



Co-producing AIED Ethics Under Lockdown: an Empirical Study of Deliberative Democracy in Action

Teresa Swist^{1,2} · Simon Buckingham Shum³ · Kalervo N. Gulson¹

Accepted: 12 November 2023 / Published online: 27 February 2024
© The Author(s) 2024

Abstract

It is widely documented that higher education institutional responses to the COVID-19 pandemic accelerated not only the adoption of educational technologies, but also associated socio-technical controversies. Critically, while these cloud-based platforms are capturing huge datasets, and generating new kinds of learning analytics, there are few strongly theorised, empirically validated processes for institutions to consult their communities about the ethics of this data-intensive, increasingly algorithmically-powered infrastructure. Conceptual and empirical contributions to this challenge are made in this paper, as we focus on the under-theorised and under-investigated phase required for ethics implementation, namely, joint agreement on ethical principles. We foreground the potential of *ethical co-production* through Deliberative Democracy (DD), which emerged in response to the crisis in confidence in how typical democratic systems engage citizens in decision making. This is tested empirically in the context of a university-wide DD consultation, conducted under pandemic lockdown conditions, co-producing a set of ethical principles to govern Analytics/AI-enabled Educational Technology (AAI-EdTech). Evaluation of this process takes the form of interviews conducted with students, educators, and leaders. Findings highlight that this methodology facilitated a unique and structured co-production process, enabling a range of higher education stakeholders to integrate their situated knowledge through dialogue. The DD process and product cultivated commitment and trust among the participants, informing a new university AI governance policy. The concluding discussion reflects on DD as an exemplar of ethical co-production, identifying new research avenues to advance this work. To our knowledge, this is the first application of DD for AI ethics, as is its use as an organisational sensemaking process in education.

Keywords Educational technology · Learning analytics · Artificial intelligence · Ethics · Deliberative democracy · Organisational learning · Co-production · Socio-technical controversies

Extended author information available on the last page of the article

Introduction: COVID-19 and AAI-EdTech

The COVID-19 pandemic triggered the fastest transition to online learning in history, as globally, educational providers from primary to tertiary, to professional and lifelong learning, were forced to find online alternatives to their usual face-to-face practices. This is reflected by the growth of the educational technology industry in recent years, with according to one estimate, venture capital investment \$10.8B in 2022.¹ While all EdTech involves data, the driver behind this rise is the rapidly expanding field of EdTech in which the hallmark is a dependence on the automated processing of increasing quantities of data. We will refer to this broad class as *Analytics/AI-enabled educational technology (AAI-EdTech)*, combining techniques developed in academic fields including EdTech, AI in education (AIED), Learning Analytics, Educational Data Mining, and Learning@Scale.

Notwithstanding the pervasive marketing hype surrounding this, academic research pre-pandemic had already demonstrated that designed and deployed carefully, with appropriate educator training, AAI-EdTech can augment the capacity of teachers by providing close tracking of learners' activity and competence, to draw attention to learning difficulties and initiate actions of different sorts. Pre-COVID-19 examples of relatively mature deployments include, for instance, the state level piloting of a mathematics tutoring tool in schools, contributing to improved outcomes (Roschelle et al., 2016), or the embedding of predictive modelling into university student-support interventions, with improved outcomes (Herodotou et al., 2019).

Given the critical role that timely, personalised, motivating, actionable feedback plays in learning, in principle, the pandemic offered an opportunity for AAI-EdTech, as educators were stretched to the limit to maintain a quality student experience. Evidence is now emerging of how AI was used to support learning during the pandemic, including the papers in this special issue, and recent evidence from the authors' own university contexts documenting the value of data-driven, personalised, automated feedback in cultivating online students' sense of belonging, despite the difficulties of being locked-down, and in many cases, international time zones apart from the university (Lim et al., 2022; Thoeming et al., 2022).

Our specific interest in the pandemic-triggered emergency transition to online learning is the way it served also to exacerbate legitimate, longstanding concerns around the ethical implications of institutions using big data, analytics and AI to track and act on learner behaviour. To take one example, Williamson (2019) critiques the implications of the cloud infrastructures underpinning much EdTech being owned by large technology corporations. Another example was the contracting of automated exam invigilation services as a solution to conducting examinations when students were locked down at home. Such tools use several forms of AI, depending on the service, including image analysis to verify examinee identity, and audio, posture and facial analysis to detect potential cheating. Accounts are emerging of positive student experience of such tools as institutions

¹ HolonIQ <https://www.holoniq.com/notes/2022-edtech-vc-funding-totals-10-6b-down-from-20-8b-in-2021>

managed this process in what they considered to be an effective, ethical manner (Sefcik et al., 2022), while others have expressed concern over the ethical risks (Coghlan et al., 2021), and questioned whether the automated invigilation of exams is the best way to reliably maintain “assessment security” (Dawson, 2021). We return later to this example in the “ProctorU” case.

The three authors came together around a shared interest in methods broadly related to *co-production*: a mode of knowledge production which recognises the complex, dynamic interrelationship between knowledge, power, and society – plus its potential to explore complex challenges, as well as shift institutional and policy arrangements (Filipe et al., 2017; Jasianoff, 2004; Wyborn et al., 2019). While co-production has many approaches, it is commonly understood as a “multilevel phenomenon occurring at the level of socio-political systems, the level of institutions, and the level of situated practices” (Bandola-Gill et al., 2022, p. 2). Our shared interest was to better understand how this multilevel phenomenon could be applied in practice within the higher education context. One of the authors (Simon Buckingham Shum) had undertaken training in *Deliberative Democracy (DD)*, a movement that has emerged in response to the crisis in confidence in how typical democratic systems engage citizens in decision making. DD was adopted as a promising way to engage the university community with the ethical issues arising from AAI-EdTech.

Against this backdrop, this paper aims to make two contributions linking AAI-EdTech, DD and co-production. First, we investigate the practicality and value of using DD as a method to engage a broad range of university stakeholders, in order to genuinely co-produce ethical principles to govern the use of such technologies. Second, and more conceptually, we reflect on DD as a particular mode of co-production that can respond to the complex ethics and controversies raised by AAI-EdTech.

To do this, the structure of this paper is as follows. The next section provides an overview of three strands of work offering current responses at different levels of detail to ethical concerns: commitment to *ethical principles*, *trustworthy algorithms*, and *human-centred design*. Since institutions procuring external services may only have direct control over the first of these, we introduce the concept of ethics co-production, and DD specifically. We detail how DD’s practical value was tested through a university-wide ethics consultation and co-production process, and explain how it was evaluated via semi-structured interviews with students, educators, and leaders, from which the key findings are presented. The discussion then considers these findings in relation to co-production, which informs our conclusion and future research directions.

Background: Responses to AI Ethical Concerns

We begin by outlining three important strands of work in response to ethical concerns around AAI-EdTech: *ethical principles*, *trustworthy algorithms*, and *human-centred design*.

AI Ethics Principles

The first strand of work seeks to articulate principles that an institution, government or professional body could commit to, as an aid to apply AI ethics policy and governance. Such principles can serve as broad reference points to guide the entire process of AI design, from requirements, design, implementation, deployment and subsequent appeals/debate about the outcomes. The last five years or so has witnessed a proliferation of lists and taxonomies for AI ethics principles² (Floridi et al., 2018; IEEE, 2017) accompanied by public endorsements by every conceivable organisational entity. A few universities have published their principles for using student data and learning analytics,³ while to our knowledge, we have yet to see similar statements on AI ethics principles, which was an outcome from the process reported here.

However, in reviewing implementation of such principles in the technology industry, (Whittaker et al., 2018) (Sec.2.3) concluded that the evidence of their impact on the behaviour of computing companies was scarce, reinforced subsequently by others emphasising that ethics principles alone do not enforce regulation (Hagendorff, 2020; Rességuier & Rodrigues, 2020). A critical distinction has been made between whether we regard principles merely as a form of *deontological ethics* declaring what should be done, but at the risk of being reduced to virtue-signalling and toothless checklists, or more productively as some have argued, as guidance for *virtue ethics* that cultivate the ethical dispositions professionals need if they are to translate abstract principles into practical action when confronted by contextualised decisions (Hagendorff, 2020). This distinction is beginning to be thought through in educational contexts, with, for instance, Kitto and Knight (2019) arguing the merits of virtues ethics for Learning Analytics.

The second and third approaches described next concern the operationalisation of principles to implement the design process.

FATE: Towards Trustworthy Algorithms

In academia, education has become the focus of critical data studies and the interdisciplinary communities dedicated to the Fairness, Accountability, Transparency & Ethics (“FATE”) of algorithms. FATE-related research spans diverse spheres of society (Boyd & Crawford, 2012; Diakopoulos, 2014; Hanna et al., 2020; Holstein et al., 2019a), and is now focusing on how these issues manifest in education. The International Journal of AIED collated examples of contemporary thinking in a recently edited collection on “the FATE of AIED” (Porayska-Pomsta et al., 2021) extended by Holmes and

² Algorithm Watch: *AI Ethics Guidelines Global Inventory*: <https://inventory.algorithmwatch.org/>

³ Examples of university Learning Analytics ethics principles:
Edinburgh University <https://www.ed.ac.uk/files/atoms/files/learninganalyticsprinciples.pdf>
Michigan University <https://ai.umich.edu/learning-analytics-guiding-principles>
The Open University <https://help.open.ac.uk/documents/policies/ethical-use-of-student-data>

Porayska-Pomsta (2022). Critical data studies perspectives are also being brought to bear on education (Prinsloo, 2019; Selwyn, 2019; Williamson & Eynon, 2020).

As illustrative examples of the kinds of approaches being adopted, Holmes et al. (2021) present a qualitative analysis of how AIED researchers perceive the response of the AIED community to the challenge of FATE, from which they distil a set of thematic challenges. Going into greater technical detail, Kizilcec and Lee (2022) offer a helpful guide to different notions of “fairness” in educational algorithms, differentiating “measurement (data input), model learning (algorithm), and action (presentation or use of output)”, each of which can lead to different biases. Baker and Hawn (2021) develop a yet more detailed taxonomy of algorithmic bias specifically in machine learning, and propose ways in which this can be mitigated, offering “a framework for moving from unknown bias to known bias and from fairness to equity”. There remains, to our knowledge, no work in education on the use of “algorithmic reparation” (Davis et al., 2021), whereby the proactive commitments associated with reparative measures are implemented through algorithms seeking to “name, unmask, and undo allocative and representational harms”.

However, as mathematical abstractions or even running code, algorithms alone do not shape the world, but must be embedded materially in social contexts, where they are encountered by people using interactive software tools, or more pervasively, through the digital infrastructure that tracks and analyses activity and other sensor data. This brings us to the third important strand of work.

Human-centred Design (HCD)

Within the human–computer interaction (HCI) community, we find rich accounts of how diverse groups of people interact with AI. This informs design principles such as *human agency* (Shneiderman, 2020), the *accountability and explainability* of AI-output (Abdul et al., 2018), and efforts to empirically evaluate the slippery quality of *trustworthiness* (Vereschak et al., 2021). Researchers are beginning to demonstrate how human-centred design methods, which give a meaningful voice to non-technical stakeholders in the design process, can be adopted and adapted. For example, Buckingham Shum et al. (2019) make the case for *human-centred* Learning Analytics, developed through design processes informed by the use of HCI concepts and methods. As illustrative examples, Holstein et al. (2019b) describe a methodology to help classroom teachers co-design of a heads-up display showing them students’ progress in an intelligent tutoring system; Dollinger et al. (2019) describe the use of participatory design techniques to empower teachers as co-designers; Richards and Dignum (2019) review the empirical evidence and ethics concerning learners’ interaction with pedagogical chatbots; and Johanes and Thille (2019) present a rare account of the ways in which technical experts — often the most powerful stakeholders in software design — conceive and engage with ethical concerns.

What is within a University’s Sphere of Influence?

Reflecting on these three approaches, the problem with the FATE and HCD responses is that educational institutions are rarely able to exercise a high degree

of control over the products they purchase. The critical decisions shaping FATE and HCD remain in the hands of the developers. Institutions with capacity to develop in-house software retain far greater agency, while also having to negotiate the challenge of scaling this for organisational innovation (Buckingham Shum, 2023; Buckingham Shum & McKay, 2018). That being said, FATE and HCD values, principles and methods can be used by universities to bring stakeholder voices into framing “the problem” from the earliest phases, which will define what might count as a satisfactory “solution”, prior to initiating procurement processes. Institutional leaders and instructors play a key role in alleviating issues of bias in relation to university students’ educational data – especially subpopulation groups (Li et al., 2021). Moreover, the value of multi-stakeholder studies in higher education is that they surface diverse stakeholder tensions about technology awareness, understanding, access and usage (Sun et al., 2019).

Notably, the strand of work falling squarely within every institution’s sphere of influence is the set of principles guiding procurement and deployment, which requires new forms of ‘collective policymaking’ (Gulson et al., 2022). It is on the co-production of those principles that we now focus, specifically, addressing the fundamental question: “*Whose principles?*” We will argue that the answer should not merely be “*principles decreed by the leadership/expert panel*”, and ask: “*How can an institution engage in ethical co-production with its diverse community about their values, concerns and expectations regarding the use of AAI-EdTech?*”.

The Potential of Ethical Co-production

Approaches to participatory forms of AAI-EdTech ethical discourse are under-theorised and under-examined. We introduce co-production as a candidate, and while approaches vary, common features are the complex and dynamic relationships, processes, and values which emerge from producing knowledge in specific contexts with diverse stakeholders to address key issues (Filipe et al., 2017). From a science and technology studies perspective, the ‘idiom of co-production’ (Jasanoff, 2004) refers to the ways in which knowledge, culture, and power intersect, which offers “a way of interpreting and accounting for complex phenomena so as to avoid the strategic deletions and omissions of most other approaches in the social sciences” (p. 3). Co-production has gained prominence in attempting to address multifaceted challenges in fields such as healthcare and sustainability science (Filipe et al., 2017; Wyborn et al., 2019). It has been identified as a strategy for knowledge-policy interactions with different approaches and interpretations, that requires “definitional clarity of what form of co-production is being carried out” – especially in relation to the style of participation and processes involved (Bandola-Gill et al., 2022).

To explore the potential of co-production as the means to articulating AAI-EdTech ethical principles, and to expand awareness of the limits and possibilities of different approaches, we focus on a specific approach, Deliberative Democracy.

Deliberative Democracy

Deliberative Democracy (DD) emerged as a “deliberative turn” in democratic theory around 1990 (Dryzek, 2010), in response to the crisis in confidence in how poorly typical democratic systems engage citizens in decision making. DD works by creating a *Deliberative Mini-Public (DMP)*. DMPs can be convened at different scales (organisation; community; region; nation) and can take many forms, including Citizens’ Juries; Citizens’ Assemblies; Consensus Conferences; Planning Cells; Deliberative Polls (Elstub et al., 2016).

A DMP has three core features (Carson & Hartz-Karp, 2007):

- 1 *Influence*: The process should have the ability to influence policy and decision-making.
- 2 *Inclusion*: The process should be representative of the population and inclusive of diverse viewpoints and values, providing equal opportunity for all to participate.
- 3 *Deliberation*: The process should provide open dialogue, access to information, respect, space to understand and reframe issues, and movement toward consensus.

Moreover, “random selection, more than any other feature, is what delivers the ‘mini-public’ aspect of a DMP” (Carson & Hartz-Karp, 2007). This “produces a certain mindset in the room, which is very different to that resulting from a selection process governed by election, by the selection of interest group representatives or by merely allowing those most interested to turn up” (Farrell et al., 2019). Stratified sampling is often used to ensure representation of important sub-populations who might otherwise be missed in purely random sampling.

Other key features highlighted by Farrell et al. (2019):

- *Who sets the DMP’s purpose/agenda* may be contentious. If there are concerns that the government/management are unreasonably biasing the outcome by the very framing of the problem, then the DMP should be conducted at arm’s length.
- DMPs are *facilitated* by a neutral person or even better, a pair/team, who have no stake in the outcome.
- Participants commit to engaging in *deliberation* which requires more than the usual modes of discussion: the ‘rules of engagement’ typically include giving reasons for views, fairness, equality of voice, and openness to difference.
- ‘*Expert witnesses*’ contribute to making the (mini)-public hearing as informed as possible, but are not directly part of the DMP decision-making process. Experts should be balanced so as not to bias deliberation unreasonably. The DMP has the power to call their own experts.
- *The DMP’s recommendations/decisions* are decided deliberatively, providing reasons for recommendations.
- The DMP should be sanctioned by government/senior leadership, with a *commitment that the DMP’s recommendations/decisions matter*. Depending on the context, the DMP’s outputs may be one of many inputs to a policy consultation,

or the primary input. What is critical is that the DMP is not seen as a tokenistic exercise.

In addition to being an active field of political theory and practice (Elstub et al., 2016), DD is a professional practice, with many companies designing and facilitating structured consultations that recognise the above principles. Within our own Australian context, the New Democracy Foundation⁴ is a primary source of information, and in partnership with the case study university, ran DD courses which catalysed the conception and co-design of the present initiative.⁵ The use of DD to conduct an *organisational* consultation (as opposed to a citizen consultation) is novel, as is the focus on *AI ethics*. We now detail the institutional case.

Institutional Case

Like most universities, the University of Technology Sydney (UTS) captures an increasing quantity and quality of student and staff activity in the form of activity traces logged in online platforms. This raises important questions about how such ‘surveillance’ capability is as transparent and beneficial as possible for all stakeholders, in order to preserve the community’s trust in a fast-moving area. The challenge of doing this well is the focus of the academic fields such as *Learning Analytics*, *Educational Data Mining*, and *AI in Education*. While educational technology (EdTech) covers the entire spectrum of software tools used to assist teaching and learning, we will use the term *Analytics/AI-enabled EdTech (AAI-EdTech)* to refer to the broad range of interactive learning and teaching software using analytics or AI to make sense of student data.

UTS is active in deploying AAI-EdTech in its teaching, including active work on the ethics of such tools (Kitto & Knight, 2019) and how diverse stakeholders can be brought meaningfully into the design process (Buckingham Shum et al., 2019). UTS is scaling up the use of automated feedback to students using a range of platforms including provoking students to write in more academically rigorous ways (Knight et al., 2020; Shiban et al., 2020), providing instant feedback on learning dispositions (Barratt-See et al., 2017; Buckingham Shum & Deakin Crick, 2012), enabling academics to provide timely, personalised feedback at scale (Lim et al., 2021, 2022) and enabling students and educators to reflect on face-to-face teamwork (Fernandez-Nieto et al., 2021, 2022). The organisational strategy and practices required to invent and scale in-house AAI-EdTech have been the subject of considerable reflection (Buckingham Shum, 2023).

While UTS had the typical Privacy and Confidentiality policies in place, and all research is subject to the Human Research Ethics Committee (HREC), with the scaling up of AAI-EdTech, UTS sought to institute:

⁴ The New Democracy Foundation <https://www.newdemocracy.com.au>

⁵ UTS Course: *Leading Deliberative Democracy* <https://open.uts.edu.au/uts-open/study-area/business-and-transformation/leadership--management/leading-deliberative-democracy>

UTS Course: *Doing Deliberative Democracy* <https://open.uts.edu.au/uts-open/study-area/public-policy--governance/doing-deliberative-democracy>

- Principles and policies to address the **particular** ethical issues that can arise with AAI-EdTech;
- A **consultation process** to engage the diverse community of students, tutors and academics in informed deliberation about their expectations and values with regard to AAI-EdTech.

With the growing use of AAI-EdTech, the Vice-President (Education & Students) requested that the research and innovation team in the division, the Connected Intelligence Centre (CIC), coordinate the development of a set of ethical principles to inform policy around their usage. CIC identified the potential of DD processes as a participatory methodology to enable stakeholders to co-develop such a set of principles, as detailed next.

A critical element in a DD consultation is to design a brief: what is the DMP being asked to do? This must be suitably open-ended, but scoped to be tractable within the given constraints, making clear the deliverable. The planning team set this as the brief: “*What principles should govern UTS use of analytics and artificial intelligence to improve teaching and learning for all, while minimising the possibility of harmful outcomes?*”.

Project Design and Process

Central to the DD model is the sanctioning of the consultation by the institution’s senior leadership. Prospective participants were assured that the significant time and effort they would invest would not be wasted. The DD project was therefore promoted and designed to make very clear the university’s interest in the outcomes. Since nobody is compelled to participate in such processes, there is an inherent response bias towards people who are willing to engage, but as noted above, random or stratified sampling is used to mitigate the biases of participants being dominated by particular sub-populations. The deliberative mini-public (DMP) was selected through stratified sampling from UTS sub-populations, as far as possible intersecting a range of demographic attributes considered to be important, as summarised in Table 1: *Gender, Indigenous (self-identified), English as a Second Language (ESL), Undergraduate (Years 1–4), Postgraduate, Staff (Academics and Tutors), and Faculty*. A DMP of 20 was recruited from 131 applicants, which was considered to be a practical number for the intensive workshops, and all were asked to commit to the 5-part workshop series. This recruitment process mitigates but clearly does not eliminate sampling bias, and the results should be interpreted within those limitations.

A DD facilitator external to UTS was appointed to design and run the process in close consultation with the UTS lead (author 2) and a core planning team. The DD process ran across five sessions over seven weeks. Due to the pandemic, the entire process was conducted under lockdown conditions, using Zoom for workshops, Google Docs for collaborative editing and Basecamp for messaging and document repository.

Over this period, the group practised skills in critical thinking, learnt AAI-EdTech terminology, how to write principles, and ways to test their level of consensus. The

Table 1 The UTS Deliberative Mini-Public recruited through stratified sampling

Gender	Indigenous	ESL	UG Yr	Staff	Faculty
Academics					
M	No	No		Lecturer	Business
F	No	No		Lecturer	Design
F	No	No		Senior lecturer	Science
F	No	No		Senior lecturer	Transdiscip
F	No	Yes		Senior lecturer	Transdiscip
Casual tutors					
M	No	Yes			Art/SocSci
M	No	No			Eng/IT
F	No	Yes			Health
F	No	No			Law
Undergrads					
F	No	No	1st year		Eng/IT
F	No	No	3rd/4th year		Transdiscip
M	No	Yes	3rd/4th year		Science
M	Yes	No	1st year		Communication
M	No	Yes	2nd year		Eng/IT
F	No	Yes	2nd year		Design
Postgrads					
F	Yes	No			Health
M	No	Yes			Art/SocSci
F	Yes	No			Transdiscip
M	No	No			Transdiscip
F	No	Yes			Business

group was also introduced to UTS systems, imaginary use cases/vignettes, as well as potential future scenarios. Guest experts based at the institution were invited who could describe how AAI-EdTech was already, or could in the future be, used at UTS, and the ethical aspects associated with each approach. Two external experts were selected to bring complementary expertise in technology ethics (author 1 and author 3). These combined activities and expertise strongly informed the co-production of principles which were iteratively developed and refined over the course of the project with the whole DMP; alongside a core group of DMP members who volunteered to help shape the wording of the principles between the workshop sessions. The output of Session 3 produced a work-in-progress draft of the principles, which went to a group of stakeholders whose teams would be expected to apply the principles, plus experts on learning analytics, ethics and social justice. In the closing workshop, four members of the DMP were elected by their peers to introduce the revised, extended principles to three senior leaders in UTS with different briefs for the responsible implementation of data, analytics, and AI. Table 2 provides additional details and the project website shows the participant recruitment details. The

recording of the DMP's introduction to the final session conveys the passion and commitment that they invested in the whole process and the outcome. This video can be viewed, together with the final set of principles on the consultation website.⁶

Positionality Statement

No educational research is value-free, far less one in which the researchers are participants. We recognise our bias towards hoping to see this pilot succeed, since we bring an academic interest in participatory approaches to educational stakeholder engagement in AI ethics. However, we consider that we have maintained appropriate critical distance in the design and execution of this research, while making no claims to complete objectivity or control. Our positionality corresponds with the idea that the implementation of DD in higher education offers a ‘critical space’ for dialogue and inquiry which both exceeds traditional and neoliberal norms and invites new modes of scholarly participation (Mourad, 2022). Firstly, this was the first test of DD in an authentic, novel context, requiring wholly online interactions under demanding pandemic lockdown conditions. There was no guarantee that this experiment would work since much depended on whether the DMP members came to trust the process, each other, the facilitator and the experts. As we discuss later, creating the conditions for genuine co-production is not a formulaic process which can be engineered.

Specifically, the following details are relevant. *Author 1* is affiliated with a different university to UTS, and was recruited to the project to observe the planning process and workshop sessions, and to co-design and conduct the interviews. She also served as an external expert in Session 2, facilitating discussion of the ethics of AI exam proctoring (Table 2). *Author 2* has institutional ‘insider’ status, employed at UTS, and served as the project lead, co-designing the sessions, but not facilitating them (a professional DD facilitator was recruited for this). His position at UTS enabled him to recruit senior colleagues to provide feedback on the DMP’s draft and revised principles, but this also biased their selection. He did not conduct any interviews, in order to maintain appropriate distance from the interviewees. *Author 3* is affiliated with the same university as Author 1, and was recruited to the project to observe the planning process and workshop sessions. He jointly presented with Author 1 in Session 2, and co-designed but did not conduct the interviews.

Evaluation Method

Participants and Recruitment

The DMP participants were informed about the option to be interviewed at the final DD workshop, and via Basecamp (an online communication platform utilised throughout the consultation). Author 1 (from outside UTS) conducted

⁶ UTS EdTech Ethics Deliberative Democracy website: <https://cic.uts.edu.au/projects/edtech-ethics>

Table 2 Deliberative democracy consultation design

Participants	Zoom Session 1 (21 Oct 2021, 1–4 pm)		Zoom Session 2 (4 Nov, 1–4 pm)		Zoom Session 3 (18 Nov, 1–4 pm)		Zoom Session 4 (25 Nov, 12–4 pm)		Zoom Session 5 (2 Dec, 2.30–4 pm)	
	DMP	DD planning team	DMP	3 guest experts DD planning team	DMP	3 guest experts DD planning team	DMP	8 experts on learning analytics, ethics and social justice	DMP	3 senior leaders UTS DD planning team
Skill-building and collective learning activities	Information pack Terminology Critical thinking video Intro to UTS systems 'Cool or Creepy' vignettes	Under the hood/bonnet session (ethics of AI exam proctoring) Presentation of future scenarios (2025 visions of learning analytics)	Process of writing principles UTS examples (tensions/ethics, OnTask)	Hear and discuss feedback on draft principles from leaders Consensus testing	Process of writing principles UTS examples (tensions/ethics, OnTask)	Prioritising principles that they want guiding the university (principle, rationale, example)	Refining principles (write, give feedback, improve, repeat until agreement reached)	Finalised principles to be integrated into project report (and inform future work)	Finalised principles to be integrated into project report (and inform future work)	Closing presentation of revised principles to UTS leaders and discussion
Co-production of principles	Response to activities/questions (thinking about how AI and analytics can improve teaching and learning): What are you comfortable with? What are you concerned about? What are you curious about?	Survey results Range of ethical considerations and principles identified to guide university's decisions for different scenarios	3 DMP volunteers (with DD facilitators) select 9 groupings of principles	UTS expert: implementation of online exam proctoring during lockdown	Ethics and UTS designed tools in practice	Feedback on v1 from UTS leaders	Work-in-progress draft principles produced for institutional feedback			
Institutional engagement	Opening welcome from the DVC Education & Students									

semi-structured online interviews, recorded via Zoom, between December 2021 and February 2022. This study was approved by the UTS Human Research Ethics Committee (ETH21-6615). Thirteen stakeholders across the three key groups were recruited for this study: students ($n=5$), educators ($n=4$), plus institutional leaders ($n=4$) who were not part of the DMP but who had engaged at least once with the process to provide insight or feedback. This range of higher education stakeholders spanned the following groups: undergraduate and postgraduate students, tutors and lecturers, plus middle and senior management.

Interview Protocol and Analysis

The design of the interview questions (Tables 3 and 4) was structured according to an analytic frame which spanned four key categories: *perceptions of deliberative democracy, experiences of the deliberative democracy process, views and visions of the principles, plus recommendations for future deliberative democracy experiments and shaping the future of AAI-EdTech*. The subsequent analysis of interviews spanned multiple phases. First, preliminary insights from the interviews were summarised according to the analytic frame's four categories. Once the individual interviews were transcribed, the next phase involved highlighting key transcript excerpts to align with the analytic framing categories, which were then clustered according to each stakeholder group (students, educators, and professionals). This iterative process resulted in an overview of multiple stakeholder group perspectives, from which key three themes emerged, detailed next. Author 1 led the interview analyses, a synthesis of which was then refined in discussion with Authors 2 and 3.

Findings

Key findings and stakeholder insights are synthesised according to the following three key themes (Table 5):

1. A uniquely structured and supportive process involving a range of higher education stakeholders;
2. Integration of situated knowledge and expertise to shape ethical principles about AAI-EdTech; and
3. An innovative approach to prioritise broader expertise and research about AAI-EdTech (both in-house and third-party designed) within a university.

The organisation and permutation of these themes — across *perspectives, contexts, and possibilities* — reflects the range of conceptual, embodied, and anticipatory aspects associated with these collective insights identified from students, educators, and leaders.

Table 3 Interview questions for students, tutors, academics

Focus	Questions
Role, motivation, prior knowledge/ experience of DD	<p>1. Could you please tell us a bit about your area of study/work at UTS?</p> <p>2. How did you hear about the opportunity—and what motivated you to take part?</p> <p>3. The advertisement linked to a webpage with further details about the DD process. Had you heard of, or participated in, a deliberative democracy approach prior to this UTS event?</p> <ul style="list-style-type: none"> • YES: what particular aspect of DD do you think makes it different from other forms of collaboration; how was the experience similar/different? • NO: what particular aspect of DD do you think makes it different from other forms of consultation?
Perspectives of process, communication/ information, learnings	<p>4. How comfortable/engaged did you feel sharing your thoughts and ideas over the course of 5 sessions? Did it change, or differ, over particular sessions/activities? E.g. during break-out rooms, whole-group discussions, in-between sessions, study/work/home commitments</p> <p>5. A range of experts provided different perspectives about AAI-EdTech. Could you share a few things you learned from the experts?</p> <p>6. Prior to this event, were you aware that UTS had designed some of the AAI-EdTech systems discussed? What do you think these in-house-designed tools offer which proprietary/for profit tools don't?</p> <p>7. What advice would you provide to improve future DD events at UTS? E.g. workshop activities, communications, information packs</p> <p>8. How do you feel about being part of the student-staff team which has co-designed this set of principles for UTS? What particular aspect/contribution makes you feel that way?</p> <p>9. How would you recommend we communicate these principles to UTS staff and students (both new and the broader population)?</p> <p>10. Who would you approach at the university if you had concerns about a member of the UTS community, or a newly introduced AAI-EdTech platform, not abiding by this policy?</p> <p>11. Would you like to be part of future DD events at UTS? Why? Why not?</p> <p>12. What is the one thing which you think is most important for shaping the future of ethical AAI-EdTech at UTS?</p> <p>13. In closing, is there anything else you would like to share about your participation/experience?</p>
Perceptions of principles and future practices	

Table 4 Interview questions for guest experts/leaders

Focus	Questions
Role, prior knowledge/experience of DD	<p>1. Could you please tell us a bit about your</p> <ul style="list-style-type: none"> • Expertise • Management/leadership role at UTS and departmental oversight <p>2. Had you heard of, or participated in, a deliberative democracy approach prior to this UTS event?</p> <ul style="list-style-type: none"> • YES: what particular aspect of DD do you think makes it different from other forms of collaboration; how was the experience similar/different? • NO: from your understanding, are there aspects of DD that you think makes it different from other forms of consultation?
Perceptions of principles and future practices	<p>3. Based on your management role/area of expertise, what do you think is most interesting about this set of proposed principles?</p> <p>4. Selection of LA/AIEd tools in higher education</p> <ul style="list-style-type: none"> • UTS management: Do the proposed principles raise any implications for EdTech procurement and its introduction to teaching at UTS? - Expert: What do you think UTS in-house-designed tools offer which external/proprietary tools don't? - What role does the procurement process play in the selection and development of AAI-EdTech tools? <p>5. What do you think are the benefits of the DD process which generated the principles?</p> <p>6. Compared to other university principles, do you think there is anything that makes the UTS principles distinctive?</p> <ul style="list-style-type: none"> • Are you aware of other universities' ethical principles around AAI-EdTech, or data/analytics/AI in general? If so, does this set of principles seem distinctive in any way? • Do you think this set of principles adds anything distinctive to existing UTS principles/policies? <p>7. We're also interested in understanding how ethical concerns related to AAI-EdTech can best be responded to within HE organisations:</p> <ul style="list-style-type: none"> • Management and experts: Principles such as these are no use if there is no accountability or process in UTS to handle concerns that they're not being followed. Can you foresee any difficulties in implementing any of these? • Experts: What do you think is best practice and/or emerging research in terms of ways organisations should be accountable <p>8. In closing, what is the one thing which you think is most important for shaping the future of ethical AAI-EdTech at UTS/in higher education?</p>

Theme 1: a Unique and Structured Process Involving a Range of Higher Education Stakeholders

Theme 1 Perspectives

Students, educators, and leaders perceived this DD project as a unique opportunity to participate in *a curiosity-driven, applied, and dynamic process*.

None of the **students** had prior experience or knowledge of DD. What sparked their interest to participate in this project ranged from: previous student participation experiences; curiosity about DD and how technology is changing; finding

Table 5 Overview of themes emerging from the interviews

Themes	Perspectives	Contexts	Possibilities
Theme 1: A unique and structured process involving a range of higher education stakeholders	A curiosity-driven, applied, and dynamic process	A culture of trust that scaffolded conscious elicitation of different stakeholder opinions	More time and space to prepare and enhance contributions within the group, alongside ways to surface broader project linkages and test the final output
Theme 2: Integration of situated knowledge and dialogue to shape ethical principles about AAI-EdTech	A sense of pride and achievement which involved unsettling views about roles, belief systems, and mandates with a set of principles distinct across organisational, sectoral, and policy levels	An interactive information exchange which relies upon testing what is reasonable by empowering participants to transgress preconceived notions and generate new knowledge	Online and in-person media, events, training/learning modules, and iterative research to build a strong, education-focused evidence base
Theme 3: An innovative approach to support ongoing research and broader expertise about AAI-EdTech (both in-house and third-party designed) within a university	Translating and adapting the principles into action with inter/transdisciplinary research that re-imagines expertise, procurement and student-staff support in light of social justice, wellbeing, and democracy	In-house, locally-designed tools that often tailored, trusted systems which prioritise teaching and learning, plus can be tuned and controlled more than third-party tools	An ongoing conversation that respects openness, diversity, human decision-making, and the stewardship of learning tools, and evaluations aligned with public trust and benefit

out more about AI and how to make it better; prior interest in democracy, development, and decision-making; plus, how ethics and AI intersect in relation to both personal and professional contexts. For instance, one student's interest was piqued to learn more about DD:

"I was like, how does that work? What's that about? Let's find out what that's like!" (S1)

In reflecting upon their experiences of the DD process, student participants largely felt comfortable engaging in the process. Most students felt very comfortable in the breakout rooms. However, one student felt a sense of discomfort discussing an unfamiliar topic in a small group, while another was acutely aware of the uneven power dynamics between students and authority figures.

Educators were motivated to take part due to a range of reasons: knowledge of researchers associated with the project, interest in AI and how it could be applied to their work and research interests, as well as particular passions for the topics of inclusion and democracy. One educator communicated their strong interest in democracy and consensus:

"I have a fascination with, and a commitment to, and a passion for, consensus, when it's understood as being also the freedom to disagree. Not simply majority rule, as most people sadly think of it as." (E3)

In reflecting upon their experiences of the DD process, educators communicated varying levels of comfort. These levels changed depending on the activity; were at times constrained by the facilitation style and strict timekeeping; feeling that confidence was gained over time; and, that issues were more comfortably raised in the breakout rooms.

Leaders had awareness of DD, but limited experiences. Only one leader had prior experience of DD; with the others communicating awareness of DD, but no associated experience. One leader pointed out that they had never considered that DD could be applied in higher education:

"so I'd heard of this term in the past, but it was more kind of a school of thought in political theory, so I never thought that it can be applicable in an educational institution, you know, for such activities as well." (L3)

All leaders highlighted the uniqueness of bringing together diverse perspectives over the course of the process.

Theme 1 Contexts

This project's particular application of DD offered students, educators, and leaders a structured and contextual process built upon a culture of trust that scaffolded the conscious/careful elicitation of different stakeholder opinions and experiences.

All of the **students** communicated that there was something distinct about DD in practice, which differed from other forms of collaboration, such as: the diversity of participants and perspectives; a lack of hierarchy; a civilised process based on shared

responsibility; in-depth coordination and involvement of participations in shaping the outcome, and; the structure, deliberation process, and quick pace. Compared to other collaborations, one student noted, where:

“it’s normally around people who know what it is, and maybe that’s the field they work in, or study... [this DD process was unique] ...you could get perspectives that would be missed out in those other situations.” (S3)

All of the *educators* communicated that there was something distinct about DD in practice, such as: being a supportive process for participants with various abilities to communicate knowledge; a productive system for sharing agreement and disagreement; a refreshing process because decisions were not made beforehand, alongside the involvement of both students and staff beyond one session; the particular importance of building trust and keeping end-view in mind; and, offering a different process and mindset to traditional product development. One educator strongly felt:

“that it was one of the best experiences I’ve had in a long time, in trying to get the committee to do work, or a group of people to do work who come from varying places with various knowledge sets, and various abilities, which is very important.” (E1)

For *leaders*, DD in practice enabled new understandings, such as: interconnecting normally segmented opinions; an educative process informed by people’s lived experiences; a process of information exchange and shared decision-making powers; and, building a culture of trust, honesty and transparency. For one leader, the DD process was supported by a:

“culture of honesty and transparency... there was clearly a strong sense of already established trust within those working groups, because people were very honest about their experience. Like, there was no holds barred. People would say, “Oh, this just doesn’t work!” and “I hated this!” and “This has completely ruined my working life!” and stuff like that” (L4).

Theme 1 Possibilities

Students and educators identified a range of possibilities to improve the process for future DD experiments, such as more time and space to prepare and enhance contributions within the group, alongside ways to surface broader project linkages and test the final output.

Recommendations from *students* included: more homework between sessions, more varied activities and opportunities to collaborate with peers and experts (such as extra use of online platforms for questions/conversation/dialogue, as well as generally more time and the value of emphasising diversity. More advance preparation and overview of the process was proposed, especially that:

“given the time frame, that was a huge challenge to being able to provide an opinion, provide your voice, and actually kind of hash out a lot of the details.” (S5)

Recommendations from *educators* included: recognising that communication platforms are necessary, but that not everyone likes using them; making clearer that the limited timeframe and process means that the final output will not be perfect; value of continually updated information pack; opportunity for more diverse facilitation; testimonials about people's experiences; plus involving experts in earlier workshops. An educator with a product development background also highlighted the importance of a broader perspective of the project, as a way of testing the principles:

"I need to get a very broad view of everything before I feel comfortable, because without that, I can't see the linkages." (E4)

Theme 2: Integration of Diverse Knowledge and Dialogue to Deliberate Ethical Principles about AAI-EdTech

Theme 2 Perspectives

Students, educators, and leaders viewed this DD project with a sense of pride and achievement which involved unsettling views about roles, belief systems, and mandates with a set of principles distinct across organisational, sectoral, and policy levels.

Students communicated a range of positive emotions about their involvement and contribution to the co-produced principles: feeling invigorated and excited, pride and enjoyment, as well as a sense of reward and empowerment. One student communicated a strong sense of responsibility and hope that the principles could inform continued conversations within the university:

"I did not have any experience with being tasked with such a big responsibility to come up with principles that would affect everyone at the University. All the stakeholders. So, it was a genuinely proud moment when we finished, but I'm just interested in how this conversation goes on, moving forward, and as we discussed in the final meeting, we would really like it not to be a full stop; rather, an ongoing conversation." (S2)

Educators also communicated a range of positive emotions and powerful feelings about the co-produced principles: especially a strong sense of satisfaction, pride and positivity in the end achievement and result, plus that such a process was long overdue. One educator felt that the DD process was a refreshing and respectful space for authentic dialogue which in their view is increasingly uncommon at universities whose cultures can inhibit open discussion:

"I think respect is everything, and I think there was something respectful about being able to participate in something and have a voice, and I also think that there was something very respectful in having students and staff at the same level." (E2)

The **leaders** viewed the principles as distinct across organisational, sectoral, and policy levels. The principles were seen as being a leading-edge, practice-based approach unique to the university. This was because:

“it was the University looking at its own practice. So, rather than just going out and telling everyone else what they should do, and teaching a whole lot of other people what they should do, it was actually about the University seeking to be best practice in this space as well, in relation to the enormous amounts of data and information and use of new technology that universities undertake.”
(L4)

Theme 2 Contexts

This DD project involved students, educators, and leaders in an interactive information exchange which relied upon contextually testing what is reasonable by empowering participants to transgress preconceived notions and generate new knowledge.

Students communicated the intense value of collaborating with diverse stakeholders, such as: hearing in-depth views from experts and others; a sense of enjoyment from hearing different views and opinions to help shape the principles; and the value of interacting with people beyond their regular study or social circle. One student highlighted the sense of enjoyment they felt about collaborating with the different educators and students in the group and that the experts helped them to understand different views to help shape the principles. In particular, they raised the example of being surprised to hear that nursing students were wearing trackers (on simulation wards), which they viewed as a good way to encourage everyone to think about other contexts, uses, and perspectives of technology to inform the principles:

“if we didn’t interact, then we could be coming to different conclusions about what we think is important. So, just being able to chat with other people and learn that kind of stuff, I think that was really useful for coming up with the principles that would benefit everyone in the University.” (S3)

When asked to specifically recall what they learnt from the process, students highlighted a range of topics, such as: hearing from the university’s examinations coordinator how the ProctorU automated exam invigilation platform was managed; how EdTech operates within the organisation; how to efficiently run meetings; plus learning about the value of ethics and multiple perspectives for data science.

Key learnings identified by **educators** about DD in practice spanned the importance of language and testing, seeing how technologies were used in practice, and that knowledge levels rise over time. For example, knowledge integration involved being introduced to new language and key concepts in discussion with experts (such as learning about the distinction between ‘access’ and ‘equity’, and the value of ‘testing’). Over time, this informed dialogue and the selection of key words for the co-produced principles, for instance:

“So, dignity was one that got captured in the document. So, these are ones that you don’t typically see in governing documents. So, I remember all of the words. And you had normal, day-to-day words that are always used in these things. Harmful, hurt, pain, suffering. You hear those. But there were particular words that are not typically been in governing documents and behaviours attached to those that we used that kind of stood out with me” (E1).

Characteristics of DD in practice observed as distinct by *leaders* involved actively challenging preconceived notions with an information exchange that empowered participants to generate knowledge and policy. In contrast to reactive ways of extracting knowledge from users, DD offered an inclusive knowledge integration approach:

“I guess the unique thing about deliberative democracy is that it is a form of knowledge generation that emerges from the subjects [...] this is placing an emphasis on the user as a generator of knowledge, as an originator of knowledge and policy.” (L2)

Theme 2 Possibilities

Students, educators, and leaders identified a range of possibilities for communicating and implementing the principles, including: online and in-person media, events, training/learning modules, and ongoing research to build a strong, education-focused evidence base.

Students shared a range of ideas for communicating and implementing the principles, such as: emails and social media; events and discussion forums to get feedback; plus, ongoing review with key decisions-makers. For example, rather than being ‘hidden’ in policy documents where no-one can find them, a student proposed the idea of integrating the principles into a module:

“I feel like in terms of using the principles, making sure they’re used when the actual technology’s implemented, I think having some sort of module in the orientation for students, and I guess for staff as well, to kind of introduce them to the ideas might be a helpful way to kind of prepare them and make sure you’re getting informed consent about the different technology when they start studying at UTS.” (S1)

Varying ideas from *educators* for communicating and implementing the principles spanned participation information, multimodal website assets, and translating the principles into action. One educator stated the need to ensure that the principles are framed as action-guiding to support decision-making and judgments driven by human learning and educative agendas:

“Because if you just leave principles as they are, there’s a terrible tendency for them to be seen as laws. There’s some kind of universal law, some kind of universal principle. No, no, no. ... Action-guiding. Circumstances are going to change. We need to see what the principles are that are going to

help us make decisions, because ultimately – and here is where I will finish; not now, but later on – ultimately, we're educating people for judgement.” (E3)

Ideas from **leaders** about communicating and implementing principles focused on addressing tensions between principles and practice. The principles were strongly recognised as aspirational in light of multiple challenges of operationalisation, including issues of digital divide and literacy issues. One leader focused on identifying the tensions associated with accountability, especially in relation to third-party software:

“the assumption there is that the University actually has the power to define the policies of educational technology when, actually, the way that it usually works is that we have vendors who are making changes all the time to the rules that govern the software that they’re deploying.” (L2)

Theme 3: an Innovative Approach to Support Ongoing Research and Broader Expertise about AAI-EdTech (Both In-house and Third-party Designed) within a University Community

Theme 3 perspectives

Students, educators, and leaders, with diverse disciplinary backgrounds and expertise, viewed next steps for this DD initiative as *translating and adapting the principles into action with continued inter/transdisciplinary research that re-imagines expertise, procurement and student-staff support in light of social justice, wellbeing, and democracy.*

All **students** expressed interest in future involvement, especially if it is a topic they are interested in, alongside broader interests to continue shaping the future of ethical tech with the broader university community. A student expressed the importance of continual adaptation and improvement in relation to students and broader social and technological change:

“I think it’s really important that we kind of have this sort of spirit of continuously improving and adapting to our students and the world around us and how...continually look at how the EdTech actually impacts stuff, right, because the last thing we should do is say, “Hey, we said these, and now these are set in stone; they’re forever.” We don’t know, you know, what changes the world, how it might impact the world in the future, so I think we should continually reassess and improve, and as new technology comes along, also reassess and improve.” (S5)

When asked who they would approach if they wanted to raise issues or concerns about an existing or new technology, students were largely not sure, but identified possibilities such as: the DD project lead; a lecturer, tutor, course director, or student

centre representative (though they might not be able to answer query); or, a course coordinator, or dean of the college/faculty.

All *educators* would participate again, especially if it was something that sparked their interest of passion; because of the well-structured process; being a ‘game-changer’ process; plus, the opportunity to contribute to change and develop not just discipline-based graduate attributes, but also inter/transdisciplinary values – such as social justice. One educator shared their hope that the University would continue to lead in this space so as to demonstrate its institutional values:

“if we are about social justice, this is a great social justice initiative as much as anything else. If we’re about social impact, it’s about social impact. You know? And we need to put our practices where are marketing documents are. We all know the words for it. We all read the words for it. But we need – things like this are the kind of thing we should be doing.” (E2)

Leaders identified the complexity of dealing with technology-related concerns and issues, such as: privacy and consent being tricky issues for people to navigate; and that the principles offer a way of supporting dialogue with students. Embedding ongoing research about the university community issues and concerns would be a valuable way to respond to people’s experiences and ideas for change, which cannot be addressed one discipline, or sector. One leader communicated the need for being more thoughtful and mindful about purchases, and the interconnections between existing and new technologies which raise ethical issues beyond privacy and bias, such as:

“the inequalities that you could be locking into the process by using EdTech in ways that weren’t thoughtful and well-understood, let alone actual access to technology, so the more you move in the EdTech direction, including tools that are AI tools, or immediate feedback tools or whatever it is you’re doing, you’re potentially excluding a vast majority of people from those processes.” (E4)

Theme 3 contexts

This DD project enabled students, educators, and leaders to learn about in-house, locally-designed tools that offer contextualised, trusted systems which prioritise teaching and learning, plus can be tweaked and controlled more than third-party tools.

Students identified a number of advantages of university designed tools/system, such as: they can be more tailored and effective, give university more control over data and transparency; prioritise support for students and teachers; offer a clear history of a tool’s development and use and how it corresponds with the dynamics and diversity of teaching and learning; plus, potentially offer more privacy, flexibility and opportunity to customise. In-house designed tools were viewed as more effective because they are tailored to specific contexts and tasks. One student identified a further value of in-house designed systems was that the history, or provenance, of a tool could be communicated:

“I do believe that tools that are designed by the University, they have a clear history. They have a clear history because the University would have identified a

problem, and out of that problem they designed that particular software. Unlike for-profit tools that people develop, even though they are developed to solve a problem, but that problem is more business-driven, and business solution, than UTS would have done their own.” (S4)

Advantages of in-house tools identified by *educators* included the advantage of designing and researching tools to specifically serve the university community spanning academics, students, and support staff. One educator also highlighted that in-house designed tools offer more adaptability and opportunities for sector-leading research:

“It also allows us an amazing thing in the research space, because we should lead in this space, and for us to design in-house, it’s not just about us being able to deploy things well and design things well and to be able to tweak things well; we need to be leaders in this space, and we need to be thought leaders in this space, so to design things in-house and to have processes like this deliberative democracy thing, and to have it ongoing, and to have rigorous discussion and a safe space for that rigorous – and respectful – space.” (E2)

Advantages of in-house designed tools recognised by *leaders* included more control over systems and rules, plus being specifically designed. For example, advantages include:

“the control that you can have over the system and the rules that are embedded within those systems. So yeah, the more locally that software is designed for, is better.” (L2)

In-house designed tools were also seen as offering a more personalised approach and feedback functionalities.

Theme 3 Possibilities

Students, educators, and leaders identified the potential of this initiative as an ongoing conversation that respects openness, diversity, human decision-making, and the stewardship of learning, tools, and evaluations aligned with public trust and benefit.

Key areas identified by *students* as central to shape the future of ethical tech at the university included an ongoing conversation, and informed consent. An important thing identified to shape the future the future of ethical AAI-EdTech was diversity, which:

“should be harnessed in the process of going into the future, where people feel more comfortable and empowered for who they are, and the community they belong to, which is the UTS community. If their diversity is respected, and if they are feeling so empowered for who they are, then the future of technology can really be great for UTS, because people feel more comfortable with the usage of any technology.” (S4)

Educators stressed the value of human decision-making, control and communication, interdisciplinary learning and public trust for shaping the future

of EdTech. One educator highlighted the importance of human agency and control, in particular, that decision-making powers are a dynamic and shared responsibility:

“That humans are decision-makers, and they need to control it. They have – they should have the final control. Right? I think that’s the one thing...to position that perspective as humans as decision-makers, it’s hard – it’s going to be ignored that decision-making means that they have the power, but not all the time. The power is a shared power, depending on who the decision-maker is.” (E1)

Furthermore, another educator noted that technologies should not be:

“at the cost of human-to-human relations. Not at the cost of human-to-human relations and learning. Not at the cost of public trust.” (E3)

Key areas identified by *leaders* as central to shaping the future of ethical EdTech at UTS including openness, changing mindsets with training, ongoing evaluation of tools, balancing between new tools and ethical considerations. One leader identified staying at the cutting-edge of being a public purpose university of technology and keeping students at the heart of that purpose as essential:

“It is actually putting the student at the heart of why EdTech exists, and that has to have an equity lens [...] So, where is there EdTech that actually can increase learning outcomes? You know. So, not being anti-technology, but being really explicit about the outcomes you want to see from that technology, and anything purchased by any university should have public benefit, and student learning at the heart of its purpose.” (L4)

Discussion

The real world impetus for this case study and co-production process was that a university needed answers to the new challenges raised by AAI-EdTech, namely: “*What principles should govern UTS use of analytics and artificial intelligence to improve teaching and learning for all, while minimising the possibility of harmful outcomes?*” While the university could have simply announced a set of principles, it instead sought to explore this challenge via ethical co-production within its diverse community. The first contribution this paper seeks to make has been to evaluate the resulting DD initiative, from the viewpoint of participants, as we have just detailed.

The second contribution we seek to make is to examine DD as a particular mode of co-production that can respond to the complex challenges and ethical issues associated with AAI-EdTech. In this concluding discussion, we aim to identify the possibilities and limits of DD as a form of co-production. This discussion is therefore framed by the following aspects: *situating co-production; outcomes of co-production; recommendations for designing co-production; and, navigating constructive interactions with society and culture* (Wyborn et al., 2019).

Situating Co-production

Co-production is not a formula which can be imposed or implemented. Co-production is always situated within particular contexts involving diverse actors in a ‘joint effort’ which “results in some product, service, or body of knowledge that contributes to addressing an issue of shared concern” (Wyborn et al., 2019, p. 331). In this study’s context, DD offered a practical process toolkit for co-producing AAI-EdTech ethics principles. It enabled a university to recruit and facilitate a representative “mini-public”, building their knowledge, skills and dispositions to reach consensual, workable agreements, avoiding the polarisation that comes when unfacilitated (and often uninformed, poor quality) argumentation is dominated by opposing activists. Participants collectively viewed their involvement in the process as dependent upon *creating a culture of trust that scaffolded conscious elicitation of different stakeholder opinions (Theme 1 contexts)*. The high quality of deliberation that is accomplished when the process works well cultivates new ideas, and a strong sense of ownership to see the insights and their rationale understood and applied. Importantly, the mini-public should be sanctioned by leadership, with a commitment that the decisions matter; depending on the context, the outputs may be one of many inputs to a policy consultation, or the primary input – what is critical is that the mini-public is not seen as a tokenistic exercise (Farrell et al., 2019). In this project we saw passionate stakeholders who grew to trust the process, were proud of their accomplishment, and wanted to stay engaged. In addition, on the evidence of this experience, it seems fair to judge the current draft principles as ‘good enough’, because the process had the necessary integrity for participants to *trust* it, and they converged on a set of principles that they felt *represented fairly the diversity of perspectives*. This served to move the conversation forward with the university’s leadership, providing a good foundation for the implementation stage, and ongoing dialogue. In doing so, this accords with a new generation of deliberative democracy research spanning a “deliberative system” (Elstub et al., 2016) which situates deliberation as a communicative activity, occurring across multiple sites, and the vital need to address interconnections between different stakeholders and spaces.

The institutional case presented in this paper offers an exemplar for other institutions to trial and test this novel methodology in their own situations, or contexts. The pandemic lockdown necessitated that the mini-public be conducted wholly online, and, recognising the stress and fatigue that the pandemic placed upon the general population meant that preparation time expectations were kept to a minimum. Importantly the scope of project-specific expectations and goals must be clearly communicated, as timeframes for mini-publics differ largely depending on resources available.

Given more time, there could doubtless have been additional expert input, deeper learning, more stakeholders, and extended deliberation, all possibly leading to a differently articulated set of principles. One could have further democratised the process by making the definition of the brief consultative, which might have questioned its assumptions. However, given the constraints which will always govern consultations, the DD process can be judged to have been a success relative to the context of

co-production: the university now had a plausibly representative expression of the community's values, interests and concerns in response to the challenge.

Outcomes of Co-production

As a multi-level phenomenon, the outcomes of co-production seek to "impact the individual, community, and even knowledge systems scale" (Wyborn et al., 2019, p. 332). AI and technology policy developments signal the need for better understanding of the contextual, on-the-ground implications of emerging technologies. No matter how successful such a co-production process is, if it leads to no change, it has failed. Participants in the DD process viewed the co-produced principles as not only unique to the organisation, but also as leading university sector responsiveness to AAI-EdTech ethics. It is significant to note that following this DD process, its findings were very positively received by the university's Data Governance Board, who further recognised the need for a wider policy articulating ethical principles to govern all uses of AI. The second author was deeply involved in the cross-university group drafting this, consulting with three members of the DMP to ensure that it aligned with the EdTech Ethics principles. In Sept. 2022, UTS formally approved its AI Operations Policy⁷ and the procedures to implement it, also forming an AI Operations Board to which the Students Association elects a representative to maintain the student voice. The subsequent emergence a year later of generative AI tools such as ChatGPT⁸ introduced new ethical dilemmas related to learning, academic integrity, and procurement, around which a new Student Partnership in AI has continued consultations.⁹

In this study, the repercussions of co-production impacted individual, community, and knowledge systems scales. For example, intensive collective learning enabled study participants to quickly learn about a range of AAI-EdTech tools and ethical issues – a unique form of specialist insight and expertise which not only informed the principles, but which can also be mobilised across participant networks beyond the study. This can unsettle traditional notions of researcher-participant roles and sites of research, as the status of expertise is mobilised in different ways for particular contexts. Study participants viewed next steps of this initiative as *translating and adapting the principles into action with inter/transdisciplinary research that re-imagines expertise, procurement and student-staff support in light of social justice, wellbeing, and democracy (Theme 3 perspectives)*. For example, the value of an interdisciplinary approach was highlighted as key to ensuring the principles were informed by multiple perspectives, so as to keep the human 'in the equation', and to provide more diverse student and staff support (in terms of existing concerns or complaints about existing or introduced technologies). Bringing in more varied expertise

⁷ UTS Artificial Intelligence Operations Policy: <https://www.uts.edu.au/about/uts-governance/policies/uts-policy/artificial-intelligence-operations-policy>

⁸ OpenAI: ChatGPT <https://openai.com/blog/chatgpt>

⁹ Student Partnership in AI Ethics Consultation: <https://cic.uts.edu.au/projects/ai-ethics-consultation-2023>

to inform the procurement process was also identified as key, which accords with growing calls for broader cross-sectoral and public involvement in transdisciplinary research to address systemic and societal issues. Notably, participants also strongly communicated that values associated with social justice, wellbeing, and democracy should guide AAI-EdTech tools and ethics. How often abstracted concepts and values can be operationalised and mobilised across the distributed expertise of a university community and associated partnerships seems a rich area for future research.

Another impact upon the knowledge-system scale was the ongoing implications of the co-produced knowledge and principles within, and beyond, the institution. Collectively, participants stressed the importance that this initiative be an ‘ongoing conversation’ that expands to involve the broader university community. In addition, the openness to diversity and the value of human decision-making were seen as central components to inform a form of shared governance – or collective stewardship – that would guide the ongoing development and evaluation of AAI-EdTech tools. Of primary importance was that public trust and benefit should be at the heart of this collective control and care of AAI-EdTech tools for learning and teaching.

Recommendations for Designing Co-production

To inform learnings about the co-production process, we discuss key insights across the preparing, managing, and sustaining phases.

Preparing

Preparing for co-production “sets the foundation for the process” (Wyborn et al., 2019, p. 332) in terms of building relationships and making both processes and expectations transparent. In this study, the DD mini-public committed to shared learning about the topic from expert witnesses, appointed not only by subject matter experts supporting the process, but also by the mini-public. The mini-public was also coached on their critical thinking and teamworking skills. Study participants viewed the process as *an interactive information exchange which relied upon testing what is reasonable by empowering participants to transgress pre-conceived notions and generate new knowledge (Theme 2 contexts)*. Examples of resources which scaffolded shared learning included: critical thinking exercises, an information pack, an online communications platform where questions could be posed and answered, as well as a variety of whole group and break-out group activities with different experts.

Managing

The logistics, relations, and resource capacity of managing co-production can “increase capacity of participants while expanding the available knowledge base to solve a problem” (Wyborn et al., 2019, p. 334). The DMP committed to ‘rules of engagement’ including giving reasons for views, fairness, equality of voice,

and openness to difference (Farrell et al., 2019). The structured, closely facilitated process shaped this deliberative space where varied stakeholder views could be aired and explored so as to infuse the co-produced principles. In this study, the novelty of hearing simultaneously from a range of higher education stakeholders was viewed as valuable for better understanding multiple perspectives of AAI-EdTech – such as the experiences of students and educators which are not commonly addressed in procurement and strategic decision-making procedures. We propose that a particular type of ‘deliberative dialogue’ about specific and situated technological practices was enabled, whereupon: pre-conceived roles and beliefs could be momentarily suspended, an openness to different perspectives and experiences were accepted, and a trusted place was created where reasons for views could be surfaced and integrated to inform cooperation toward a common goal.

Practically speaking, DD brings financial costs, firstly of a facilitator (unless this is volunteered within the institution), plus fair recompense to students and casual staff for their time. Possible objections to this cost must be weighed against the risks and costs of any or all of the following: (i) losing the trust of the university community that the institution is deploying AAI-EdTech in an ethical manner; (ii) procuring technologies that inadvertently violate the university’s ethical principles due to inadequate consultation with students and staff on the use cases; and (iii) a breakdown in technology-enabled ethical behaviour on the part of staff or students, whether through ignorance or intent. We would assert that this is a conversation worth investing in, worth sustaining, and worth conducting to a high standard.

Sustaining

This phase draws critical attention to the broader contexts of co-production, in particular: “Failure to account for the institutional context in which a given intervention is situated may result in ostensibly successful projects meeting a dead end when it comes to creating lasting change” (Wyborn et al., 2019, p. 335). A related key aspect of DD is opening up the decision-making process to diverse stakeholders so as to increase the diversity of perspectives, ideas, and action. For example, as an idea to inform the future of AAI-EdTech tools and ethics, in this study participants learnt about *in-house, locally-designed tools that offer tailored, trusted systems which prioritise teaching and learning, offering greater control than third-party tools (Theme 3 contexts)*. These insights signal the potential of future research and funding opportunities focused upon the co-design of system/tool development within – and between – universities. While in-house AAI-EdTech *research* is widespread, scaling innovations to enterprise grade deployment with staff development is complex (Buckingham Shum, 2023) and most universities procure commercial services. In this regard, the DD consultation identified potential in *more participatory approaches to AI procurement* which involve a wider range of higher education stakeholders in the decision-making process. However, a key issue raised by some participants was the need to understand broader project linkages, implications of cost-effectiveness, and the increasing power of EdTech vendors over organisational

autonomy: this is a political economy dimension of AAI-EdTech that could be explored in future studies.

Co-production in this study is a demonstration of testing – and transgressing – existing preconceptions and knowledge claims about AAI-EdTech tools and ethics. For example, a participant recalled an expert who proposed the value of ‘testing’ to find out what is reasonable in terms of shaping the principles. We argue that the combined critical thinking and teamwork aspects of this study’s methodology helped to shape a ‘testing-ground space’ for cooperative learning about what is collectively reasonable in terms of institutional responsiveness to AAI-EdTech tools and ethical concerns. DD mini-publics are typically discrete events responding to particular dilemmas and controversies; however, scaffolding means that shared learning occurs not only within a given timeframe of a project or controversy, but also embeds supportive structures well beyond the initial consultation and participants.

Navigating Constructive Interactions with Society and Culture

Co-production in practice highlights the value of taking into account the “broad-reaching social and cultural norms that underlie knowledge systems, interventions, and even interactions between individuals” (Wyborn et al., 2019, p. 335). We propose that the COVID-19 pandemic brought to sharp relief an ongoing socio-technical controversy: how society, and educational institutions, utilise and consider emerging technologies for new forms and scales of personal and public surveillance, communication, and decision-making. A key imperative for the university sector in the COVID-19 era is institutional *responsiveness* to the introduction of AI-driven educational practices. Even post-lockdown, many universities are continuing with remote proctoring: which is viewed as the ‘new normal’. This particular controversy reflects broader concerns about technologies introduced during emergencies/crises. During such intense events, the rapid implementations of socio-technical innovations are often initially thought of as temporary measures – which then become normalised, and eventually mainstreamed. Even before the pandemic, there were growing calls for new ways to involve broader stakeholders in deliberation and decision-making about increasingly sophisticated AI and data-driven technologies which promise to offer new levels of responsiveness based upon data-driven personalisation, prediction, and pedagogy. This study responds to the limits of existing ethical approaches to emerging analytics and AI *enabling technologies* that underpin innovation across a range of products and domains – which are distinct from *stand-alone technologies* that are limited to one application domain (Brey, 2017).

Overall participants collectively perceived navigating the DD process as a *curiosity-driven, applied, and dynamic process* (*Theme 1 perspectives*). In doing so, the value of this paper’s focus upon dilemmas and specific socio-technical controversies foregrounds three key elements: first, discovering how emerging technologies function for particular education-related tasks; second, that there is often no neat solution to deal with the complexity of social and technical concerns, both foreseen and unforeseen; and third, that ethical implications are messy, uncertain, and always

contextual for particular stakeholders and organisations. A specific example was the procurement by many institutions of commercial services providing online, remote, and in some cases AI-automated, invigilation of examinations, which sparked widespread debate when the technology was rapidly applied worldwide to monitor university students during the pandemic lockdown (Coghlan et al., 2021; Sefcik et al., 2022). This is because we had the university's lead practitioner to explain the safeguards that were put in place for students who could not, or did not want to, use the technology – alongside researchers to explain how the system worked in relation to ethical implications. That session changed some DMP participants' perspectives about the role of ProctorU, especially the ways in which universities can manage the introduction of third-party software. Significantly, the expert witness noted that the session's productive dialogue stemmed from a clear focus upon deliberating the functionality, ethical issues, and application of the technology in a specific university context. The example of the ProctorU session demonstrates how an ethically sensitive tech product can be interrogated productively and pragmatically, when the right deliberative context is established.

Conclusions

The COVID-19 pandemic triggered a range of ethical concerns and responses to AAI-EdTech. The controversial dilemmas associated with AAI-EdTech span multiple issues, such as privacy, consent, access, and inequalities. This was the impetus for UTS to identify the need to institute principles and policies to address the particular ethical issues that can arise with AAI-EdTech; and a co-production process to engage the diverse community of students, tutors and academics in informed deliberation about their expectations and values with regard to AAI-EdTech.

Methodologies to investigate ethical issues associated with the rise of AAI-EdTech via a consultation process with diverse higher education stakeholders are under-theorised and under-examined. In this paper we have presented the potential of DD as a co-production approach, which can: (i) generate principles to address the particular ethical issues that can arise with AAI-EdTech; and (ii) facilitate a consultation process to engage a diverse community of students, tutors and academics in informed deliberation about their expectations and values with regard to AAI-EdTech, building rather than eroding their trust.

To our knowledge, this is the first application of DD for AI ethics, as is its use as an organisational sensemaking process in education. This work responds to calls for more interdisciplinary explorations of co-production and how it operates in practice (Bandola-Gill et al., 2022). In doing so, there is potential for future research to explore other co-production processes, such as 'technical democracy' (Callon et al., 2001) to understand its similarities, and differences to DD within educational contexts (Gulson et al., 2022; Thompson et al., 2022). There is also a recognised need to further pilot and extend the knowledge network of DD 'critical spaces' in higher education, which could involve collaboration between universities and communities across multiple countries (Mourad, 2022). More broadly, feminist and decolonial perspectives could also help to examine the conditions and limits of iterative design

approaches with communities, which requires reconfiguring research beyond traditional project funding and implementation life-cycles (Dourish et al., 2020).

On the evidence to date, Deliberative Democracy, even when conducted wholly online under COVID-19 lockdown conditions, would appear to offer educational institutions an approach to address the urgent need for meaningful student/staff consultation on the ethical implications of introducing AAI-EdTech into teaching and learning. The implementation process is now beginning, which we are tracking with equal interest. Members of the DD researcher/practitioner community have advised that the application of DD for institutional consultation, wholly online, and about this topic, is a combination of novel features that has not been seen within the DD research literature, so we are considering how best to engage that community. The DMP's work helped to catalyse a new university policy on AI ethics, and work is now well under way designing/evolving the governance instruments and procedures to ensure that this translates into action.

In closing, we propose that our evaluation of this novel methodology offers valuable co-production insights to inform university sector responsiveness to AAI-EdTech. We have discussed the specific context and constraints under which this study was conducted, but maintain that the issues and findings transcend the specific context, which operates in a similar manner to many other universities. We hope this multi-level process and its outcomes will be of wider interest to different university communities and the education sector more broadly, as a novel and productive way to co-produce AAI-EdTech ethics with diverse stakeholders.

Acknowledgements We would like to thank the participants interviewed for this study, as well as all the stakeholders involved in the Deliberative Democracy mini-public. We also appreciate the organisational support of Gabrielle Gardiner, Senior Manager at the UTS Connected Intelligence Centre, and the workshop design and facilitation expertise of Chad Foulkes from Liminal By Design. Our thanks also to Susan Gibson for her leadership of the UTS AI Operations Policy that followed the deliberative democracy workshops. In addition, we acknowledge the funding support of University of Technology Sydney, University of Sydney, and the Australian Research Council.

Author Contributions All authors contributed to the study conception and design. Interview data collection and analysis were performed by Teresa Swist. All authors contributed to drafting, reviewing, and approving the final manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions This project was funded by University of Technology Sydney with additional funding provided by University of Sydney and the Australian Research Council.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission

directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., & Kankanhalli, M. (2018). *Trends and Trajectories for Explainable, Accountable and Intelligible Systems*. Paper presented at the Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/3173574.3174156>
- Baker, R. S., & Hawn, A. (2021). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00285-9>
- Bandola-Gill, J., Arthur, M., & Ivor Leng, R. (2022). What is co-production? Conceptualising and understanding co-production of knowledge and policy across different theoretical perspectives. *Evidence & Policy*, 1–24. <https://doi.org/10.1332/174426421x16420955772641>
- Barratt-See, G., Cheng, M., Deakin Crick, R., & Buckingham Shum, S. (2017). *Assessing resilient agency with CLARA: Empirical findings from piloting a visual analytics tool at UTS*. Paper presented at the proceedings UniSTARs 2017: University students, transitions, achievement, retention & success, Adelaide, 1–4 July, 2017.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brey, P. A. E. (2017). Ethics of emerging technologies. In S. O. Hansson (Ed.), *The ethics of technology: Methods and approaches* (pp. 175–192). Rowman & Littlefield International.
- Buckingham Shum, S. (2023). embedding learning analytics in a university: Boardroom, staff room, server room, classroom. In O. Viberg & Å. Grönlund (Eds.), *Practicable Learning Analytics*. SpringerNature.
- Buckingham Shum, S., & Deakin Crick, R. (2012). *Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada. <https://doi.org/10.1145/2330601.2330629>
- Buckingham Shum, S., & McKay, T. (2018). Architecting for learning analytics: Innovating for sustainable impact. *Educause Review, March/April*, 25–37. Retrieved from <https://er.educause.edu/articles/2018/3/architecting-for-learning-analytics-innovating-for-sustainable-impact>. Accessed 31 Jan 2024.
- Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, 6(2), 1–9. <https://doi.org/10.18608/jla.2019.62.1>
- Callon, M., Lascoumes, P., & Barthe, Y. (2001). *Acting in an uncertain world: An essay on technical democracy (Translation by Graham Burchell. 2009)*. The MIT Press.
- Carson, L., & Hartz-Karp, J. (2007). Adapting and combining deliberative designs: Juries, polls, and forums. In J. Gastil & P. Levine (Eds.), *The deliberative democracy handbook: Strategies for effective civic engagement in the twenty-first century*. Jossey-Bass.
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good proctor or “big brother”? Ethics of online exam supervision technologies. *Philosophy & Technology*. <https://doi.org/10.1007/s13347-021-00476-1>
- Davis, J. L., Williams, A., & Yang, M. W. (2021). Algorithmic reparation. *Big Data & Society*, 8(2), 205395172110448. <https://doi.org/10.1177/20539517211044808>
- Dawson, P. (2021). *Defending assessment security in a digital world*. Routledge.
- Diakopoulos, N. (2014). Algorithmic accountability. *Digital Journalism*, 3(3), 398–415. <https://doi.org/10.1080/21670811.2014.976411>
- Dollinger, M., Liu, D., Arthars, N., & Lodge, J. (2019). Working together in learning analytics towards the co-creation of value. *Journal of Learning Analytics*, 6(2), 10–26. <https://doi.org/10.18608/jla.2019.62.2>
- Dourish, P., Lawrence, C., Leong, T. W., & Wadley, G. (2020). On being iterated: The affective demands of design participation. In *Paper presented at the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)* (pp. 1–11). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376545>

- Dryzek, J. S. (2010). *Foundations and frontiers of deliberative governance*. Oxford University Press.
- Elstub, S., Ercan, S., & Mendonça, R. F. (2016). Editorial introduction: The fourth generation of deliberative democracy. *Critical Policy Studies*, 10(2), 139–151. <https://doi.org/10.1080/19460171.2016.1175956>
- Farrell, D., Curato, N., Dryzek, J. S., Geißel, B., Grönlund, K., Marien, S., . . ., & Setälä, M. (2019). *Deliberative mini-publics: Core design features*. Retrieved from <https://www.governanceinstitute.edu.au/magma/media/upload/ckeditor/files/Deliberative%20Mini-Publics%20Core%20Design%20Features.pdf>. Accessed 31 Jan 2024.
- Fernandez-Nieto, G., Echeverria, V., Buckingham Shum, S., Mangaroska, K., Kitto, K., Palominos, E., . . ., & Martinez-Maldonado, R. (2021). Storytelling with learner data: guiding student reflection on multimodal team data. *IEEE Transactions on Learning Technologies*, 14(5), 695–708. <https://doi.org/10.1109/tlt.2021.3131842>
- Fernandez-Nieto, G., Kitto, K., Buckingham Shum, S., & Martinez-Maldonado, R. (2022). *Beyond the learning analytics dashboard: Alternative ways to communicate student data insights combining visualisation, narrative and storytelling*. Paper presented at the LAK22: 12th international learning analytics and knowledge conference. <https://doi.org/10.1145/3506860.3506895>
- Filipe, A., Renedo, A., & Marston, C. (2017). The co-production of what? Knowledge, values, and social relations in health care. *PLOS Biology*, 15(5), e2001403. <https://doi.org/10.1371/journal.pbio.2001403>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., . . ., & Vayena, E. (2018). AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gulson, K. N., Thompson, G., Swist, T., Kitto, K., Rutkowski, L., Rutkowski, D., . . ., & Knight, S. (2022). *Automated essay scoring in australian schools: Collective policymaking*. <https://doi.org/10.25910/e8nn-q994>
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Hanna, A., Denton, E., Smart, A., & Smith-Loud, J. (2020). *Towards a critical race methodology in algorithmic fairness*. Paper presented at the Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain. <https://doi.org/10.1145/3351095.3372826>
- Herodotou, C., Hłosta, M., Boroowa, A., Rienties, B., Zdrahal, Z., & Mangafa, C. (2019). Empowering online teachers through predictive learning analytics. *British Journal of Educational Technology*, 50(6), 3064–3079. <https://doi.org/10.1111/bjet.12853>
- Holmes, W., & Porayska-Pomsta, K. (Eds.). (2022). *The ethics of artificial intelligence in education: Practices, challenges, and debates*. Taylor & Francis.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., . . ., & Koedinger, K. R. (2021). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00239-1>
- Holstein, K., Vaughan, J. W., Daumé, H., Dudík, M., & Wallach, H. (2019a). Improving fairness in machine learning systems: What do industry practitioners need? In *Paper presented at the 2019 CHI Conference on Human Factors in Computing Systems* (Vol. 600, pp. 1–16). <https://doi.org/10.1145/3290605.3300830>
- Holstein, K., McLaren, B. M., & Aleven, V. (2019b). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics*, 6(2), 27–52. <https://doi.org/10.18608/jla.2019.62.3>
- IEEE. (2017). *Ethically aligned design*. Retrieved from <https://standards.ieee.org/industry-connections/ec/autonomous-systems/>. Accessed 31 Jan 2024.
- Jasanoff, S. (2004). *States of knowledge: The co-production of science and the social order*. Routledge.
- Johanes, P., & Thille, C. (2019). The heart of educational data infrastructures = conscious humanity and scientific responsibility, not infinite data and limitless experimentation. *British Journal of Educational Technology*, 50(6), 2959–2973. <https://doi.org/10.1111/bjet.12862>
- Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Journal of Educational Technology*, 50(6), 2855–2870. <https://doi.org/10.1111/bjet.12868>
- Kizilcec, R., & Lee, H. (2022). Algorithmic fairness in education. In W. Holmes & K. Porayska-Pomsta (Eds.), *The ethics of artificial intelligence in education: Practices, challenges, and debates*. Taylor & Francis.

- Knight, S., Shibani, A., Abel, S., Gibson, A., Ryan, P., Sutton, N., . . . , &, 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>
- Li, W., Sun, K., Schaub, F., & Brooks, C. (2021). Disparities in students' propensity to consent to learning analytics. *International Journal of Artificial Intelligence in Education*, 32, 564–608. <https://doi.org/10.1007/s40593-021-00254-2>
- Lim, L.-A., Gasevic, D., Matcha, W., Uzir, N. A. A., & Dawson, S. (2021). *Impact of learning analytics feedback on self-regulated learning: Triangulating behavioural logs with students' recall*. Paper presented at the LAK21: 11th International Learning Analytics and Knowledge Conference, Irvine, CA, USA. <https://doi.org/10.1145/3448139.3448174>
- Lim, L.-A., Atif, A., & Farmer, I. (2022). 'Made good connections': Amplifying teacher presence and belonging at scale through learning design and personalised feedback. Paper presented at the Proceedings ASCILITE 2022: 39th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education, Sydney.
- Mourad, R. (2022). Deliberative democracy in higher education: The role of critical spaces across universities. *Journal of Deliberative Democracy*, 18(1), 1–10. <https://doi.org/10.16997/jdd.954>
- Porayska-Pomsta, K., Woolf, B., Holmes, W., & Holstein, K. (2021). The FATE of AI in education: Fairness, accountability, transparency, and ethics. *International Journal of Artificial Intelligence in Education (Special Collection)*. Retrieved from https://link.springer.com/journal/40593/topic/aiCollection/AC_dcac58fbff2e68a27dd420b8fa69ba47. Accessed 31 Jan 2024.
- Prinsloo, P. (2019). A social cartography of analytics in education as performative politics. *British Journal of Educational Technology*, 50(6), 2810–2823. <https://doi.org/10.1111/bjet.12872>
- Rességuier, A., & Rodrigues, R. (2020). AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data & Society*, 7(2), 2053951720942541. <https://doi.org/10.1177/2053951720942541>
- Richards, D., & Dignum, V. (2019). Supporting and challenging learners through pedagogical agents: Addressing ethical issues through designing for values. *British Journal of Educational Technology*, 50(6), 2885–2901. <https://doi.org/10.1111/bjet.12863>
- Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open*, 2(4), 1–12. <https://doi.org/10.1177/2332858416673968>
- Sefcik, L., Veeran-Colton, T., Baird, M., Price, C., & Steyn, S. (2022). An examination of student user experience (UX) and perceptions of remote invigilation during online assessment. *Australian Journal of Educational Technology*, 38(2), 49–69. <https://doi.org/10.14742/ajet.6871>
- Selwyn, N. (2019). What's the problem with learning analytics? *Journal of Learning Analytics*, 6(3), 11–19. <https://doi.org/10.18608/jla.2019.63.3>
- Shibani, A., Knight, S., & Buckingham Shum, S. (2020). Educator perspectives on learning analytics in classroom practice. *The Internet and Higher Education*, 46. <https://doi.org/10.1016/j.iheduc.2020.100730>
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Sun, K., Mhaidli, A. H., Watel, S., Brooks, C. A., & Schaub, F. (2019). It's my data! Tensions among stakeholders of a learning analytics dashboard. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 594, 1–14. <https://doi.org/10.1145/3290605.3300824>
- Thoeming, A., Baird, A., Weeks, R., Liu, D. Y. T., Flitcroft, H., & Bridgeman, A. J. (2022). *Diverse definitions of engagement: Personalised learning analytics to support staff and students*. Paper presented at the Proceedings ASCILITE 2022: 39th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education, Sydney.
- 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*, 1–13. <https://doi.org/10.1080/17439884.2022.2126495>
- Vereschak, O., Bailly, G., & Caramiaux, B. (2021). How to evaluate trust in AI-assisted decision making? A survey of empirical methodologies. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), Article 327. <https://doi.org/10.1145/3476068>
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., . . . , & Schwartz, O. (2018). *AI Now Report*. Retrieved from https://ainowinstitute.org/AI_Now_2018_Report.pdf. Accessed 31 Jan 2024.
- Williamson, B. (2019). Policy networks, performance metrics and platform markets: Charting the expanding data infrastructure of higher education. *British Journal of Educational Technology*, 50(6), 2794–2809. <https://doi.org/10.1111/bjet.12849>

- Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, 1–13. <https://doi.org/10.1080/17439884.2020.1798995>
- Wyborn, C., Datta, A., Montana, J., Ryan, M., Leith, P., Chaffin, B., . . ., & Kerkhoff, L. v. (2019). Co-producing sustainability: Reordering the governance of science, policy, and practice. *Annual Review of Environment and Resources*, 44(1), 319–346. <https://doi.org/10.1146/annurev-envir-on-101718-033103>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Teresa Swist^{1,2}  · Simon Buckingham Shum³  · Kalervo N. Gulson¹ 

 Teresa Swist
teresa.swist@sydney.edu.au

Simon Buckingham Shum
simon.buckinghamshum@uts.edu.au

Kalervo N. Gulson
kalervo.gulson@sydney.edu.au

¹ Sydney School of Education and Social Work, University of Sydney, Camperdown, NSW, Australia

² Institute for Culture and Society, Western Sydney University, Parramatta, NSW, Australia

³ Connected Intelligence Centre, University of Technology Sydney, Ultimo, NSW, Australia