligence (AI), which are systems that use machines to make predictions and recommendations that can inform human decision-making. These systems aim to supplement teaching practices, increase access for students in remote areas, and improve system-level information management.

The automation of grading and scoring goes by various names, such as "machine marking", "automated essay grading", and "automated essay scoring". We use "automated essay scoring" (AES) to describe any technology that uses methods from machine learning to classify, score or rank essays[5]. Where the technology provides formative rather than summative feedback, we identify this as "automated essay feedback". Systems that provide formative feedback have shown potential in higher education contexts. These systems highlight key patterns or textual features and provides the student with information on how an essay meets or fails to satisfy the expectations of a discourse community[6]. For a general overview of the different approaches to AES and their various strengths and weaknesses, we encourage the reader to consult the systematic review undertaken by Ramesh and Sanampudi[7].

Current international landscape

The adoption of AES is increasing globally. In some countries, such as Switzerland and Germany, pilot programs were trialled before AES was considered for widespread use[8]. Notably, AES is used extensively in countries at the forefront of EdTech development, such as China and the United States (US). The rise of Chinese AES systems provides a counter to the predominance of AES products that score writing for European languages. For present purposes, it is instructive to look at the use of AES in the US as it shows the impact of AES in high-stakes tests in a context that is comparable to the Australian one.

The distinction between the use of AES in high- and low-stakes tests is an important one. A high-stakes test has serious consequences for the student and/or educators. This can include effects on progress to the next stage of schooling, entry into higher education or financial consequences for a school. A low-stakes assessment has less consequence and is seen as playing a formative role for the test-taker, such as providing feedback or exam practice.

5 Ke, Z., & Ng, V. (2019). Automated essay scoring: A survey of the state of the Art. *IJCAI*, 19, 6300–6308. https://www.ijcai.org/proceedings/2019/0879.pdf

6 Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sándor, Á., Kitto, K., Liu, M., Vijay Mogarkar, R., & Buckingham Shum, S. (2020). AcaWriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research*, 12(1), 141-186. https://doi.org/10.17239/jowr-2020.12.01.06

7 Ramesh, D., & Sanampudi, S. K. (2021). Automated essay scoring systems: a systematic literature review. *Artificial Intelligence Review*, 1-33. https://link.springer.com/article/10.1007/s10462-021-10068-2

8 Rupp, A. A., Casabianca, J. M., Krüger, M., Keller, S. & Köller, O. (2019). Automated Essay Scoring at Scale: A Case Study in Switzerland and Germany. *ETS Research Report Series*, 1-23. https://doi.org/10.1002/ets2.12249

'A direct attack on teaching': NSW rules out NAPLAN robo marking

https://www.smh.com.au/education/a-direct-attack-on-teaching-nsw-rules-out-naplan-robo-marking-20171207-h00jle.html

In the US, automated marking has a long history in educational testing. In 1964, Ellis Page, widely regarded as a historical pioneer in automated marking, developed an early machine-scoring algorithm that demonstrated high inter-rater reliability and was purportedly able to better assess creativity in writing than human judges[9]. Since 2001, e-rater, a machine scoring system developed by the Education Testing Service, a non-profit organisation, has been used to make high-stakes decisions around graduate school entry in US colleges and universities[10]. At the time of writing, a version of e-rater is still used to assess the quality of the analytical writing section of the Graduate Record Examinations. Similarly, Pearson Online Learning Services[11] uses automated marking in their Test of English Academic and English Benchmark assessments. This is just one example of the way the landscape of AES provision is becoming dominated by EdTech.

Why marking essays by algorithm risks rewarding the writing of 'bullshit'

https://theconversation.com/why-marking-essays-by-algorithm-risks-rewarding-the-writing-of-bullshit-85910

Smarter Balanced, a high-stakes assessment tool used in 17 states in the US, uses automated scoring to measure the attainment of state standards alongside human validation for interim and summative assessments[12]. In the case of interim assessments, the student scores generated by Smarter Balanced are accompanied by a confidence level, which, if too low, triggers a review by the teacher. In the case of summative assessments, low score confidence levels trigger a review by a professional scorer. Recently, the US Institute for Education Statistics adjudicated an automated scoring challenge on a key test assessment among automated marking companies[13]. In this context, the AES systems judged best were alike in using natural language processing models to successfully identify language patterns in student responses.

In 2018, the Programme for International Student Assessment (PISA) implemented automated marking for short-answer responses[14]. PISA has proposed the use of neural networks to code graphical responses in the Problem Solving and Inquiry section of Trends in International Mathematics and Science Study (TIMSS).

Global developments indicate that AES is likely to become attractive to education system leaders in Australia in coming years.

9 Page, E. B. (1968). The use of the computer in analysing student essays. *International Review of Education, 14*, 210-225. https://doi.org/10.1007/BF01419938

10 Herrington, A. & Moran, C. (2001). What happens when machines read our students' writing? *College English, 64*(4), 480-499. https://doi.org/10.2307/378891

11 Pearson Online Services. (2021). https://www.pearson.com/en-au/educator/university/pearson-online-learning-services

12 Smarter Balanced (2022). https://smarterbalanced.org

13 National Centre for Education Statistics [NCES] (2022). https://nces.ed.gov/surveys/annualreports/overview

14 Organisation for Economic Co-operation and Development [OECD] (2018). Coding design, coding process, coding reliability studies, and machine-supported coding in the main survey. 2018 Technical Report. https://www.oecd.org/pisa/data/pisa2018technicalreport/PISA2018-TecReport-Ch-13-Coding-Reliability.pdf

While AES use in NAPLAN in 2018 was controversial, the international experience suggests that its appeal to education system leaders and politicians will continue to increase. Based on the controversies that have arisen around previous AES rollouts in the US, Canada, and Australia, US researchers have made several recommendations to enhance public communication around machine scoring and its strengths and limitations. This includes a phased implementation process that spans design, piloting, and review activities[15]. Greater public and professional understanding of how AES is adapted and applied in practice may help Australian researchers, educators, AES vendors, and policymakers navigate the challenges and opportunities associated with its implementation.

Controversy – an Australian case study

In Australia, aside from previous trials and applications in the higher education sector, AES became known in the school sector after the controversy surrounding NAPLAN in 2018. NAPLAN is a standardised test administered by the Australian Curriculum, Assessment and Reporting Authority (ACARA). Since 2008, it is taken by all students in years 3, 5, 7, and 9. Initially promoted to the schooling sector as a diagnostic test to aid in student learning, NAPLAN quickly became a high-stakes assessment when the publication of results became a means to compare school performance across Australia.

The controversy about the use of AES in NAPLAN began in 2016 when ACARA announced that it would shift from being a "pencil and paper" test to something delivered online. This transition was connected to the adoption of Automated Essay Scoring. From the beginning, teacher unions in multiple states and territories opposed the introduction of AES[16].

In 2017, the NSW Teachers Federation (NSWTF) commissioned its own research by Professor Les Perelman, an AES scholar from the US. Perelman challenged a report from ACARA that recommended the use of AES in NAPLAN. Perelman argued that AES should not be used in tests like NAPLAN since further evidence was required across several areas, including pilots to ensure AES does not discriminate against any groups[17]. NSWTF also conducted research that indicated that NSW schools had vastly different capacities to conduct high-stakes testing involving AES[18]. In December 2017, after sustained opposition from the NSWTF and other teachers' unions, the Education Council at the time comprised of all state and territory ministers, decided to halt the proposal to use AES to mark NAPLAN[19] .

15 Shermis M. D. & Lottridge, S. (2019). Communicating to the public about machine scoring: what works, what doesn't. Paper presented at the annual meetings of the National Council of Measurement in Education https://www.air.org/sites/default/files/CSSC_Communicating%20with%20the%20Public_White%20Paper.pdf

16 New South Wales Teachers Federation [NSWTF] (2017). Opposition to NAPLAN Online and robot marking grows. https://news.nswtf.org.au/blog/news/2017/12/opposition-naplan-online-and-robot-marking-grows

17 Ibid.

18 NNew South Wales Teachers Federation [NSWTF] (2017). 2017 NAPLAN Readiness Trial. The Response of the Teaching Profession. https://www.nswtf.org.au/pages/2017-naplan-online-readiness-trial-response-teaching-profession

19 Robinson, N. (2018) NAPLAN: Robot marking of school tests scrapped by education ministers. 29 January, ABC News. https://www.abc.net.au/news/2018-01-29/push-to-have-robots-mark-naplan-tests-scrapped/9370318

Issues relating to AES in Australia

Issue 1: AES system complexity and contexts

Issue 1: AES system complexity and contexts

AES systems within academic and corporate sectors continue to innovate as research and development progresses. The rapid integration of AES across schooling sectors requires greater understanding of how AES systems work and the contexts in which they are applied.

AES system complexity

There are a number of different approaches to AES that are used globally. To date, state-of-the-art solutions tend to employ supervised learning where a corpus of essays already marked by humans is used as a training data set. The process of training and then using an AES system is outlined in Figure 1.

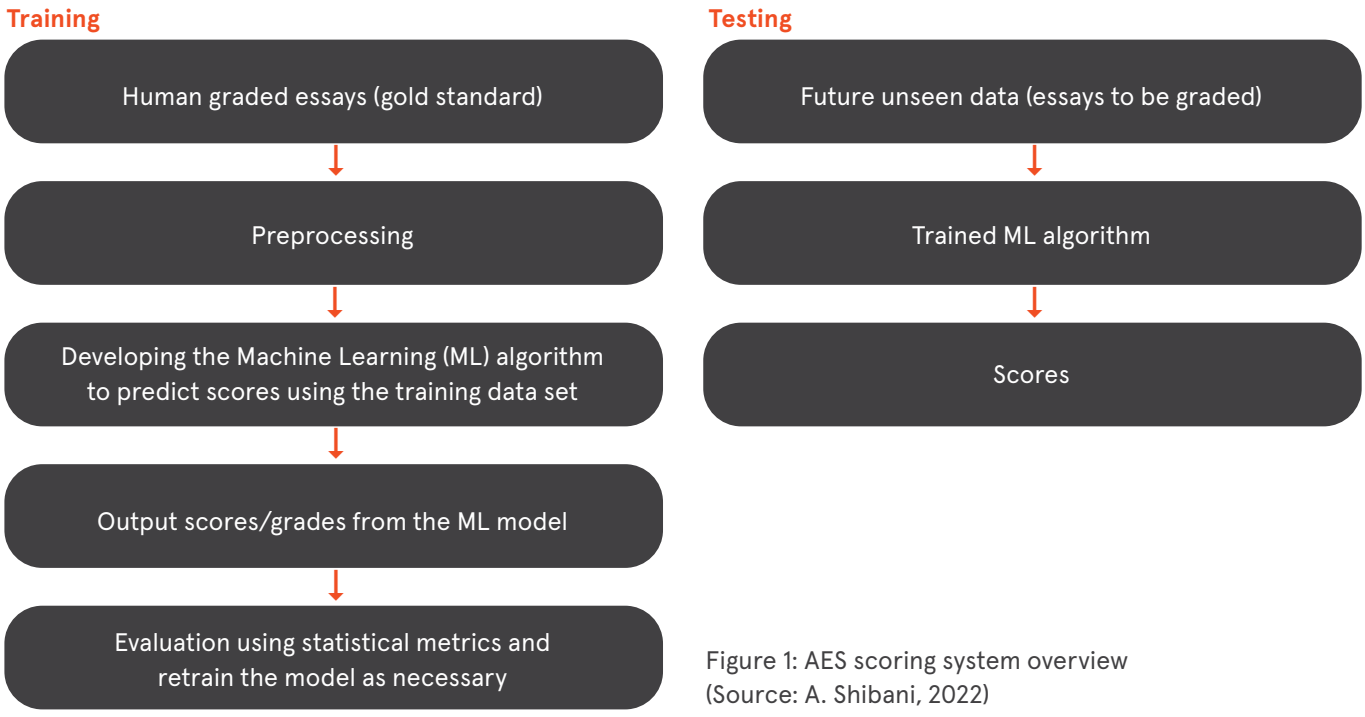


Figure 1: AES scoring system overview
(Source: A. Shibani, 2022)

Several factors influence the accuracy of the scores, classifications, or rankings produced. First, the range of different training sets available, which may include essays written by different age groups, or by writers with different language competencies. These variables all affect how the machine learning algorithm learns to score essays. Second, different types of algorithms produce different scores. This is because different algorithms make use of different features in a training data set. While supervised learning models require features to be chosen (e.g., the length of sentences), state-of-the-art approaches increasingly rely on deep learning methods which extract features automatically. If deep learning algorithms are used, there is a loss in transparency as it is difficult to understand the precise characteristics that led to a score. This is known as a black box system.

Most AES applications are proprietary products provided by a range of EdTech companies. The use of proprietary EdTech raises concerns about independent evaluation of their functionality and impact, including potential harms to different populations[20]. In addition, when EdTech uses black box systems it exacerbates concerns about the transparency of AES systems and undermines trust in the scores produced.

Because of these factors, there is a need for users to be able to audit and assess systems in relation to specific testing contexts. EdTech companies rarely provide this type of information. One way to secure this is through "algorithmic audits"[21]. An algorithmic audit allows users to check whether a system is working in the ways that are claimed and to identify the outputs produced (e.g., scores on essays). These audits may help stakeholders understand how data is analysed, and collated (e.g., what algorithms are used) and the modelling used in the AI system, as well as insight into the strengths and weaknesses of different models. However, while algorithmic audits are an option for an education department, the technical expertise required to undertake these is often beyond the purview of schools.

While the information provided by an algorithmic audit requires a high level of technical literacy, it is more appropriate for education departments to undertake them in collaboration with technical experts and schools so that AES systems can be examined from multiple perspectives.

An additional option is to raise awareness of AES via automated writing evaluation systems developed in-house by universities[22]. These in-house systems provide substantial information on how they are designed, developed, and deployed.

Key Takeaway

More learning about AES system complexity and contexts is needed

20 Williamson, B. (2019). New power networks in educational technology. *Learning, Media and Technology* 44(4), 395-398. https://doi.org/10.1080/17439884.2019.1672724

21 Metaxa, D., Park, J.S., Robertson, R. E., Karahallos, K., Wilson, C. Hancock, J. & Sandvig, C. (2021). Auditing algorithms: understanding algorithmic systems from the outside in. Foundations and Trends in *Human-Computer Interaction, 14*(2), 272-344. http://dx.doi.org/10.1561/1100000083

22 Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., Wight, R., Lucas, C., Sándor, Á., Kitto, K., Liu, M., Vijay Mogarkar, R., & Buckingham Shum, S. (2020). AcaWriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research*, 12(1), 141-186. https://doi.org/10.17239/jowr-2020.12.01.06

Issues relating to AES in Australia

Issue 2: School infrastructure and capacity to deploy AES

Issue 2: School infrastructure and capacity to deploy AES

The implementation of AES depends on reliable digital infrastructure and new sets of skills for various users. There continues to be inequality in the provision of digital infrastructure and capacity-building opportunities across Australian schools.

Infrastructure

Within the field of information technology, infrastructure is commonly understood to encompass structures on which networked technologies operate. This includes physical elements, such as computers or internet connectivity, as well as human elements, such as expertise[23]. In addition to technical infrastructure, a high level of organisational capacity within schools or education departments is necessary to ensure the stability and functionality of AES. This includes, for instance, those responsible for administering tests within authorities like ACARA. It is important that administration authorities assume the responsibility of overseeing the implementation of AES systems, and proactively remedy problems with inadequate infrastructure in schools before AES programs are rolled out.

Inadequate infrastructure can have negative consequences in high-stakes assessment environments. For example, poor digital infrastructure can lead to delays in software loading or other impediments for students in accessing and completing tests delivered online[24]. This may lead to penalties if the test is time-sensitive, or the time taken to answer is included in the score. Depending on the chosen AES method, the application of penalties may not be transparently reported (e.g., in a black box system it may be impossible to see which features are used to determine a score).

Reliability of internet

Many schools in Australia lack a sufficiently reliable broadband internet connection to enable the effective use of EdTech like AES. The government-funded National Broadband Network (NBN) that began to be rolled out in Australia in 2011 has been described as slow in regional areas[25]. Furthermore, internet provision in Australia is delivered through fibre, fixed wireless, and satellite, with significant differences in quality and speed between all three. Broadband access and bandwidth in schools is a critical problem disproportionately affecting public schools, schools in low-income communities, regional, rural, and remote schools, and in particular, schools with high numbers of Aboriginal and Torres Strait Islander students[26].

23 Sellar, S. (2017). Making network markets in education: the development of data infrastructure in Australian schooling. *Globalisation, Societies and Education*, *15*(3), 341-351. https://doi.org/10.1080/14767724.2017.1330137

24 Cavanagh, S. (2018). *Tennessee, Measurement Incorporated Tussle Over Blame for Testing Woes*. EdWeek Market Brief https://marketbrief.edweek.org/marketplace-k-12/tennessee-measurement-incorporated-tussle-testing-woes/

25 Randell-Moon, H. E. K., & Hynes, D. (2022). 'Too smart': Infrastructuring the Internet through regional and rural smart policy in Australia. *Policy & Internet*, *14*(1), 151-169. https://doi.org/10.1002/poi3.286

26 Flack, C. B., Walker, L., Bickerstaff, A., Earle, H., & Johnson, C. L. (2021). *Principal perspectives on the impact of COVID-19: Pathways toward equity in Australian schools*. Pivot Professional Learning. https://broadcast.schooltv.me/sites/default/files/2021-08/Pivot_Principal%2BPerspectives%2Bon%2Bthe%2BImpact%2Bof%2BCOVID-19%2BWhitepaper_February%2B2021_1.pdf

'2017 NAPLAN Online Readiness Trial: The Response of the Teaching Profession'

https://www.nswtf.org.au/files/2017_naplan_online_readiness_trial_-_the_response_of_the_teaching_profession_-_final.pdf

Accessibility

Accessibility to digital devices in schools is essential for everyday teaching and learning, as well as testing. However, outdated hardware in schools continues to be a significant challenge. In 2008, the Digital Education Revolution National Partnership, an initiative of the federal government under former Prime Minister Kevin Rudd, announced $807 million in federal funding to contribute to the installation and maintenance of computers in Australian schools. Over a decade later, schools still experience unequal access to digital hardware, despite many adopting Bring Your Own Device (BYOD) protocols or "1:1" schemes that seek to transfer the cost of the provision and maintenance of digital devices to parents[27]. There is a significant "digital divide" among Australian schools, with a gap in technology access between wealthy and disadvantaged schools (particularly rural and remote schools) as well as the skills and knowledge required to navigate technology[28].

Key Takeaway

The range of infrastructure and capacity-building required to fully support AES across Australian schools is underexplored.

27 Maher, D., & Twining, P. (2017). Bring your own device–a snapshot of two Australian primary schools. *Educational Research, 59*(1), 73-88. doi: 10.1080/00131881.2016.1239509

28 Heffernan, A., Magyar, B., Bright, D., & Longmuir, F. (2021). *The impact of COVID-19 on perceptions of Australian Schooling*. Monash University. https://www.monash.edu/education/research/downloads/Impact-of-covid19-on-perceptions-of-Australian-schooling.pdf

Issues relating to AES in Australia

–

Issue 3: Impact of AES on professional practice

Issue 3: Impact of AES on professional practice

The impact on assessment and professional practices must be identified by education departments and boards before introducing AES. Training opportunities should be provided for teachers to develop new skills in judging whether AES systems can and should be used. Ideally AES systems should be co-designed with a range of direct and indirect users, including those who depend on the outputs and those who support the system.

The displacement of professional judgement and workload

The expertise required to understand and use AES systems can unsettle established routines of professional judgement and workload among teachers and administrators. Thompson and colleagues[29] define displacement as the phenomenon in which people who are impacted by complex technical systems are often denied opportunities to develop a critical understanding of them or contribute to design and implementation decisions. Selwyn and colleagues[30] found that the substitution of teacher judgement with automated systems was commonly seen to undervalue teacher expertise and professionalism. Historically, a lack of opportunities for teachers and school communities to understand and meaningfully contribute to the oversight of complex testing tools has had significant implications for both the implementation and effectiveness of AES tools.

The implementation of AES in conjunction with standardised testing may lead to additional complications and workloads for teachers, including pressure to teach how to write to meet the criteria on which the marking algorithm is based (e.g., what is valued in writing skills, structural features, or higher-order thinking). This is called "working to the algorithm", whereby teachers must teach students to write in ways that are "readable" by the machine[31]. This criticism finds parallels in those of standardised testing, which, according to its opponents, requires that teachers 'teach to the test' rather than teaching a broad, dynamic range of writing styles.

Stakeholders in education need to understand the competing values associated with AES systems and their implications. For example, valuing the efficiency of using machine-learning can take precedence over teacher expertise. Another issue to consider is the contrast in what EdTech vendors value when trying to sell products versus the needs of education departments and schools. With the overwhelming emphasis on EdTech products and the technical knowledge involved in developing and applying AES, there is growing concern that teachers' professional knowledge and experience may be side-lined.

29 Thompson, G., Rutkowski, D., & Sellar, S. (2018). Flipping large-scale assessments: bringing teacher expertise to the table. In D. M. Netolicky, J. Andrew, C. Paterson (Eds.). *Flip the System Australia: What Matters in Education*, Chapter 6 (pp. 55-63). Routledge, USA.

30 Selwyn, N., Hillman, T., Bergviken Rensfeldt, A., & Perrotta, C. (2021). Digital Technologies and the Automation of Education—Key Questions and Concerns. *Postdigital Science and Education*, 1-10. https://doi.org/10.1007/s42438-021-00263-3

31 Ibid, 28

Expanding socio-technical expertise

The displacement of assessment expertise in schools, alongside concerns around de-professionalisation, are additional potential impacts of AES systems. New expertise is needed to understand the technical aspects of AES, as well the social aspects of decisions made by algorithms. Professional development aimed at "algorithmic literacy" may help empower educators to develop the necessary expertise required to critically evaluate new technologies[32]. AI literacy can encompass multiple objectives, including the understanding, use, evaluation, or creation of AI, as well as an understanding of the ethical questions posed by AI[33]. Developing AI literacy requires a multifaceted approach, including a focus on 'learning artefacts' to support the development of technological knowledge, collaborative learning for pedagogical knowledge, and content knowledge such as ethics[34]. Proper funding and resourcing of professional development will be vital in expanding socio-technical expertise and facilitating collective learning and experimentation among school communities and teachers.

Co-designed AES models for collective learning and experimentation

Currently, there is a significant mismatch between what EdTech products promise and the difficulties encountered by users in their effective implementation, maintenance, and repair[35]. Involving diverse stakeholders in the co-design of AES systems could help minimise this gap through collective learning and experimentation. At the simplest level, learning about the failures and successes of previous systems could shape the co-design of new models in a process known as "human-machine teaming", which involves practitioners and machines working together to communicate, coordinate, and adapt tasks[36].

Ideally, this co-design process would assist the development of AES prototypes that value and integrate teacher expertise. For instance, the machine focuses on specific criteria, such as spelling or sentence structure, while the teacher focuses on higher-order processes like creativity. AES human-machine teams could be built to integrate expertise and address needs across multiple scales: individual assessment, school-level workload, and education governance. To achieve this, teams of educators, students, leaders, and administrators would all need to be engaged in the co-design process.

Key Takeaway

Assessment and workload practices are being displaced with AES, so there is a need to expand teachers' socio-technical expertise and to build co-designed systems.

'NAPLAN: Differences do exist between online and pen-and-paper test results, document reveals'

https://www.abc.net.au/news/2018-10-18/differences-exist-naplan-online-and-pen-and-paper-results/10388156

32 Ciccone, M. (2021). Algorithmic literacies: K-12 realities and possibilities. *Algorithmic Rights and Protections for Children* https://wip.mitpress.mit.edu/pub/algorithmic-literacies/release/1?readingCollection=646d0673

33 Ng, D.T.K., Leung, J.K.L., Chu, K.W.S & Qiao, M.S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: *Artificial Intelligence, 2.* https://doi.org/10.1016/j.caeai.2021.100041

34 Australian Data Science Education Institute (nd). Why Data Science? https://adsei.org/why-data-science/

35 Pangrazio, L., Selwyn, N., & Cumbo, B. (2022). A patchwork of platforms: Mapping data infrastructures in schools, Learning, Media and Technology. doi: 10.1080/17439884.2022.2035395

36 Stowers, K. Brady, L. L., MacLellan, C. Wohleber, R., & Salas E. (2021). Improving teamwork competencies in human-machine teams: Perspectives from team science. Frontiers in Psychology 12, 590290. https://doi.org/10.3389/fpsyg.2021.590290

Issues relating to AES in Australia

Issue 4: Cross-sectoral interests and values associated with AES

Issue 4: Cross-sectoral interests and values associated with AES

Varying assessment priorities, cultures and student demographics can influence unequal automated scoring outcomes. Most AES systems are proprietary products, which raises issues related to data ownership, privacy, and commercialisation.

Varying test designs, values, and student backgrounds AES can be applied across different assessment contexts including those that span both low- and high-stakes testing applications. For example, a standardised test that uses similar criteria for comparison becomes high-stakes when "the scores of the test are used to make important decisions about a student's educational pathway (e.g., graduation, grade promotion, educational track) or about other areas like school funding, teacher or principal pay, and staff performance"[37]. Where low-stakes assessment tends to focus upon more incremental and formative tests, the use of AES in a high-stakes test could lead to the modification of teaching and learning practices, a constrained curriculum, and a narrowing of students' educational experiences[38]. These outcomes can mean students are inadequately prepared in other skills for work and further education.

It is crucial to understand that AES instruments and applications do not exist in isolation but are shaped by assessment "micro-cultures" and "macro-cultures," from localised support of student learning to systemic measures of school performance, accountability, and funding[39].

Discussions about AES systems should be framed in relation to a spectrum of micro-macro assessment priorities and values.

Diverse student needs, arising from diverse cultural and linguistic backgrounds, must be accounted for when schools consider AES systems. For example, AES systems produce differences in scoring depending on the writing characteristics of different demographic groups due variations in the use of vocabulary and language[40].

37 Au, W. (2022). *Unequal by Design: High-Stakes Testing and the Standardization of Inequality*. Routledge, New York.

38 Polesel, J., Rice, S., & Dulfer, N. (2014) The impact of high-stakes testing on curriculum and pedagogy: a teacher perspective from Australia. *Journal of Education Policy,* 29(5), 640-657. https://doi.org/10.1080/02680939.2013.865082

39 DeLuca, C., Rickey, N., & Coombs, A. Sammy King Fai Hui (Reviewing editor) (2021). Exploring assessment across cultures: Teachers' approaches to assessment in the U.S., China, and Canada, *Cogent Education, 8*:1, doi: 10.1080/2331186X.2021.1921903fau

40 Feathers, T. (2019). Flawed Algorithms Are Grading Millions of Students' Essays. *VICE*, August 20. https://www.vice.com/en/article/pa7dj9/flawed-algorithms-are-grading-millions-of-students-essays

https://www.school-news.com.au/news/naplan-online-glitches-challenges-and-opportunities/

https://www.janison.com/customer-stories/naplan-online/

The use of commercial AES systems

Extensive research has identified issues and challenges associated with educational technologies that are developed and provided by private companies. In particular, researchers have raised concerns about issues of trust and transparency in regard to patents and intellectual property in proprietary AES technology[41]. Others have argued that commercial EdTech companies are solely motivated by profit and that algorithms do not provide sufficient transparency or accountability for automated decision making.[42,43,44]

Furthermore, there are concerns surrounding data collection and ownership[45], as well as accountability regarding data management, particularly as data breaches are a risk within EdTech companies[46]. According to Williamson, many data privacy statements are deliberately ambiguous so that data may be sold on to third parties or used to inform internal processes of research and development[47].

Key Takeaway

Opportunities to explore multiple stakeholder interests and values about AES are currently lacking.

41 Williamson, B. (2017). *Big data in education: The digital future of learning, policy and practice*. Sage.

42 Cohen, D. (2022). Any time, any place, any way, any pace: Markets, EdTech, and the spaces of schooling.*Environment and Planning A: Economy and Space* https://doi.org/10.1177/0308518X221084708

43 Macgilchrist, F. (2019). Cruel optimism in EdTech: When the digital data practices of educational technology providers inadvertently hinder educational equity. *Learning, Media and Technology, 44*(1), 77-86. https://doi.org/10.1080/17439884.2018.1556217

44 Wright, N., & Peters, M. (2017). Sell, sell, sell or learn, learn, learn? The EdTech market in New Zealand's education system–privatisation by stealth? *Open Review of Educational Research, 4*(1), 164-176. https://doi.org/10.1080/23265507.2017.136562

45 Willis, J. E., Slade, S., & Prinsloo, P. (2016). Ethical oversight of student data in learning analytics: A typology derived from a cross-continental, cross-institutional perspective. *Educational Technology Research and Development, 64*(5), 881-901. https://www.jstor.org/stable/45018695

46 Fouad, N. S. (2022). The security economics of EdTech: vendors' responsibility and the cybersecurity challenge in the education sector. *Digital Policy, Regulation and Governance* (ahead-of-print).

47 Williamson, B. (2017). Educating Silicon Valley: Corporate education reform and the reproduction of the techno-economic revolution. *Review of Education, Pedagogy, and Cultural Studies*, *39*(3), 265-288. Doi: 10.1080/10714413.2017.1326274

Issues relating to AES in Australia

–

Issue 5: Policy uncertainty regarding AES and emerging EdTech

Issue 5: Policy uncertainty regarding AES and emerging EdTech

AES is currently being used in Australia with little policy guidance or clarity at either state or federal levels of government. There is a lack of regulation or guidelines about recourse options and appealing automated scoring in high-stakes testing that has implications for future opportunities and pathways for students.

Education sector policy developments

In December 2018, a federal meeting of Australian state and territory education ministers announced that "robot marking" of student essays should not proceed[48]. While automated essay scoring did not occur, the rollout of the NAPLAN online infrastructure proceeded in 2018 via a private company that delivered scalable online tests and learning solutions that included "advanced marking tools" for auto-marking, semi-auto marking, and paper marking. Features included 'adaptive testing' (trained to adapt to a student's ability) and novel 'authoring possibilities' (to vary the design of various test formats – with interactivity, audio, video, and drawing possibilities[49]. Online tests have since been delivered to 1.2 million students in Australia in more than 12,000 schools. Notably, the answer to "Is automated marking used in NAPLAN online?" (one of the frequently asked questions on the NSW Department of Education website) is:

> *Automated marking was used in the paper NAPLAN test to mark numeracy, reading and conventions of language and remains in use for these purposes in NAPLAN online. The automated essay scoring system is not used to mark NAPLAN online writing tests. Instead, trained markers are employed to mark student writing.*[50]

Automated marking is now routine for some aspects of NAPLAN online, yet trained human markers still remain central for marking online writing tests. The delineation of which parts of digital infrastructure are automated, and which should prioritise human decision-making, is significantly underexplored especially in policy contexts. The Australian Government's National School Reform Agreement (2018-2023) includes several new policy initiatives, such as Online Formative Assessment, a National Unique Student Identifier for School Students, and Improving National Data Quality[51]. Where the future of AES fits into these policy priorities remains to be seen especially in relation to acceptable and non-acceptable levels of automation across school infrastructure and assessment. What is clear, however, is that there is a need for clearer governance and policy to be applied to AI technologies in education, such as AES.

48 Koziol, M., Singhai, P., & Cook, H. (2018). Plan for 'robot marking' of NAPLAN essays scrapped. The Examiner, January 29 https://www.examiner.com.au/story/5196827/plan-for-robot-marking-of-naplan-essays-scrapped/?cs=8

49 Janison (2022) NAPLAN Online allows educators to tailor essential teaching for Australian students. https://www.janison.com/customer-stories/naplan-online

50 New South Wales Department of Education (2022). Frequently asked questions. https://education.nsw.gov.au/teaching-and-learning/student-assessment/naplan-online/frequently-asked-questions

51 Australian Government (2022). The National School Reform Agreement. Department of Education. https://www.education.gov.au/quality-schools-package/national-school-reform-agreement

Emerging policy developments

In Australia, policy regarding the use of AI in education is underdeveloped compared to other countries and regions. To keep pace with policy trends internationally, a national framework should regulate the use of automated technologies, specifically AES and other AI systems. In contrast to other jurisdictions like the European Union and Canada, Australia has largely taken a soft-law approach to the regulation of AI and automated decision-making[52]. There are existing whole-of-economy frameworks that impact the development and use of AI, such as the federal *Privacy Act*[53], plus laws that apply to the use and sharing of data like the *Data Availability and Transparency Bill*[54].

To date, there have been a range of different policy approaches regarding AI and regulation. Some of these include:

- AI Action Plan: the goal of which to make Australia a leader in responsible and inclusive AI[55]
- Australian AI Ethics Framework: which guides businesses, governments, and organisations to design, develop and deploy AI responsibly[56]
- Human Rights and Technology Final Report: which recommends human rights impact assessments, regulatory sandboxes, and the creation of an AI Safety Commissioner to advise on AI and adjudicate consumer complaints[57]
- A review of digital economy regulation: facilitated by the Department of Industry, Science and Resources[58].

The lack of a coordinated policy approach to AI is a matter of concern as an overwhelming majority of Australians (96%) expect governments to regulate AI, and more than 68% have moderate-to-high confidence that government and regulatory agencies will govern AI[59]. The lack of sector-specific policy frameworks means there is widespread regulatory uncertainty in Australia, increasing the risk that AI may be misused including in education contexts. A range of governance and policy ideas from initiatives in other countries and regions have the potential to be applied to AES and emerging technology developments in Australia, as outlined in the policy brief which accompanies this white paper[60].

Key Takeaway

Existing policies are not keeping pace with rapid technological change, such as AES, in Australian schools.

52 Burton, T. (2022) Why it's important for Australian to get data regulation right. Financial Review, July 11. https://www.afr.com/politics/federal/why-it-s-important-for-australia-to-get-data-regulation-right-20220711-p5b0mq

53 Australian Government (1988) Privacy Act. Attorney-General's Department. https://www.ag.gov.au/rights-and-protections/privacy

54 Parliament of Australian (2022). Data Availability and Transparency Bill 2022.

55 Australian Government (2021) Australia's AI Action Plan: discussion paper. Department of Industry, Science and Resources. https://consult.industry.gov.au/australias-ai-action-plan-discussion-paper

56 Australian Government (2019). Australia's Artificial Intelligence Ethics Framework. Department of Industry, Science and Resources. https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework

57 Australian Human Rights Commission [AHRC] (2021). Human Rights and Technology Final Report. https://humanrights.gov.au/our-work/rights-and-freedoms/publications/human-rights-and-technology-final-report-2021

58 Australian Government (2022). Positioning Australia as a leader in digital economy regulation (automated decision making and AI regulation): issues paper. Department of Industry, Science, and Resources. https://consult.industry.gov.au/automated-decision-making-ai-regulation-issues-paper

59 Curtis, C., Gillespie, N., & Lockey, S. (2020). Australians have low trust in artificial intelligence and want it to be better regulated. The Conversation, October 29. https://theconversation.com/australians-have-low-trust-in-artificial-intelligence-and-want-it-to-be-better-regulated-148262

60 Gulson, K., Thompson, G., Swist, T., Kitto, K., Rutkowski, L., Rutkowski, D., Hogan, A., Zhang, V., Knight, S. (2022). *Automated Essay Scoring in Australian Schools: Collective Policymaking* . Policy Brief, November 2022. Education Innovations Policy Brief Series ISSN 2653-6757 Sydney Social Sciences and Humanities Advanced Research Centre (SSSHARC), University of Sydney, Australia.

Critical questions stakeholders should ask

The following critical questions will support teachers, leaders and policymakers to address the issues identified in this white paper. These questions act as valuable prompts for collective action in three ways: i) to guide and support dialogue about AES systems; ii) to extract key information associated with the social, technical, and ethical dimensions of AES; and iii) to inform multi-scalar action with diverse stakeholders.

Issue 1: AES system complexity and contexts

Has the AES vendor been transparent about any immediate, or longer-term, social and ethical impacts?

Are you aware of publicly available learning tools which would help people to understand AES systems?

Issue 2: School infrastructure capacity to deploy AES

Is there adequate resourcing, internet reliability, and labour to maintain the digital infrastructure and introduced system?

Do all people in your school or organisation have access to the required technology and skills to implement AES?

Issue 3: Impact of AES upon professional practice

Does your school or organisation provide opportunities to discuss the positive and negative impacts of new technology, such as AES, on professional practices?

Would you attend professional development that provided opportunities to learn about and experiment with automated technologies like AES?

Issue 4: Cross-sectoral interests and values associated with AES

Do you know the details of why, when, and where an AES system is introduced?

Who decides if and how an AES system is introduced into your jurisdiction, organisation or school?

Issue 5: Policy uncertainty regarding AES and emerging EdTech

Do you know what policies frame the introduction and use of AES or other education technologies in your jurisdiction, organisation, or school?

What avenues are there for appeals, or new approaches, to be made about AES decisions and systems?

Recommendations

Issue 1: AES system complexity and contexts

Make learning tools that enable diverse stakeholders to understand how AES systems work.
These learning tools would surface AES system complexity, plus expose associated tensions and possibilities, to generate ongoing stakeholder dialogue. Examples of such potential tools include, flow diagrams, games, a common vocabulary, case studies, and collaborative auditing tools.

Issue 2: School infrastructure capacity to deploy AES

Identify the digital infrastructure and skills required to support the use of AES across urban, regional, and remote schools.
This recommendation focuses upon not only the technical readiness of schools to deploy AES, but also the skill-based readiness required to support digital and social inclusion. For example, this could be achieved through a national survey and mapping of infrastructure and capacity-building needs across urban, regional, and remote schools.

Issue 3: Impact of AES upon professional practice

Prioritise professional development and co-designed systems which value, and build upon, teachers' judgement and socio-technical expertise.
Such workforce strengthening and co-design strategies can help to ensure that teacher assessment and workload practices are augmented, and not displaced, with emerging technologies like AES. For example, the design of professional development which includes critical AI/algorithmic literacies and learning tools, plus funding the co-design of AES human-machine teams for beneficial communication and coordination of expertise.

Issue 4: Cross-sectoral interests and values associated with AES

Provide opportunities for sharing knowledge and decision-making about the use of AES between diverse stakeholders. These opportunities would inform cross-sectoral decision-making about where, when, and if, to use AES systems. This means not just when they are introduced, but also over the course of the product lifecycle. For example, this could be achieved via in-person hybrid forums (such as the method used for this study) and expanded to online forums as well.

Issue 5: Policy uncertainty regarding AES and emerging EdTech

Connect and integrate policies for the use of AES in high-stakes education contexts.
These governance developments would harmonise, where possible, with emerging national and international frameworks and initiatives. This could include the take-up of collective policymaking ideas enacted across multiple policy settings and scales. Please refer to our related policy brief for further details[61].

61 Ibid.

Conclusion

Global developments indicate that AES is likely to become attractive to education system leaders in Australia in coming years.

The introduction of AES in education must be informed by stakeholder expertise across multiple locations and decision-making levels. **Discussions about AES systems should be framed in relation to a spectrum of micro-macro assessment priorities and values.** These discussions should not only include policymakers and ministers across state, territory, and national jurisdictions but must recognise and incorporate the expertise of educators in classrooms and schools.

A cooperative process would ensure that diverse stakeholder expertise is integrated across education sector innovation and reforms, such as future AES developments. **While the information provided by such things as an algorithmic audit requires a high level of technical literacy, it is more appropriate for education departments to undertake them in collaboration with technical experts *and* schools so that AES systems can be examined from multiple perspectives and usages.**

Why is such cross-sectoral collaboration so urgent and vital? **AES cannot be approached from one dimension or scale** Educators, policymakers, and EdTech companies must work together **to frame the use of AES in schools as a multi-scalar issue with interrelated ethical, social, technical, and political implications.** While it is not recommended that AES be used in high-stakes testing, it is likely that AES will be widely adopted across time. There is now an opportunity for Australia to lead the way in the collective development of AES guidance, policy, and regulation.

Discussions about AES systems should be framed in relation to a spectrum of micro-macro assessment priorities and values.

Notes

THE UNIVERSITY OF SYDNEY

Sydney Social Sciences and Humanities Advanced Research Centre

:Education Futures Studio]

More information
If you would like further information about our Education Innovations White Paper Series or the Education Futures Studio, please contact:

Kalervo Gulson
kalervo.gulson@sydney.edu.au

Greg Thompson
g6.thompson@qut.edu.au

Teresa Swist
teresa.swist@sydney.edu.au

We acknowledge the tradition of custodianship and law of the Country on which the University of Sydney campuses stand. We pay our respects to those who have cared and continue to care for Country.

CRICOS 00026A