

NSW Department of Education

# FUTURE EDgE

Issue 3 / December 2020



**Teaching and technology:**  
educating in a machine age







© The authors and the State of New South Wales (Department of Education), 2020. The views expressed in these articles are solely those of the authors and do not represent the views of the NSW Department of Education.



# FUTURE EDge

## Contents

About the Authors.....	4
<b>Foreword</b>	
<b>Sarah Mitchell MLC</b> .....	8
<b>Advancing learning through AI:</b> insights from a NSW teacher-educator and emerging technology researcher <b>Erica Southgate</b> .....	12
<b>Teaching privacy and ethical guardrails for the AI imperative in education</b> <b>Evan Selinger and Amelia Vance</b> .....	30
<b>In conversation with the 3A Institute:</b> designing ethical technology for a better future <b>Genevieve Bell, Amy McLennan and Leslie Loble</b> .....	56
<b>Learning together:</b> using technology to support expert teaching practice <b>Mark Scott</b> .....	72



# About the Authors



## **Sarah Mitchell MLC**

Sarah Mitchell is the Minister for Education and Early Childhood Learning and Deputy Leader of the Government in the NSW Legislative Council. Sarah has been a member of the NSW Legislative Council since March 2011.



## **Erica Southgate**

Erica Southgate is Associate Professor of Emerging Technologies for Education at the University of Newcastle, Australia. She is the Lead Researcher on the VR School Study, the first research to embed highly immersive virtual reality into school classrooms. Erica is also a maker of award-winning computer games for literacy learning, and the lead author of the recent Australian Government commissioned report, 'Artificial Intelligence and Emerging Technologies in Schools'



## **Evan Selinger**

Evan Selinger is a Professor of Philosophy at the Rochester Institute of Technology and Affiliate Scholar at Northeastern University's Center for Law, Innovation, and Creativity. His research focuses on technology governance, primarily ethics and privacy. The Observer newspaper selected 'Re-Engineering Humanity', which Evan co-authored with Brett Frischmann, as one of Thirty Books to Help Us Understand the World in 2020.



## Amelia Vance

Amelia Vance is a Director of Youth and Education Privacy at the Future of Privacy Forum. Amelia advises on child and student privacy law and best practices to ensure the responsible use of student data and education technology in schools in the US and internationally. She is part of the OECD expert group currently reviewing the draft Recommendation on Children in the Digital Environment.



## Genevieve Bell

Genevieve Bell AO FTSE is the Director of the 3A Institute (3Ai), Florence Violet McKenzie Chair, and a Distinguished Professor at the Australian National University as well as a Vice President and Senior Fellow at Intel Corporation. She is a cultural anthropologist, technologist and futurist best known for her work at the intersection of cultural practice and technology development.



## Amy McLennan

Dr Amy McLennan is a Research Fellow at the 3A Institute, where she works at the intersections of technology, society and wellbeing. She is also one of four Tuckwell Fellows at the ANU. Amy is trained in biomedical science, human nutrition and medical anthropology, and has been affiliated with the University of Oxford for over a decade, where her work has focused on food systems, nutrition, obesity, health policy and human ecology.



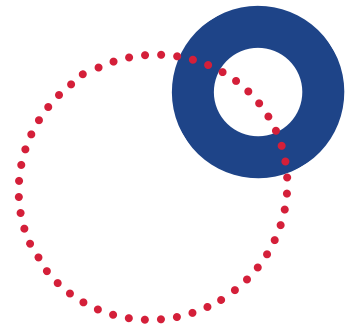
## Leslie Loble

Leslie Loble, recently retired as Deputy Secretary in the NSW Department of Education, has led strategy, reform and innovation in Australia's largest and most diverse education sector for nearly two decades. Leslie instigated Education for a Changing World, a major initiative to ensure that education meets students' needs in a future of artificial intelligence and global change.



## Mark Scott

Mark Scott AO is the Secretary of the NSW Department of Education and has a distinguished record in public service, education and the media. Mark's career began as a teacher in Sydney. Mark held a number of senior editorial roles at Fairfax. From 2006 to 2016, Mark was Managing Director of the ABC and led the organisation's transformation to be a public broadcaster in the digital era.







# Foreword

Sarah Mitchell MLC

It's hard to believe we've already reached the end of the 2020 school year. We started the year gripped by bushfires, intently focused on the immediate safety and wellbeing of students in affected regions. Little did we imagine the challenges that lay ahead with COVID-19 – social distancing, pivoting to remote learning and ensuring we were meeting the needs of students during a period of almost daily flux. The pandemic meant our education system needed to respond at speed to maintain continuity of learning for children and young people.

The fact that we achieved this is a credit to all of the early childhood educators, teachers, principals, school support staff and those working behind the scenes who made sure that students' learning came first. Thousands of milestones were still reached in 2020: children wrote their names without help, read chapter books cover to cover, carried out scientific investigations and completed their HSC.

One of the unexpected upsides of remote learning for many parents was the window it afforded into the classroom of 2020. For many of us, it allowed us to watch our children explore new ways of learning and new technologies. The vast majority of NSW schools used digital platforms to keep teaching and learning on track and, as Minister, emerging educational technology is something I have been thinking about more than ever. I know our school leaders and teachers





**Student safety, privacy and wellbeing is paramount before we introduce any new learning tool into the classroom – whether it’s metalworking equipment or a machine learning algorithm.**

have been doing the same. This publication is timely in exploring the great potential of these technologies, including artificial intelligence (AI), for us in education. At the same time, it reinforces that student safety, privacy and wellbeing is paramount before we introduce any new learning tool into the classroom – whether it’s metalworking equipment or a machine learning algorithm.

Sophisticated programs that can help students understand and practise essential concepts across maths, science, English and the humanities are increasingly available. Through machine learning, voice and text recognition, these programs adjust to suit a student’s level and learning pace. With regard to assessment, the OECD already uses algorithms to improve the precision of its Programme for International Student Assessment (PISA) testing: adapting and selecting questions in real time to pinpoint students’ strengths and weaknesses.

Having recently launched the Department of Education’s Inclusive Education Statement, I am very interested in the capacity of advanced digital tools to support students with disability and additional learning needs. Then there is AI’s

**It is great teachers who inspire students and unlock doors to new knowledge in a way that no machine ever will.**

unprecedented power to process and sort large amounts of data, which we are seeing across many industries. We have done a lot of work already to reduce the time our school leaders and teachers spend on administration – so they can devote all their energy to the core business of teaching and learning – and new technologies can no doubt help us to cut administrative tasks even further.

We will use these new technologies in our schools and across Government wherever it's appropriate to do so, but not before we have considered all the implications. That means asking some fundamental questions of any new technology and knowing where to draw the line. Can we absolutely ensure the privacy and security of students and staff members? Are we certain it is the best and most appropriate tool for the job of helping students to learn and improve, or making our processes more efficient? Will it help to make education fairer and more inclusive? And can we very clearly explain to students and community members how it works, just as we explain anything else we use to support learning in our classrooms?

It is also important to stress that technology is just a tool, as Department Secretary Mark Scott points out here. This year brought transformations in the 'how' and the 'where' of teaching, but it did not change what's at the heart of learning. When I think of myself at school in country NSW, I can hardly remember a time I didn't have a book in my hand. It was my teachers who ignited that passion for reading, which then became a gateway to new learning for me. For today's students, it is no different. Core skills like literacy and numeracy will be their enduring passport to further education, rewarding jobs and lifelong learning in a changing world. Great teaching is the key to those skills; and it is great teachers who inspire students and unlock doors to new knowledge in a way that no machine ever will.

We owe it to every student and teacher across our state to embrace and provide equitable access to what new technologies can offer, and we will. However, we will do so with prudence and care: never losing sight of what, and who, is most important.







# Advancing learning through AI:

insights from a NSW teacher-educator and emerging technology researcher

Erica Southgate





The Future EDge team recently spoke with Associate Professor Erica Southgate, University of Newcastle, about ethical and other implications of emerging technologies for teachers and students. In this interview, Erica shares her insights about artificial intelligence (AI), some ethical implications it raises, and how it might be used to help enrich students' learning. Erica also makes the case for why teachers and students, alongside technology experts and policymakers, should be active contributors to the conversation on how AI should be used in our schools. This article provides an edited form of that interview.

**Erica – you're a teacher educator and researcher who's thought a lot about how we might use technology to assist with education. Could you tell us a little bit about yourself, and how you became interested in this area of research?**

I'm an Associate Professor of emerging technologies for education, which is a really cool and wonderful field, at the University of Newcastle. I'm the lead researcher of the VR (Virtual Reality) School Project, which

was a world-first study to embed high end virtual reality into school classrooms – working with Callaghan College and Dungog High School here in New South Wales. I'm really interested in technology ethics, particularly around the ethical use of immersive learning experiences – things like virtual, augmented and mixed reality – as well as the use of AI, and how we might govern the use of AI in schools. I'm also interested

in how emerging technologies like AI use biometric ('of the body') data, and what this might mean for privacy and human rights. Whatever is on the horizon catches my eye – it might be interesting or efficacious for learning, anything that has ethical implications – then I'm on it.

**You are the lead author for a recent report for the Australian Government on the implications of AI and other emerging technologies for schools (Southgate et al., 2018). Could you describe what makes AI-based technology so different to 20th-century educational software and other technologies?**

It's always good to start by defining what AI is. The OECD defines AI as a machine-based system that can make predictions, recommendations, or decisions, for a given set of human objectives. AI can influence real or virtual environments with varying levels of autonomy. Machines can learn and think by themselves, with varying levels of independence from human oversight.

AI has been around since the 1950s, but it has recently taken off because of technological improvements, such as cloud computing and the ability of software engineers and programmers to share code and learn online together. There have also been significant advances – for example, in the way machines sense and capture data visually and process language – which have supported this acceleration.

**I think machine learning raises profound questions about the role of machines in our lives and the extent to which they should make decisions for and about us.**

An important subfield of AI is machine learning, which is devoted to getting machines to learn by themselves by harvesting data (Southgate, 2020). Sometimes humans might label the data and sometimes the machines can label it and interpret the data themselves, to make categorisations and predictions. So, machine learning helps AI to learn and make decisions by itself.

I think machine learning raises profound questions about the role of machines in our lives and the extent to which they should make decisions for and about us. There are certainly other technologies that have also influenced human behaviour – there's no doubt about that. Anything from the pencil, pen or even the ink nib – that's technology. Books are also a form of technology that have certainly had a huge influence. And of course, computers as well. Now, however, we really need to think about what it means to live in a machine age, when machines can influence what we know and can do and, some would argue, even affect our life opportunities.





## To what extent is AI already in our schools?

It's important to understand that AI can be used in different ways, including in recommender, predictive and adaptive systems (Berendt, Littlejohn & Blakemore, 2020). AI powers the data and analytics within systems, the data mining, interpretation, categorisation, clustering of data and the modelling of these data. So, AI can sit at the back end of a system as well as having user-facing applications.

AI is at play in noise suppression, for instance, when we use teleconferencing. It also powers the everyday applications that we all use, such as internet search engines. AI does this very well and at scale. AI organises, captures and presents information at a speed humans couldn't keep up with. AI is very good at using data. Just the other day, when I opened up a PowerPoint presentation for a lecture I was developing, a new function called 'design ideas' popped up and this is powered by AI. The machine had read the document, interpreted the visual layout and the algorithm presented me with a beautiful piece of graphic design to enhance the presentation of my slides. Then I became so interested that I kept on looking at different design ideas, and didn't finish the PowerPoint presentation!

AI is powering recommender systems, like chatbots, that can help us navigate information. If we want to find out where something is, a chatbot might pop up, might

vocalise or communicate through text on a screen. It will tell us where we can find information and make recommendations. We're all very familiar with this in terms of online advertising – recommender AI is everywhere. It harnesses our data in real time, and it recommends stuff to us. This kind of AI could be used in schools to make recommendations to students or teachers about where to find information. A recommender system, for example, might assist an educator to find out about policy. AI can provide something a bit more fine-tuned than an ordinary key word search function and might recommend which policies we need to look at or the potential policy implications of something.

Like many other technologies, AI is being adopted by higher education before school education. This includes AI in plagiarism detection software and its use in organisational functions like learning platforms. In these platforms, university student data is gathered, analysed and turned into data analytics by the learning management system. It's visualised in terms of a dashboard and it may include predictive functions. The educator can see which students may be behind in their online work, who may be ahead in terms of listening to or viewing information and who hasn't logged in for a while and might be at risk of failure. These kinds of analytics are available in higher education and they're coming through learning management systems to school education.

**AI is already here. It's being experimented with, prototyped, rolled out to various degrees in education ... and we can see more on the horizon.**

Then there are adaptive systems. These are the sort of systems which computer scientists dream of building, intelligent tutoring systems, for instance, which adapt and personalise learning. So, let's say we do our learning online: the tutoring machine will know by analysing our responses when we don't exactly understand something; it'll question us; it'll then adapt the curriculum or the pedagogical approach so that we learn better. These types of systems have already been built, some of them work reasonably well within certain domains of knowledge, and some of them don't. They are definitely out there and being developed right now.

An extension of this is the idea of developing pedagogical agents. These would be 'helpers' in particular applications or systems that would assist us in learning in a very personalised way, possibly used throughout our whole lives not just in school. In order to help us learn what we need to know or do, these would be tracking and recording our lifelong learning journey – and not everyone might want that – but there's

work going on now on developing pedagogical agents within systems and virtual learning environments to support learning.

So, whether it's user-facing or powering the back end of applications, AI is already here. It's being experimented with, prototyped, rolled out to various degrees in higher education. Some AI is used in schools too through everyday applications such as PowerPoint or teleconferencing. It's here and we can see more on the horizon. The speed at which it will come into schools, I can't predict. But I don't think we should wait too long to develop our foundational understanding of what AI is and how machine learning works so that we can demystify the technology and ask informed and even critical questions about it.

**As a teacher yourself (of adults), are there things you wouldn't want AI to be doing or automating for you? Are there risks in allowing it to do too much?**

An important thing to note is that not all AI presents the same level of risk. When I'm interacting with the PowerPoint design ideas function, the only risk I have is that I'll be diverted off my task of finishing my lecture presentation. There's a risk that I'll become so interested in how beautiful it is, I won't be able to finish it. However, if we were looking at using particular systems with students for learning, we need to ask a lot of questions about that system for learning before we start using it.





I think we've got some good recent examples such as 'Robodebt'<sup>1</sup>, when people didn't think through automated decision-making enough in advance. They didn't think through the complexity of the technology and how to understand and interrogate it critically. And they didn't foresee the kind of impact that could have on humans. As a result, there are now some very good guidelines on using automated systems in public administration, which were produced by the Commonwealth Ombudsman. But these would have been better developed before the technology was deployed and people negatively impacted, of course.


I'm also reminded of the example of the robo-marking controversy from a few years ago, when ACARA proposed that computerised essay marking could be used for NAPLAN (Australia's national literacy and numeracy tests). There were quite potent arguments made about whether the scoring of a skill such as writing was best done by a teacher or a computer. In the end, teachers continued to do the marking. Really, what should have taken place was a much broader and deeper community and school-based conversation first. This conversation needed to include the teaching profession, technologists and ethicists, to decide what would be useful, alongside evidence of effectiveness for learning and appropriateness for assessment.

**We always need to ask: what are the consequences if that task is undertaken by a machine or automated system, rather than a teacher? What's the evidence that using these technologies is effective for learning? What are the human implications? What are the equity implications?**

I always say that the very foundation of education is the ability to explain stuff. As a teacher, I need to explain content. I need to explain skills-based learning or procedural learning. I need to explain my pedagogical decision-making, and relate that to evidence. I need to explain assessment and my grading and my whole assessment dynamic, for instance, through the development of criteria-based assessment with rubrics. I need to explain how that works to students. I need to explain why I might develop curriculum around a particular model versus another model. There's a whole lot of stuff I need to explain: not just content, but the very craft of my teaching.

If we use systems in education, we need to be able to explain why a machine made a particular decision. If we can't explain that – whether

1. 'Robodebt' refers to an automated debt recovery system used by the Australian Government's Services Australia in 2016. The automated data-matching system linked Centrelink records with averaged income data from the Australian Taxation Office to issue debt notices. It was heavily criticised, the subject of Commonwealth Ombudsman investigation and two Senate committee inquiries, and ultimately abolished by the Australian Government in 2020.



that's due to the complexity of the machine learning, a lack of access to data used to train the machine to understand potential error or bias, or a lack of transparency from companies that have proprietary algorithms – then this is problematic. We need to understand the decisions made by an intelligent tutoring system – why one student was given a particular curriculum pathway, where another wasn't. As teachers, we need to be able to explain why in a system certain students were, for example, flagged as 'at risk' or categorised in potentially stigmatising ways, because that can have real impacts on them as humans and their life opportunities. If companies or vendors creating this software have got proprietary algorithms and won't open them up transparently for independent expert analysis, and there aren't functions built into those systems where the machine can outline or audit its own decision-making process, then fundamentally, the technology undermines explainability as a foundation of education. As teachers, we need to guard against that.

Also, everyone has a right to privacy. Everyone has a right, for instance, to bodily integrity. When there is a suggestion to use facial recognition technology for roll call in schools, is that a good use of AI? Is this a good use of harvesting people's bodily information and behavioural information for a task which teachers do in the morning? Teachers also use that time to connect and learn about

their students and what's going on. They do it efficiently. We always need to ask: what are the consequences if that task is undertaken by a machine or automated system, rather than a teacher? What's the evidence base that using these technologies is effective for learning? What are the human implications? What are the equity implications?

So, I would argue that we need to be curious, we need to have questions, and we need to have dialogue when it comes to using these emerging technologies in our classrooms. If a vendor can't explain, using a robust peer-reviewed evidence base, the effectiveness of a particular application for learning, then we shouldn't use the product. We, as teachers using the technology, need to be able to explain and potentially also predict when things may go wrong.

### **Do you see potential for these emerging technologies to support students with additional learning needs and disability in particular?**

AI is quite amazing in terms of what it might do for people with disability or who are differently abled. For example, it could be used as an adaptive system or to provide pedagogical agents – those little helpers that could be embedded in particular applications or systems, which I mentioned earlier. These helper agents could assist in learning in a very personalised way, and this may be particularly beneficial for people with different cognitive capacities.

Things like computer vision (a type of AI) will be really useful for people with visual impairment, for example, to be able to tell them what's around the room. In fact, there are already applications like that.

And there's the potential that natural language processing can be further developed so that people who speak differently because they have a disability can use online tools very effectively, by searching verbally rather than through typing.

But, of course, using any type of new technology raises potential issues too.

In terms of natural language processing, for example, its functionality depends on the integrity and the diversity of the data sets used and the data that are being harvested to train the AI model. There is evidence that people who have accents or who speak differently because they have a disability, or because they come from a different socio-cultural background, or even because they're children and have higher voices, aren't recognised or understood by AI systems. This is because these systems are typically trained on culturally normative data sets and only with adults.

We need to understand that these types of technologies are built on data, they're built on big data, built on big data flows, harvested, and interpreted by both humans and machines. This means, when it comes to the potential of AI technologies to offer tools to help with disability, we really do need to very clearly understand the data


process: the data that goes in, the inputs, and then what's coming out in terms of the decision-making the machine does to assist the person, as well as what's going on between the input and the output, that is, the algorithmic process.

**In your recent report for the Australian Government, you've pointed to the types of risks that AI and other emerging technologies raise, including bias, privacy, error and transparency, which you've also touched on in this interview. When using these technologies, what else do we need to keep front of mind?**

There's a lot of literature on AI ethics; and many have argued, for example, that 'black box AI' should not be used in sensitive, human domains, such as criminal justice, welfare and education. Black box AI is where we don't understand why the machine or algorithm is making the decisions that it is. For example, if the algorithms are proprietary – where a business or government won't let us look inside the algorithmic process 'box' (between data inputs and outputs) and undertake independent expert analysis of this. This is because of the vulnerability of the humans in these systems, where people impacted most by these technologies aren't necessarily powerful enough or knowledgeable enough to challenge it – there are power and informational differentials.

Even if you understand this technology, it's very difficult to challenge the system, because you need to be able to understand how algorithms work. You need a very





**There's a pedagogical project here, where we start to learn about what AI is, what it isn't, what it can do, what it can't do and what the evidence base is.**

deep knowledge of statistics and mathematics to be able to explore this. Or you need access to someone who's an independent expert to translate that for you – someone who's not part of the manufacturing or vendor industry process. If we haven't got a level of independent expertise to do that in schooling systems, then that's problematic because the system leaves itself open to regulatory capture. This means it trusts the judgement and the information provided by the vendor because it hasn't got the depth of technical expertise to be able to make or access independent advice. I think we really do need to think hard about this, particularly for vulnerable populations. Who are the independent experts here who can provide oversight?

Interestingly, many of the debates around ethics and AI are being driven by people who are vulnerable to being negatively impacted by these technologies – and that includes people with disability or who are differently abled, as well as other minority or under-represented groups, including people of colour and Indigenous people, and women. People in these communities are

working very hard to highlight the problem of how AI can learn and amplify bias in its use of data, to ensure that we do get unbiased AI – or at least we can have informed conversations about it. We absolutely must include people from diverse backgrounds and different points of view in these conversations, when seeking to democratise the ethical design, use and governance of these emerging technologies.

**What sort of skills do you think teachers will need, in that case, if we're going to have more AI in our classrooms?**

Well, I think it's about learning as a profession. There's a pedagogical project here, where we start to learn about what AI is, what it isn't, what it can do, what it can't do and what the evidence base is. We begin to demystify it, so it doesn't feel like magic, but rather, a particular machine-human process. We can't expect individual teachers to master this overnight. But we can begin a pedagogical project. As teachers, we're good at pedagogy. We understand teaching, we know when new things come along that we need to understand and be able to explain and engage with them. And so, the teaching profession is very well-placed to begin to grapple with the technology and the introduction of it into classrooms.

We also need to ensure teachers have access to independent experts to provide technical advice on the computer science and statistical aspects but also the ethical and

**We need to deeply engage with the complexity around these emerging technologies, so that we get good outcomes for the teaching profession, for our students that we care so much about, and for our communities more broadly.**

governance aspects of this type of technology. For instance, to help ask and answer important questions when school leaders or educators are looking to purchase new technologies. Whether it's a schooling system or individual schools procuring a particular application or platform that uses AI – are we aware if AI is present and understand how it's used? People need to understand, for instance, that from a legal perspective, biometric data is considered sensitive data and has to be handled in particular ways. People will need to be aware of the regulatory frameworks and privacy impacts relating to the harvesting, storing and use of sensitive data. In terms of schooling systems, we need to look at privacy policies and ask whether they sufficiently address these kinds of issues. What do the data sharing arrangements with vendors look like for new technologies? If I was a parent, could I independently go and find out, on a website or by contacting someone, what the processes and dynamics are? It's a complex field and these are


complex issues. We can't just treat this as a 'ticking a box' exercise, and so leadership and transparency are important in this space.

I think it's critical that the teaching profession starts to become really active in these conversations. There are many issues, such as bias and error and accountability that we need to interrogate so that we can start to carefully introduce this type of technology into classrooms. To support this, systems need to be thinking about the kinds of structures needed to educate not only teachers and students through curriculum, but also parents and school communities about this. We need to deeply engage with the complexity around these emerging technologies, so that we get good outcomes for the profession, for our students that we care so much about, and for our communities more broadly. It's really about taking a more engaged and transparent approach. I'm hoping that we, as a profession, can band together to get that done.

### **What do you think might be next in AI for education? What excites you and what worries you the most?**

AI is so ubiquitous now and comes in so many forms: user-facing, back end and biometric. Some of it is really innocuous and cool. Some of it is very helpful, and some of it raises red flags for me.

I find it exciting to think about this in terms of the Australian Curriculum – of how we can develop an AI curriculum for Kindergarten to Year 12 and to



**It's important that our data doesn't become our destiny. ... Data shouldn't be used to categorise us in ways which can stigmatise, especially when there's no scientific basis behind the claims.**

higher education. The issue around big data, automated and intelligence systems, is here now. It's really time to deal with this in terms of curriculum development and pedagogy. It's exciting – and we should remember that children and young people will want to know how it works and how to take control over their interaction with this type of technology.

It's also exciting to think about how we can start to develop regulatory, legal and policy frameworks, which could be world leading. It's exciting to think about this, at a school systems level, a state and national level, and then at an international level. There is the General Data Protection Regulation (GDPR) on AI, which is a kind of European Union regulatory framework around AI, and there's a model there around data sovereignty – but how might we engage with that, so we have better data rights in this country? There are a lot of really interesting and exciting aspects to this.

In terms of worries, there's an issue around who gets access to emerging technologies and who doesn't. It

can be very exciting, interesting and pedagogically challenging and creative to use these types of technologies. So how can we use some of these technologies powered by AI, virtual or augmented reality for instance, for distance learning, and how can we use it more effectively? What's the potential there for that? How can we innovate in this space to reduce equity issues? Because if you only get access to the technologies because you're in a well-resourced school or your family can provide access to it, then that's problematic. COVID is highlighting the huge digital divide in this country. It's always been there, but COVID has put a spotlight on it. It's shown how much more we need to do to ensure equity of access and opportunity in the technology space.

It's also important that our data doesn't become our destiny, or come to be used as a sort of digital 'phrenology' as one of my colleagues has suggested. The family you're born into, the community you're born into, shouldn't be your destiny in this country. Everyone should have a fair go. We should have fair opportunity. When we use data, it shouldn't be used to determine what opportunities we have. Data about our gaze pattern or if we read test questions out loud shouldn't be harvested to try to determine whether we are cheating on an online exam any more than the bumps on our forehead should be used to classify people as criminals (as the debunked pseudo-science of



**Skills like ethical understanding can be hard to isolate, and harder still to measure. That's why great teaching is so essential to our safe passage into a digital future.**

phrenology did). Data shouldn't be used to categorise us in ways which can stigmatise, especially when there's no scientific basis behind the claims. Educational researchers and teachers have been very concerned and active around labelling learners and the effects that can have. When we're presented with an analytic dashboard which labels a learner, we should really question and understand what that label means and the unintended effects of that. I don't want to be viewed as a meat sack of data – that's against my right to bodily integrity which has been hard fought for. And yet, we have systems now who see us as these big meat sacks of data that can be harvested, broken down into pixels and micro-movements, to process and try to understand our intent and our emotional state. It's a complicated field and there are no easy answers. But what can we do about it as a profession? I think we need to build our knowledge, and really try to understand the science and evidence base behind it. We need a pedagogical project.

I also worry about discrimination. There's a very interesting UK paper that was published this year, by Wachter and colleagues (Wachter, Mittelstadt & Russell, 2020). They argue that we may now be entering a world where machines will discriminate in ways that are different from the way humans do, in ways that we can't predict. This becomes a problem for humans if we can't see or understand the ways machines discriminate, or if they're discriminating in ways that aren't conventionally understood by law. It becomes a particular problem if you're the human from a profiled group being discriminated against. If we can't collect data or information to create a prima facie case around discrimination, because the algorithm is proprietary or in a black box type of AI, then this is really worrying.

But that's not to say that we shouldn't try to harness emerging technologies to change, improve or support what we do. We need to have caution, ask questions, and have dialogue with different groups of people so that we can come to good solutions for the ethically tough questions that arise. We need good, transparent governmental systems. I'm often thought of as the 'grumpy cat' of AI when it comes to these matters – but we do need a bit of scepticism here! How do we use technology fairly? Will it be possible, and what are the human and technical approaches



to understanding this? We've got to ask the question – where's the evidence of effectiveness for learning? Does the technology have equitable consequences for students? These are the tough questions we need to ask.

As we argued in our report for the Australian Government, the important thing is that we very carefully incubate these technologies in place, with very strong ethical frameworks around them, and governance, oversight and accountability mechanisms. We need to ensure access to independent experts and expertise to avoid regulatory capture. And we all really need to understand what our human rights are in relation to technology. Students need to understand the digital rights of the child in this space. And we haven't had enough conversation about that in this country, yet – but there's no time like the present.

### **Thanks for your fascinating insights, Erica. Do you have any concluding thoughts?**

As someone who is researching cultures of ethics and AI in schooling in Australia and elsewhere – I'm very interested in knowing more about how teachers and computer scientists and policy makers are thinking about these big issues. If you're currently engaging with these issues – I would love to hear from you.

I would also just acknowledge the pedagogical project that lies ahead for education as a profession, the need to have a deeper understanding of AI, and the importance of including educators from diverse backgrounds in the conversations around the use of AI both inside and outside of the educational context.

I'm really excited by the opportunities AI can provide for teachers, students and entire communities. As long as we keep asking questions and having conversations, collectively, about issues of ethics and equity as we design, select and use AI-driven technology, I think we have a really exciting future to look forward to.

---

## References

- Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: learner choice and fundamental rights. *Learning, Media and Technology*, 45(3), 312-324. <https://doi.org/10.1080/17439884.2020.1786399>
- Southgate, E., Blackmore, K., Pieschl, S., Grimes, S., McGuire, J. & Smithers, K. (2018). *Artificial intelligence and emerging technologies (virtual, augmented and mixed reality) in schools: A research report*. Commissioned by the Australian Government Department of Education. Newcastle: University of Newcastle, Australia. <https://docs.education.gov.au/node/53008>
- Southgate, E., (2020). Artificial intelligence, machine learning and why educators need to skill up now', *Professional Voice Journal*, 13 (2), <https://www.aeuvic.asn.au/professional-voice-1322>
- Wachter, S., Mittelstadt, B. & Russell, C. (2020). *Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI*. <https://ssrn.com/abstract=3547922>

### Interested in reading more about Associate Professor Erica Southgate's research?

---

Visit Erica's website to find out more:  
[www.ericasouthgateonline.wordpress.com](http://www.ericasouthgateonline.wordpress.com)

If you are a teacher, school leader or policymaker who would like to be interviewed for Erica's research into the ethics of AI in education, you can connect with Erica via email: [Erica.Southgate@newcastle.edu.au](mailto:Erica.Southgate@newcastle.edu.au).





# We asked some NSW students:

## what do the words “Artificial Intelligence” mean to you?

To me ‘artificial intelligence’ conjures up images of robots, facial recognition - something that collects data and processes it. I associate it with corporations getting data on you, so they can market things to you – so a bit creepy ... very ‘1984’ – George Orwell vibes – it could become another form of surveillance – attempting to control people. But also useful – in that AI can process data faster than people can.



**Female student, 17 years old**

Artificial intelligence is intelligence in a computer. It’s a computer that can think for itself and make complicated choices – likes humans. Most people when they think of AI think of something like the Matrix, where computers have come to life.



**Female student, 11 years old**

A non-human intelligence that can perform a specified task, it can make decisions in a controlled environment, with controlled variables and settings. It can do something without a human making it do it, to some extent.



**Male student, 16 years old**

AI means the ability to step into the future. Robots that would assist in learning and work.



**Male student, 15 years old**

I don’t know what artificial intelligence means.



**Male student, 9 years old**







# Teaching privacy and ethical guardrails for the AI imperative in education

Evan Selinger and Amelia Vance





In 1956 computer scientist John McCarthy coined the phrase ‘artificial intelligence’ (AI) to describe ‘the science and engineering of making intelligent machines’ (McCarthy, 2007). Over time, the term has evolved to cover a variety of technologies, including ones widely used in education, from plagiarism detectors to voice-activated virtual assistants leveraged to enhance campus information distribution and classroom pedagogy (Arizona State University, 2018).

Contemporary AI discussions are about ‘a variety of methods and tools which mimic cognitive functions across three areas: 1) perception/vision, 2) speech/language, and 3) learning analysis’ (Family Online Safety Institute, 2019). Experts further distinguish between ‘narrow AI’ and ‘artificial general intelligence’. Narrow AI performs well in discrete tasks – think of algorithms that are good at playing chess but wouldn’t help someone drive a car, and vice versa – and is currently having a big impact

on society. By contrast, artificial general intelligence can learn across domains and think for itself, possibly in astonishingly more sophisticated ways than humans do. For now, artificial general intelligence is the stuff of science fiction and unresolved debates rage over whether it will ever be created.

Presently, the advancement of AI is associated with machine learning and its subset deep learning. Machine learning involves algorithms using

statistics to computationally identify patterns in data sets that are often so large they are referred to as big data (Emerging technology from the arXiv, 2013). Big data is the foundation on which modern AI is built. This vast quantity of data gives scientists and statisticians more opportunities to create AI that can identify increasingly intricate and nuanced patterns. Since these patterns are frequently about how humans think and act, they can be used to make sense of present human behaviour. For example, AI can analyse student facial expressions to infer whether students are bored or engaged while doing an assignment. AI can also be used to more accurately predict future human behaviour. For example, it might be used to estimate which students will be academically successful and which candidates will perform best at a job by scanning for certain keywords in a resume that match or are similar to the resumes of current successful accepted students or employees when they were initially hired. Without big data, these AI applications would not exist. In this article, we use the term AI to refer not only to the definition provided in the previous paragraph, but also to the infrastructure that is used to create the AI, which includes big data, statistical modelling, and machine learning.

Because this article focuses on privacy and ethics issues in K-12 education, we won't get into the specific details of technical discussions about AI. Instead, our goal is to identify

**To put it mildly, high hopes exist for AI ... Experts believe AI will significantly impact how teachers instruct, how students engage with learning, and how administrators set priorities and select policies.**

practical educational uses of tools and systems associated with AI and big data. We will offer clear principles for responsibly using AI and its related big data in educational settings, and make recommendations for how instructors can effectively teach students about the ethical and privacy risks associated with AI and big data in and beyond the educational context.

Although we won't review any products, from the start we would caution educators and administrators to use care when selecting technologies that are marketed as powered by AI. As discussed earlier, the term AI can encompass a variety of technologies and be interpreted in many ways. Journalistic coverage of AI products and services has been riddled with misinformation. Consequently, companies have enough wiggle room to manipulatively peddle their wares and abuse the term to create exaggerated impressions of how cutting-edge their products are (Marcus, 2019).

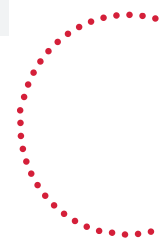




## How can educators meet the AI imperative?

To put it mildly, high hopes exist for AI. Beyond widely-covered victories of AI over human champions in games requiring immense knowledge, skill, and creativity, such as IBM Watson in Jeopardy!, Google AlphaGo in Go, and AlphaZero AI in chess, AI is being adopted in a range of domains, from medical to military applications to the educational sector. One AI-powered educational testing system is so advanced that it scored over 90% on the multiple-choice questions found on the eighth grade standardised science exam administered in New York State (Clark et al., 2019).

Experts believe AI will significantly impact how teachers instruct, how students engage with learning, and how administrators set priorities and select policies. Consider the following thought experiment about an advanced AI tutor. The scenario has a few distinctive features: it takes place in the future where technological capabilities are deployed that don't yet exist, it envisions university students learning in virtual classrooms and, crucially, it represents what some in the education technology sector consider a positive step forward.



“The AI tutor will design personalised learning plans that optimise each student's outcome. Should one student watch their lecture at breakfast time, or in the evening? Where should their first test pop up in a busy schedule? How much preparation will they need to understand a certain concept? While a skeleton crew of humans would be needed initially to design curriculums (the creative bit) and film lectures (CGI is still too expensive), AI tutors could do the rest” (Haw, 2019).

Regardless of debates about whether the scenario described above is desirable or advantageous for students, the most important pedagogical tasks related to AI don't simply keep humans in the loop, but at the centre of the action. Educators must figure out how to best ensure students are prepared to succeed in a world where AI is deeply embedded in infrastructure and its use influences what jobs are available, what civic life is like, how markets function, and how people communicate and socialise. Simply put, 21st century education requires preparing students for an AI-intensive future. Educators in primary and secondary schools need




to create new lessons and experiment with new educational technologies that can help them meet the challenge of the AI imperative.

The AI imperative is the requirement that educators adequately understand what AI is and the key privacy and ethical issues surrounding its use, effectively teach grade-appropriate versions of that knowledge, and put AI tools to good use in and around the classroom. Failure to meet the AI imperative will leave students unprepared for the increasingly AI-reliant real world. With the stakes so high, the pressure to meet the AI imperative eventually will come from everywhere: from parents, politicians, employers, technology companies, and even the students themselves.

What does it take to meet the AI imperative? Clearly, technical proficiency is essential. For example, educators will need to teach students

what algorithms are and gain experience creating programs that involve algorithmic problem-solving and computational thinking. However, this is not enough: without due consideration of privacy and ethics, world-class computer science lessons on machine learning and other forms of AI will fall short. Technology companies themselves acknowledge that this is the case by publishing company-backed 'AI principles' that outline vision statements infused with ethical values and commitments (Future of Privacy Forum, n. d.). Research institutions and public sector organisations also recognise the importance of privacy and ethics by proposing principles and guidelines for ethical AI applications (Jobin et al., 2019).

These formal initiatives are also driven by headline-grabbing controversies that have raised the public's consciousness about the importance



of algorithmic accountability. For example, the Australian Online Compliance Intervention program, colloquially referred to as Robodebt, found inconsistencies between Australian Taxation Office income and welfare payments. The program was deployed to generate debt notices and secure 'up to \$4.5 million in Centrelink debt a day' (Dean, 2019). After, after 'tens of thousands of welfare recipients' claimed they were 'overcharged for alleged debts,' the federal government decided to change the program. As initially implemented, the program left the burden on wrongly targeted recipients to prove they were treated unfairly (Karp, 2019).

Now, human officers are no longer permitted to exclusively rely on the automated system and are required to conduct 'further investigations to determine whether a debt exists' (Marsh, 2019). Mandating human oversight over important decisions is a recurring concern expressed in debates over AI accountability.

Without proper human oversight, education-based AI initiatives can also result in privacy and ethical violations. Respecting student privacy should be a critical component of any effort to meet the AI imperative. Protecting student privacy is not only the right thing to do – it can be legally required. And, because promising programs can be destroyed by poor privacy optics, respecting privacy is instrumentally valuable. If schools or schooling systems communicate

**Clearly, technical proficiency is essential. However ... without due consideration of privacy and ethics, world-class computer science lessons on machine learning and other forms of AI will fall short.**

poorly with parents about the privacy precautions they are taking and the privacy policies they are following, they may shoulder some of the responsibility if adverse outcomes follow. Opportunities that are genuinely good for students can be misconstrued as dangerous and get shut down if judged irresponsible.

Meeting the AI imperative also requires educators to teach students about the ethical risks of big data and AI systems. Without this knowledge, students won't be prepared to understand and navigate an increasingly AI-centric world. Quality approaches to teaching ethical risks, including privacy pitfalls, require educators to practice what they preach. If students and parents believe that educators are not modelling good behaviour, their authority could be undermined.

A key lesson that educators should stress is that the most basic way AI will change society is through the choices people make about which AI systems to adopt and reject, and how to wisely use the ones that are selected.



By acting as if the AI imperative requires fatalistically acquiescing to a contingently popular technological trend, people neglect their agency and abdicate basic responsibility.

It is simple for educators to rationalise away accountability by acting as if a technologically determined pathway exists, mandating how students must learn. For example, a group of students expressed discomfort about being required to use a classroom engagement detection system that scans students' faces and uses AI to infer whether they are engaging with material in an attentive manner. When asked to comment on this, the CEO of the company that produced the system bluntly replied that the students' concerns would not change anything; channelling Darwin, he noted "[e]verybody is doing this. It's really early and shocking, but we cannot go against natural laws of evolution" (Eveleth, 2018).

What a striking response! Starting with the age-old question, "Would you jump off a bridge just because everyone else is doing it?", we teach young children that something is not necessarily right just because other people are doing it. But adults may fail to heed their own advice when the AI imperative seems to only provide two options: accommodate and adapt all potential applications of AI or perish and give up hope of being relevant.


Technological trends are not natural laws that force individuals and institutions to go in any particular direction. When people, including

students, articulate justified ethical reservations that push back against popular positions, seriously considering their concerns can help us forge a better future. Educators should not lose sight of this fact by acting as if they are powerless to contribute to decisions about how to go about meeting the AI imperative.

In the case of the automated engagement detection systems, it is wrong to automatically dismiss student concerns. Without seriously considering the merit of their criticisms, it is impossible to identify students' values, whether the values deserve the attention of educators and administrators, and whether greater emphasis on these values can motivate the educators using these systems to meet their pedagogical goals through better alternatives. Students who are dismayed about AI surveillance and analysis in the classroom are often mirroring broader concerns in societal controversies about policing and workplace dynamics. Giving concerned students the impression that advocating for privacy is passé and that authorities should be given unrestricted power over their lives diminishes student agency and threatens their autonomy.

### **Ingrained privacy and ethics guardrails**

We've already provided some of the reasons why student privacy should be respected. But let's think about this issue on a more fundamental level. Clearly, students aren't entitled



to absolute privacy. For safety reasons, teachers and administrators might be justified in searching a locker or a student's bag for drugs or responding to threats made over social media. If disciplinary problems arise in schools, educators can be obligated to contact parents rather than keep the information confidential. Nevertheless, there are many reasons why students deserve to have some privacy interests protected.

Students are a vulnerable population. Consequently, lack of adequate privacy safeguards can lead to students experiencing embarrassment and shame, even trauma, and suffering lasting reputational damage that can unfairly damage or diminish their future opportunities. Insufficient privacy protections also can lead to exploitation; that's why many believe that technology companies providing educational services should be prevented from benefitting in any way from student data, even from using it to improve their products. Inadequate privacy measures can also have a chilling effect, inhibiting student self-expression. This harm can impede learning because students who are anxious about privacy issues in schools may have difficulty focusing on scholastic tasks, like lectures and exams. In turn, this can compromise their grades and intellectual maturation. Furthermore, the chilling effect can compromise students' ability to socialise. Socialisation is critical to the healthy development of students that schools are supposed to foster. Since schools

**Ideally, schools will provide students with tools for having thoughtful conversations and making thoughtful decisions about privacy in both analogue and digital spaces with their parents, teachers and peers.**


are responsible for preparing students to become citizens, schools should strive to empower them to think critically about when it is appropriate or inappropriate for authorities, like government officials, employers, and technology companies to engage in privacy diminishing activities like surveillance. If students and parents believe that schools are not adequately protecting their privacy, they might intervene in ways that compromise, if not shut down entirely, promising educational programs.

Since schools should teach students how to grow into responsible, autonomous adults, they should avoid smothering forms of monitoring that lead students to abdicate responsibility for their own behaviour as well as avoid forms of surveillance that will inhibit students from full participation in educational activities. Ideally, schools will provide students with tools for having thoughtful conversations and making thoughtful decisions about privacy in both analogue and digital spaces with their parents, teachers and peers.

For the reasons just provided, it is difficult to neatly distinguish between privacy and ethics. Many privacy protections are rooted in ethical ideals, like respecting autonomy and fostering personal and social development. A useful way to think about the relationship between privacy and ethics is to revisit what happened when data digitisation began in the 1970s. At that time, government officials and policymakers expressed concern, often ethically-minded, about the implications the innovation would have on privacy. Worldwide, committees and commissions were formed to determine the rules for public and private entities that collect, process and share computerised data. These efforts culminated in the Organisation of Economic Cooperation and Development (OECD) releasing the 'OECD Guidelines on the Protection of Privacy and Transborder Flows of Person Data' in 1980, drafted by a group chaired by the then-Chairman of the Australian Law Reform Commission (Gellman, 2019). The guidelines included eight core ideas (paraphrased below) that now lie at the foundation of many privacy laws and proposals around the world. Australia adopted them in 1984, and today these established principles can help provide privacy and ethical guardrails for K-12 institutions to responsibly adopt AI (Greenleaf et al., 2013).

1. Information collection should be limited. It should be collected in a fair way, ideally with the knowledge or consent of the person whose data it is.
2. Information needs to be accurate, complete, and up-to-date.
3. Information should only be collected for a specific purpose that is clearly specified before collection begins. The information should not be reused later in ways that are incompatible with the original specific purpose.
4. Information should not be shared or reused in ways incompatible with the original specific purpose without either consent of the person whose data it is or a law providing the authority to do so.
5. Information should be protected through reasonable security safeguards that limit the risk of an unauthorised person accessing, using, changing, sharing, or destroying that information.
6. There should be transparency about how information is collected, used, shared and protected.
7. People have certain rights about their own information, including the right to know what information others have about them, who has it, and the right to request that information be corrected, amended or erased.
8. Those holding data must be held accountable for the above principles.





**Wisdom is based on tacit knowledge. To understand how to act wisely in a particular situation, one must have tacit knowledge about that situation and the context in which it has evolved.**

To see how these principles matter in practice, we offer two case studies where privacy and ethical missteps compromised potentially positive advancements in data sharing and educational technology.

### **Case Study 1 – inBloom**

In Australia, inBloom might not be a well-known initiative. However, its rise and fall fundamentally shaped the educational technology landscape in the United States. Even though it is not a cautionary tale about the use of AI, the manner in which the debacle unfolded holds significant lessons for educators everywhere who want to responsibly use and communicate about new educational technologies, particularly in contexts where big data is involved. Indeed, cases like inBloom are especially important to consider now in the early days of using AI in education. Some of the clearest lessons about privacy and ethics can be learned by critically analysing predecessor cases that raise similar issues.

InBloom was a \$100 million non-profit educational technology initiative that ran from 2013 to 2014 and aimed to improve American schools by providing a centralised platform for data sharing, learning applications and curricula (Bulger et al., 2017). At the time, many education reformers and student advocates argued that education data should be democratised, and that students should be able to ‘carry’ their data with them from grade-to-grade and post-graduation in a ‘digital backpack’. Despite lofty ambitions, scholastic digital data systems adopted in the 1990s and 2000s were clunky (Denver Post, 2016). Sometimes, data was not even transferable between schools in the same state. Moreover, the systems were often built around minimally useful data points that were used to evaluate students and educators at a single moment in time, instead of continually throughout the year, such as grades and test scores rather than digital portfolios.


The data inBloom could store and the technical architecture of the software were portrayed as game-changers. InBloom was developed to provide secure, semi-instantaneous access to student information, able to facilitate personalised learning, identify learning issues before they manifested in failing grades, provide students and their parents easy access to their own continually updated data, and, overall, help create a more data-driven educational system. Many of these objectives underlie the current push for schools to take advantage of big data and AI.

Soon after inBloom launched, it secured contracts with several states and school districts (Bulger et al., 2017). For many parents, inBloom's launch was the first time they heard about how data collection and use in the educational sector had changed since the passage of the major U.S. education law, No Child Left Behind (NCLB). NCLB required schools to collect more data to ensure that all students were being served equally (H. R. 1, 2001). While laudable in principle, the reality was far more complex. A 2009 study conducted by Fordham Law School's Center on Law and Information Policy found that schools often collected more data than necessary to comply with NCLB and evaluate a school's overall progress (Reidenberg & Debelak, 2009). In fact, the study noted that state databases included students' detailed academic, disciplinary, health, and family information, 'often in a non-anonymous student record' (Reidenberg & Debelak, 2009). For example, California kept records of students' parental educational levels, and Iowa and Illinois included data points for students' criminal histories. With this mismatch between aspiration and behaviour in the background, parents feared that schools adopting inBloom would end up creating a digital permanent record that tracks, ranks, and categorises students for the rest of their lives (Study session regarding InBloom Inc, 2013). While such a threat was, in itself, ominous, the anxiety was further exacerbated by concern about due

process. Parents lacked a clear process for correcting, amending, and erasing incorrect and misleading information contained in the record.

InBloom's own messaging did little to dissuade fears. Its website listed pages and pages of possible data fields that districts could choose to use (Bulger et al., 2017). While inBloom employees probably perceived this to be a neutral configuration – after all, they were simply listing the fields available without telling districts what specific data should be collected – privacy advocates worried that districts would use all the available fields and collect excessive amounts of data. The Electronic Privacy Information Center (EPIC) raised concerns that inBloom advertised that it allowed for the creation of 'principal watch lists' with no description of why a student might be added to the list. EPIC also criticised inBloom for allowing children to be labelled if they violated 'norms of behaviour' (an undefined classification) and categorised as a 'perpetrator' or 'accomplice' for disciplinary incidents, terms that would normally be used in criminal, not school matters (Study session regarding InBloom Inc, 2013).

Ultimately, inBloom's leadership erred in their faith that the platform would be seen as sufficiently neutral and that all of the privacy and ethical responsibilities would be transferred to the administrators and teachers at the schools that adopted it. InBloom's failure to incorporate core Fair Information Principles (another




term for the principles we introduced on page 38, like specifying a clear purpose for any data collection and committing to only collect the minimum amount of data needed for that purpose, likely exacerbated privacy concerns.

Public perception was further damaged by inBloom's privacy policy stating that it could not 'guarantee the security of the information stored in inBloom or that the information [would] not be intercepted when it [was] being transmitted' (Ash, 2013). Furthermore, as was the norm for many technology companies, inBloom had a data privacy and security policy that they could unilaterally modify "from time to time, as approved by [an] independent advisory board" (Study session regarding InBloom Inc, 2013). This meant that whatever protections their policy did offer could not be guaranteed; instead, they could change at any time. While these disclaimers and disclosures are standard practice in corporate contracts, this was likely the first time parents had seen these types of clauses. Parents expected that their children's information would be reasonably protected and expressed unease when it appeared inBloom was dodging responsibility (Ash, 2013).

Finally, and perhaps most importantly, representatives of inBloom did not communicate with parents and advocates effectively (Solove, 2018). An article in 'Politico' noted that many educational technology entrepreneurs and school reformers were

'bewildered by and anxious about the backlash' against the use of data in education. They had assumed parents would automatically support mining 'vast quantities of data for insights into what's working, and what's not, for individual students and for the education system as a whole,' and that the benefits were 'self-evident' (Simon, 2014). Indeed, the lack of clear answers given to parents and advocates when they raised questions only added to public mistrust and anxiety. Without clear information from inBloom about the platform's privacy guardrails, parents and advocacy groups narrowly focused on possible abuses of the system.

InBloom lacked a communication plan for addressing the concerns of parents and privacy advocates. It expected school administrators and educators to do the heavy lifting on informing parents about the advantages of the platform. Such overconfidence kept inBloom from providing any communications resources to their education partners. Indeed, they didn't even inform schools that proactive communication would be necessary. This lack of foresight left schools unprepared to explain how, exactly, the new technology would be beneficial, such as helping parents understand what was happening in their child's classroom and making it exponentially easier to transfer records when their child transferred schools. Without plausible and positive depictions of how inBloom could responsibly be used, parents





had no reason to trust that adopting it would be beneficial. Instead, they had plenty of reasons to be concerned that it would be abused and result in too much information about students being collected. And so, with the bad impressions making a larger impact than the positive ones, every state and district cancelled their contracts with inBloom within one year of its highly anticipated launch (Madda, 2014).

Like inBloom, initiatives around AI in education are often premised on assumptions that ‘everyone’ supposedly agrees with: everyone agrees that data should be easier to share and analyse; everyone agrees that mining data for insights that could help students succeed is worthwhile; everyone wants students to be able to carry their data with them through grades and across districts; everyone wants to take advantage of new technologies to serve students. But if schools or school systems cannot clearly explain how data will be collected, used, shared, and protected, and cannot explain the benefits of initiatives that are based on new data collection or sharing, it will not matter how much that initiative could help students, how much the initiative solves a problem, or whether the initiative has privacy and security built-in. When there is not enough information about an endeavour and privacy and ethics are on the line, people are likely to assume the worst.

## Case Study 2 – Mount St. Mary’s

As with the previous case study, this one does not involve the use of AI. Nevertheless, it’s a remarkable cautionary tale for educators considering adopting AI. This is because the fiasco that occurred at Mount St. Mary’s University revolves around predictive analytics, a use of big data that routinely applies AI. Moreover, the problem with how predictive analytics were used at Mount St. Mary’s gets to the heart of a fundamental issue in education itself – equity. The case of Mount St. Mary’s University shines a spotlight on how privacy and ethical guidelines are not only necessary for managing optics, but also for ensuring that the underlying goal of helping each and every student lies at the foundation of all educational initiatives.

Most educators would like to believe that the fundamental goal of education – and associated jobs – is to help every student succeed. It is the guiding ethos of the teaching profession and the true purpose of the work that many view as a vocational calling. However, in the real world, idealism and pragmatism regularly clash. Teaching a diverse student body inevitably requires making compromises and accepting trade-offs. For example, the wellbeing of the majority of students – the students who use the fewest resources and need the fewest interventions – may be prioritised



over students with disabilities and students of lower socioeconomic status, who may need more resources and attention. Further complications arise when schools risk losing funding due to lacklustre test scores, low graduation rates and high levels of disciplinary incidents.

While a strictly utilitarian view might justify schools prioritising the interests of the majority instead of trying to apply limited resources equally to everyone, most educators would be repulsed by the prospect of giving up entirely on some children in order to serve the good of the many. If society transitions towards an increasingly data-driven educational model that focuses on maximising efficiency and places increasing priority on AI services, the tension between what is aspirational and what is possible will become more evident.

Mount St. Mary's University is the second-oldest Catholic university in the United States. In 2014, Simon Newman, a private equity CEO and entrepreneur (Svrluga, 2016), was hired as Mount St. Mary's new president. His goal was to "raise a lot of capital and ... start the university on a more aggressive growth trajectory" (Bowie, 2014). With the school board's approval, he planned to increase the university's first-to-second year retention rates. Increasing retention rates is significant because the 'U.S. News and World Report', an influential publication that ranks colleges and universities, weighs 'average freshman retention rate' as one of its highest evaluative categories (Morse, 2013). Schools ranked highly are perceived as prestigious and, as a result, attract quality applicants. Furthermore, alumni can leverage the university's high profile and ranking when looking for jobs.

During the fall of 2015, Mount St. Mary's freshmen were asked to take a survey that the president's office stated would help the school 'develop better advanced metrics for accepting students' (Mount President's attempt, 2016). The survey's introduction described it as a 'very valuable tool that will help [students] discover more about themselves' (Svrluga, 2019). The survey also purported to be 'based on some of the leading thinking in the area of personal motivation and key factors that determine motivation, success, and happiness' (Mount President's Attempt, 2016). Crucially, students were told that 'there are no wrong answers' (Svrluga, 2019).

Some questions were innocuous, asking whether students might be interested in certain majors if the school added them. Others were extremely sensitive, asking whether the student had a learning disability, whether a close friend or family member had died in the past year, or whether the student felt they 'could not shake off the blues, even with help from ... family and friends' (Svulga, 2019). Unbeknownst to the students and most faculty members, Newman planned to use the survey results to inflate the school's retention rates by asking select students to leave the university (Schisler & Golden, 2016). The administration would contact some of the students whose survey results labelled them at risk of leaving the university and offer them a tuition refund if they dropped out before the federal reporting deadline for the school's first-year enrolment. When some faculty protested, Newman

stated, "[t]his is hard for you because you think of the students as cuddly bunnies, but you can't. You just have to drown the bunnies ... put a Glock to their heads" (Schisler & Golden, 2016).


Despite Newman's confidence, his plan failed. The faculty refused to identify the quota of twenty to twenty five students to talk into dropping out, the student newspaper reported on the scheme, and, eventually, enough outrage resulted that Newman resigned (Johnson, 2017).

Newman argued that he was doing the right thing for both the college and students who were likely to fail. According to an op-ed he wrote in 'The Washington Post':

'[r]ather than continuing to collect exorbitant sums of money from the families of continually failing students, the kindest, most responsible option for institutions may be to return students their paid-in tuition and offer guidance for other paths [such as] other universities or technical schooling . . . many students aren't always willing to raise their hand and say, "I need help." So, it is our obligation to identify warning signs that can appear as early as a student's first semester that the academics and college life is not the right fit' (Svrluga, 2017).

But a sociological analysis of this case found that the survey questions lacked an 'established connection to educational settings' and, further, lacked 'evidence of any effort by the university to test validity or reliability' to 'measure of the likelihood of





academic success' (Johnson, 2017). The questions focused largely on 'non-cognitive characteristics of students', which included sections on 'resilience and grit, personality inventories...religious beliefs' and even included a section for evaluating students for clinical depression (Johnson, 2017).

Another problem with Newman's calculus – beyond the lack of scientific rigour underlying the survey and prior statements where he had expressed anything but sympathy for struggling students – is that the underlying assumptions were not predicated upon universally shared assumptions. While his choice was to push students who seem to have a higher likelihood of failure out of the university, there were other choices, including providing students with more support to prevent them from failing. One first-generation Mount St. Mary's student interviewed after Newman resigned described her first year as "a struggle for me, it was traumatic ... [e]very day I would call my mom, sobbing, 'I don't think I can do this.'" Fortunately, with a faculty advisor's help, she persisted, and was a successful third-year student at the time of the interview. "If somebody would have told me when I was a freshman, 'this is not a good fit for you, you should probably go,' I would have gone" (Scott, 2019).

No matter how much information a school has about students, and no matter how smart a human or machine's analysis of it becomes, it

**Students are the largest stakeholders in the context of student privacy conversations. They deserve complete honesty when it comes to why their data is being collected and how it will be used.**

is impossible to predict with 100% certainty who will decide to drop out of college. As society aims to create more educational opportunities for lower income and minority students, we should not lose sight of the fact that, since fewer students like them have historically succeeded, predictive algorithms trained on historical data will label these students less likely to succeed here and now. Imagine how such analytics would have rated the chances of women graduating from medical school when universities first began admitting women.

Beyond the potential for discrimination, automated decision-making processes like predictive analytics are not neutral. They are imbued with the values of their creators. In the case of Mount St. Mary's, the administration's values were not aligned with faculty or students. The survey's wording led students to believe there were no wrong answers, encouraging them to answer truthfully and giving them the impression that their answers would not be used against them.




Consequently, students who took the survey consented to the use of their personal information for the survey's stated purpose – improving the school and student experience – but not its actual purpose – culling the student body to meet the president's goals for the university. Not only does this case study illustrate what happens when an administration fails to be transparent with data collection and analysis, but it also illustrates the importance of fairness. Students are the largest stakeholders in the context of student privacy conversations. They deserve complete honesty when it comes to why their data is being collected and how it will be used.

Imagine another universe where Mount St. Mary's created the exact same survey, but instead of pressuring students to leave, they presented them with opportunities for more

guidance counselling, tutoring, or a re-ordered class schedule designed to improve student success. If the administration had adhered to privacy and ethical guardrails – such as maintaining a primary focus on the best interests of each student, data minimisation, equity, and transparency – it is unlikely that this initiative would have become a case study in how predictive analytics could harm students.

### Teaching AI and privacy ethics

In the previous sections, we highlighted many interrelated ethical and privacy risks associated with big data and AI. We'll expand the discussion here and offer suggestions for how instructors can effectively teach some of the topics.



As with all subjects, teachers should design age-appropriate lessons. Until recently, it was nearly impossible to find educational material on AI ethics that weren't created for university students or professionals. Fortunately, things are starting to change. Blakely Payne, a graduate student at the Personal Robotics Group at the MIT Media Lab, created an important resource that instructors at various levels can use and modify. It's an open-source, module-based curriculum for middle school students on AI that can be integrated into science and humanities classes. Access the material at [MIT Media Lab](#).

### Engaging lessons that combine technical and ethical content

Payne's curriculum is important for three reasons. First, it covers both technical and ethical concepts. Second, it provides examples that students will find resonant. Third, it encourages students to view AI systems as changeable.

By combining technical and ethical education, Payne's approach provides a model for instructors at all levels to emulate. Including technical information is valuable because it helps students develop the technical literacy needed to understand AI. For example, students begin by learning the basics – understanding what algorithms are. Payne recognises that a useful entry point is for students to view algorithms as recipes containing instructions for generating outputs after acting upon inputs in designated ways. To make this point in an

engaging manner, Payne includes an exercise for kids to brainstorm which instructions an algorithm should follow to make the 'best' peanut butter and jelly sandwich.

Since the sandwich only has a few ingredients, students might presume it's a basic task. However, in short order, they'll realise it's actually a surprisingly complicated endeavour. Moreover, with help from instructors, students can begin to appreciate that the complexity in this seemingly mundane case actually speaks to deep issues that lie at the heart of many consequential applications of AI.

Making the best peanut butter and jelly sandwich isn't an endeavour that can be reduced to automating a basic process because what 'best' means is undefined. Is it an aesthetic ideal, guiding the algorithm to make the best tasting or the best-looking sandwich? Or should 'best' be understood as an ideal about safety, guiding the algorithm to make a sandwich that tastes like peanut butter and jelly, but that kids with peanut allergies can eat without getting sick? By helping students appreciate that the definition of concepts like 'best' are rooted in ethical values with diverse interpretations, they can be guided towards the eureka moment of grasping that programming algorithms inherently comes with the risk of, to continue with the culinary metaphor, baking in questionable preferences and unfair biases. Becoming aware of this risk – which fundamentally links



technical and ethical components of AI – is a teachable skill that enhances the broader skillset of ethical attunement, which educators at every grade level should aspire to develop in their students.


Instructors can cover more advanced material by modifying this engaging approach combining technical and ethical lessons. For example, consider the ethical issue of fairness in the context of AI. As we have already noted, the ideal of fairness is mentioned in the OECD privacy guidelines. However, the concept of fairness, like the concept of best, has additional definitions that correspond to a range of ideals. Fairness encompasses everything from equality of opportunity in situations where candidates are applying for jobs and AI plays a role in determining who is and isn't hired, to equality in outcome in cases where AI helps determine who gets pulled over for questioning during airport security checks (Binns, 2018).

In an ambitious study of prominent AI ethics guidelines developed around the world (including a discussion paper from the Australian Government Department of Industry, Science, Energy and Resources) Anna Jobin, Marcello Lenca, and Effy Vayena discovered a 'global convergence' around five ethical principles: 'transparency, justice, and fairness, non-maleficence, responsibility and privacy' (Jobin et al., 2019). (Note: The fact that privacy is listed as one of the most fundamental AI ethics issues

corroborates the approach we have been taking here of treating ethics and privacy as deeply connected.) The authors succinctly summarise key findings of an extensive literature review that traverses public and private sectors, carefully explain what the most important AI ethics concepts are and why they matter, and identify some of the most important commonalities and differences in how ethical principles are interpreted.

Let's bring the discussion of fairness back to teachable examples. The controversy surrounding the use of the U.S. Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system is an example illustrating the difficulty of translating diverse ethical interpretations of fairness into mathematical models. COMPAS, developed by the private company Northpointe, is a recidivism risk assessment system that predicts the likelihood a defendant will re-offend in the future. U.S. judges use COMPAS as an aid in conjunction with other information when deciding whether or not a defendant should be confined to jail or released on bail until the criminal trial takes place. The hope was that the system and others like it could improve justice by helping generate fairer predictions.

COMPAS generates its results from a questionnaire on the defendant's criminal history and attitudes about crime, explicitly not including race as a variable for analysis. However, investigative reporting by ProPublica



found that COMPAS was twice as likely to erroneously flag black defendants as being high-risk for committing future crimes, with the opposite holding for white defendants, who were twice as likely to be flagged as low-risk (Angwin et al., 2016). The higher risk scores generated by COMPAS do, in fact, correlate with the likelihood that someone will be rearrested, and is nearly identical across racial lines – a result that, taken on its own and without regard for broader context, appears fair because predictions about future criminal offences should apply equally to everyone in a risk category regardless of their race. So, why is it that COMPAS is harsher on black defendants? It appears to be because in the U.S. black defendants have historically been rearrested at higher rates than white ones, and proxies for race, such as prior arrests, can lead COMPAS to classify black defendants as higher risks than white defendants (Corbett-Davies et al., 2016). Although COMPAS was designed to analyse risk without racial bias, historical prejudice nevertheless contributed to biased predictions.

It can be more difficult for students to grasp the idea that different, sometimes conflicting, conceptions of fairness exist than appreciating that there are different ways to make the best peanut butter and jelly sandwich. The danger of algorithms using data tainted by histories of injustice, such as de-contextualised statistics that make minorities seem

like criminals or poor job candidates, makes issues surrounding AI and fairness especially timely. Modelling educational exercises that follow Payne’s engaging model of combining technical and ethical lessons will help educators effectively communicate both concepts to students. For example, Karen Hao and Jonathan Stray created ‘Can you make AI fairer than a judge?’ for the ‘MIT Technology Review’. It is an interactive online article that explains the main issues in the COMPAS example in an engaging way, encouraging readers to play a courtroom algorithmic game where they explore issues like trying to make the fairest possible risk-prediction algorithm (Hao & Stray, 2019).

### Lessons that focus on examples students can relate to

Payne is right to realise that K-12 educators should primarily focus on examples that will resonate with students based on their personal experiences and professional aspirations. Most conversations about AI focus on potential existential risks set in the distant future with technology that does not yet and may never exist. Payne focuses on relatable examples because she recognises that children are growing up in a digital world increasingly mediated by AI. Algorithms suggest what they should watch, read and listen to. Since kids are a vulnerable demographic, and the technologies they use in and outside of the classroom will shape everything from what they know to

who they are and aspire to be, it's absolutely critical they are empowered to use them wisely and be critical consumers. To this end, Payne includes a module on redesigning YouTube, a platform recently criticised for nudging kids towards dangerous, conspiratorial content (Roose, 2019). Her lesson plan has students identify the various stakeholders that YouTube impacts and construct an ethical matrix that highlights where their values overlap and conflict. This exercise is ideal for learning how prioritising different stakeholders with different preferences and values can lead to designing recommendation algorithms that perform in different ways. Crucially, this analytic and pedagogical approach can be usefully applied and adapted to cover a range of AI systems that students should be expected to care about.

We expect students will be interested in how AI is used in schools since these applications can directly impact their lives and futures. This makes examples like the Mount St. Mary's case especially relevant. The scenario can be adapted to K-12 schools considering using AI to predict who will and won't benefit from participating in a special program. Students could write an essay that imagines the school administrators proposing the program to an AI ethics board. Students could pretend to be ethics board members, and

writing from this perspective, they could determine who the relevant stakeholders are and what values they embody. Specifically, they should decide to approve or reject the proposal based upon considerations related to them, as well as the OECD principles and AI ethics principles that Jobin, Lenca, and Vayena cover. An important issue for students to grapple with in this context is what definition of fairness is implicated. We already alluded to it in our previous discussion – it's the problem of statistical discrimination that fails to treat people as individual decision-makers who might or might not behave like other members of the groups they are identified as belonging to (Binns, 2018).

Similarly, we also expect students will be interested in topics that concern how AI will be used to impact the lives of their parents and other family members. Thus, teachers might want to design assignments that have students think carefully about examples like the Australian controversy surrounding Robodebt that we discussed earlier. Students could reflect on how they would feel if their parents were incorrectly flagged by such a system, and how they would design a better system to prevent the injustice from occurring.





## Training tomorrow's responsible leaders today

Since today's children are tomorrow's leaders, it's important to expose them to the profound ethical and privacy consequences of AI as early as possible. This way, as they get older and start working on projects involving AI, they'll hopefully be inclined to care about ethical issues from the very start of a project. This ethically attuned outlook is preferable to seeing ethics as an afterthought. Making ethically important changes after a project is completed can be burdened by undue expenses or difficulties that could have been avoided through better planning. The ethically attuned outlook is also preferable to passing responsibility for addressing ethical concerns to other bodies, like regulators. Forgoing this responsibility implicitly rests on the outdated and fallacious presumption that scientists and engineers are not at all responsible for how their devices, systems and ideas are applied (Green, 2018).

We believe it is especially important, as Payne declares that “the ultimate goal is to enable students to see artificial intelligence as manipulatable – from a technical and societal standpoint – and to empower students with tools to design AI with ethics in mind.” Pushing this point further, we would like to suggest that students be given the opportunity to

consider cases where the most ethical outcome might be to prevent AI from being used in the first place.

Consider the example we previously discussed of students objecting to being required to learn by using a face scanning engagement system. As mentioned, the CEO of a company that makes one of these systems dismissed their concerns as irrelevant because they are out of step with a presumably inevitable technological evolution. In the classroom, students could be given an opportunity to articulate precise objections to an AI-informed product, like the engagement scanner, and brainstorm ways that the product could be changed or policies could be put in place to promote quality education while mitigating against student concerns. Students should also be given the chance to explain if and why changes or policies won't suffice and how educational goals could be better met without a particular form of AI or possibly without using AI at all.

By helping students consider the full range of possibilities, educators can truly prepare students to meet the AI imperative – an imperative that does not require solving every possible problem with AI.

### Acknowledgements

The authors would like to thank the following individuals for their help and support with this article: Sara Collins, Brenda Leong, Jasmine Park, Anisha Reddy, Alexis Shore, and Katherine Sledge.

## References

- Angwin, J., Larson, J., Kirchner, L., & Mattu, S. (2016, May 23). Machine bias. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Arizona State University. (2018). *Connecting the dots at ASU*. <https://fullcircle.asu.edu/external/connecting-the-dots-at-asu/>
- Ash, K. (2013, April 16). InBloom aims to increase data flow despite controversy. *Education Week*. [http://blogs.edweek.org/edweek/DigitalEducation/2013/04/inbloom\\_aims\\_to\\_increase\\_data\\_.html](http://blogs.edweek.org/edweek/DigitalEducation/2013/04/inbloom_aims_to_increase_data_.html)
- Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of machine learning research* 81, 1-11. <http://proceedings.mlr.press/v81/binns18a/binns18a.pdf>
- Bowie, L. (2019, May 28) Simon Newman named Mount St. Mary's president. *Baltimore Sun*. <https://www.baltimoresun.com/education/bs-md-mount-st-marys-president-20141208-story.html>.
- Bulger, M., McCormick, P., & Pitcan, M. (2017). The Legacy of InBloom. *Data & Society*. [https://datasociety.net/pubs/ecl/InBloom\\_feb\\_2017.pdf](https://datasociety.net/pubs/ecl/InBloom_feb_2017.pdf)
- Clark, P., Etzioni, O., Khashabi, D., Khot, T., Mishra, B. D., Richardson, K., Sabharwal, A., Schoenick, C., Tafjord, O., Tandon, N., Bhakthavatsalam, S., Groeneveld, D., Guerquin, M., & Schmitz, M. (2019). *From 'F' to 'A' on the N.Y. Regents Science exams: an overview of the Aristo project*. Allen Institute for Artificial Intelligence. <https://arxiv.org/pdf/1909.01958.pdf>
- Corbett-Davies, S., Pierson, E., Feller, A., & Goel, S. (2016, October 17). A computer program used for bail and sentencing decisions was labeled biased against Blacks. It's actually not that clear. *The Washington Post*. <https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/>
- Dean, L. (2019, November 19). Government announces major robodebt changes. *Yahoo! Finance*. <https://au.finance.yahoo.com/news/robodebt-shake-up-as-government-backs-down-042716284.html>.
- Denver Post. (2016, April 29). InBloom enables great teaching tools. *The Denver Post*. <https://www.denverpost.com/2013/07/18/inbloom-enables-great-teaching-tools/>
- Eveleth, R. (2018, October 29). Facing tomorrow's high-tech school surveillance. *Vice*. [https://www.vice.com/amp/en\\_us/article/j53ba3/facial-recognition-school-surveillance-v25n3](https://www.vice.com/amp/en_us/article/j53ba3/facial-recognition-school-surveillance-v25n3).
- Emerging technology from the arXiv. (2013, October 3) The big data conundrum: how to define it? *MIT Technology Review*. <https://www.technologyreview.com/s/519851/the-big-data-conundrum-how-to-define-it/>
- Family Online Safety Institute. (2019). *Online safety in the age of artificial intelligence*. <https://www.fosi.org/policy-research/online-safety-artificial-intelligence/>
- Future of Privacy Forum. (n.d.). *Ethics, governance, and compliance resources*. <https://sites.google.com/fpf.org/futureofprivacyforumresources/ethics-governance-and-compliance-resources?authuser=1>
- Gellman, R. (2019). *Fair Information Practices: A Basic History*. <https://bobgellman.com/rg-docs/rg-FIPshistory.pdf>
- Green, Ben. (2018). Data science as political action: grounding data science in a politics of justice. *arXiv preprint*. <https://arxiv.org/abs/1811.03435>
- Greenleaf, G., Clarke, R., & Waters, N. (2013). International Data Privacy Standards: A Global Approach (Australian Privacy Foundation Policy Statement). *UNSW Law Research Paper 2013(62)*. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2327325](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2327325)
- Hao, K., & Stray, J. (2019, October 17). Can you make AI fairer than a judge? Play our courtroom algorithm game. *MIT Technology Review*. <https://www.technologyreview.com/2019/10/17/75285/ai-fairer-than-judge-criminal-risk-assessment-algorithm/>
- Haw, M. (2019, September 6). Will AI replace university lecturers? Not if we make it clear why humans matter. *The Guardian*. <https://www.theguardian.com/education/2019/sep/06/will-ai-replace-university-lecturers-not-if-we-make-it-clear-why-humans-matter>.
- Jobin, A., Marcello I., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–99. <https://doi.org/10.1038/s42256-019-0088-2>
- Johnson, J. (2017, February 24). Structural justice in student analytics, or, the silence of the bunnies [Paper presentation]. Digital sociology mini-conference: Eastern Sociological Society annual meeting, Philadelphia. <https://the-other-jeff.com/wp-content/uploads/2017/02/Structural-Justice-in-Learning-Analytics-ESS.pdf>
- Karp, P. (2019, November 19). Robodebt: government abandons key part of debt recovery scheme in major overhaul. *The Guardian*. <https://www.theguardian.com/australia-news/2019/nov/19/robodebt-government-abandons-key-part-of-debt-recovery-scheme-in-major-overhaul>
- Madda, M. (2014, April 21). What InBloom's Shutdown Means for the Industry. *EdSurge*. <https://www.edsurge.com/news/2014-04-21-what-inbloom-s-shutdown-means-for-the-industry>
- Marcus, G. (2019, November 30). An Epidemic of AI Misinformation. *The Gradient*. <https://thegradient.pub/an-epidemic-of-ai-misinformation/>
- Marsh, S. (2019, November 19). Federal Government immediately halts part of Centrelink robodebt scheme. *9News*. <https://www.9news.com.au/national/centrelink-robodebt-to-be-halted-immediately-debts-frozen/c9e2dc06-efda-4e58-a9d2-cdd15a43c856>
- McCarthy, J. (2007). *What is artificial intelligence?* Stanford University. <http://www-formal.stanford.edu/jmc/whatisai/>
- Morse, R. (2013, September 9). Best colleges ranking criteria and weights. *US News & World Report*. <https://web.archive.org/web/20140802233928/http://www.usnews.com/education/best-colleges/articles/2013/09/09/best-colleges-ranking-criteria-and-weights>

Mount President's attempt to improve retention rate included seeking dismissal of 20-25 first-year students. (2016, January 19). *FIRE*. <https://www.thefire.org/mount-presidents-attempt-to-improve-retention-rate-included-seeking-dismissal-of-20-25-first-year-students/>

No Child Left Behind Act 2001, H. R. 1, 107th Cong., 1st Sess (2001).

Sainato, M. (2015, August 19). Stephen Hawking, Elon Musk, and Bill Gates warn about artificial intelligence. *Observer*. <https://observer.com/2015/08/stephen-hawking-elon-musk-and-bill-gates-warn-about-artificial-intelligence/>

Schisler, R., & Golden, R. (2016, January 19). Mount President's attempt to improve retention rate included seeking dismissal of 20-25 first-year students. *The Mountain Echo*. <http://msmecho.com/2016/01/19/mount-presidents-attempt-to-improve-retention-rate-included-seeking-dismissal-of-20-25-first-year-students/>

Scott, A. (2019, April 29). Mount St. Mary's president resigns in the wake of 'bunnies' comment. *Marketplace*. <https://www.marketplace.org/2016/02/29/mount-st-marys-u-quest-better-numbers/>

Selinger, E. (2019, October 14). Why technologists fail to think of moderation as a virtue and other stories about AI. *Los Angeles Review of Books*. <https://lareviewofbooks.org/article/why-technologists-fail-to-think-of-moderation-as-a-virtue-and-other-stories-about-ai/>

Simon, S. (2014, June 8). Big brother: meet the parents. *Politico*. <http://www.politico.com/story/2014/06/internet-data-mining-children-107461.html>

Solove, D. (2018, November 30). Why did InBloom die? A hard lesson about education privacy. *TeachPrivacy*, <https://teachprivacy.com/inbloom-die-hard-lesson-education-privacy/>

Svlurga, S. (2016, January 19). University president allegedly says struggling freshmen are bunnies that should be drowned. *The Washington Post*. <https://www.washingtonpost.com/news/grade-point/wp/2016/01/19/university-president-allegedly-says-struggling-freshmen-are-bunnies-that-should-be-drowned-that-a-glock-should-be-put-to-their-heads>

Svlurga, S. (2019, April 29). The freshman survey that rang alarm bells for some at Mount St. Mary's. *The Washington Post*. <https://www.washingtonpost.com/news/grade-point/wp/2016/02/12/the-freshman-survey-that-rang-alarm-bells-for-some-at-mount-st-marys/>

Study session regarding InBloom, Inc. Colorado State Board of Education (2013) (testimony of Khaliah Barnes).

Reidenberg, J., & Debelak, J. (2009). Children's Educational Records and Privacy. *Fordham Center on Law and Information Policy*. [https://www.fordham.edu/info/23830/research/5922/childrens\\_educational\\_records\\_and\\_privacy](https://www.fordham.edu/info/23830/research/5922/childrens_educational_records_and_privacy)

Roose, K. (2019, February 19). YouTube unleashed a conspiracy theory boom. Can it be contained? *The New York Times*. <https://www.nytimes.com/2019/02/19/technology/youtube-conspiracy-stars.html>







# We asked some NSW students

what they think about the importance of privacy when it comes to using technology...

"I need to get a VPN – this means that I would have a private browser presence pretty much."



**Male student, 16 years old**

"I think privacy is very important and whenever you're playing games online with other players you should never tell anyone any details. Not even if you're a boy or a girl. And if someone asks you a bad question, take a picture of it and show an adult."



**Female student, 12 years old**

"Privacy is very important. If you don't protect your privacy, it can be quite bad. You wouldn't go around with a giant sign in real life telling everyone your full name and where you live and here's the PIN to all my bank accounts! When I'm playing multi-player games online, I never tell anyone what my name is, where I live or my passwords."



**Female student, 11 years old**

"People can hack and get to know where you live and expose your personal information. It is really important to not wear your school uniform on social media."



**Male student, 15 years old**

# In conversation with the 3A Institute:

designing ethical technology  
for a better future

Genevieve Bell, Amy McLennan and Leslie Loble





The COVID-19 crisis has brought into focus the role that technology now plays in our society, economy and even our education system. The rapid development of artificial intelligence (AI) systems raises a range of questions about the safety, responsibility and sustainability aspects of their design and use. Questions such as: What is so different about artificial intelligence? How do we ensure that smart technologies serve the common good? How can education empower students to shape a future that we all want to live in?

Leslie Loble, who recently retired from her role as Deputy Secretary in the NSW Department of Education, joined Professor Genevieve Bell and Dr Amy McLennan from the 3A Institute at the Australian National University to discuss these questions. This paper provides an edited form of that interview.

**Leslie Loble:** I'll start by asking each of you how you ended up working in AI, and what keeps you engaged in this field?

**Genevieve Bell:** I'm a cultural anthropologist by training and I did my PhD at Stanford. I'm also the child of an anthropologist and grew up

on my mother's field sites in Central and Northern Australia in the 1970s and 1980s.

In the late 1990s, I left my tenure track job at Stanford and became a researcher at Intel, which is the first place I encountered computational technology and digital technology

**We now have systems around us that have various pieces of technology at their core that can take information, make sense of it, and act without reference to a human.**

at scale. It was a moment in time where digital technologies were at the forefront of a lot of conversations in the United States. It was the promise of both an adventure and an opportunity to take what I was good at and make it relevant in a very different sphere.

I've spent 20 plus years in that company – I'm still there part time. My job has been to take rich insights about human practice and use them to shape next generation technologies. Over the last 20 years, I've looked at everything from Bluetooth and Wi-Fi to next generation mobile services and social media, and over the last 10 years big data, and then AI. For me, it's less about AI in and of itself, and more about what it represents both as a technical system and as something that human beings are going to have to encounter repeatedly over time.

**Amy McLennan:** My background also starts in the Australian outback, although a very different part of it. I grew up on a farm where lots of conversations were had about the

things that both divide and unite people, such as fences, and whose responsibility the fence was when a sheep jumped through it. Usually, it turned out it was the responsibility of the person who needed the fence fixed the most.

I started my study in the medical sciences and went on to work in policymaking and international development. Something that I learned really quickly was that science works really well in a lab and worked less well in those other spaces. At the advice of one of my mentors at the time, I looked into anthropology and ended up pivoting to medical anthropology. I went over to the University of Oxford where I ended up spending ten years, almost unlearning everything I learned about science and relearning how to think about the world from the perspective of a social scientist and from the perspective of all number of different cultures and peoples around the world. It didn't give me a lot of answers to the questions that I came in with, but I think it did give me a better appreciation of how to frame those questions.

Since then, I have continued to remain affiliated with the University of Oxford, but I've also spent some time moving back into the policy and consulting worlds. I returned to Australia around four years ago to work with the Department of the Prime Minister and Cabinet on all number of complex policy issues. While I was there, a friend of mine called me and said:





“Look, there's a woman at ANU and she seems to be doing some really interesting thinking about the future of technology and the future as we know it. I think you'd really get along, and she needs someone to run a workshop for her.”

So I was really fortunate to be involved in running one of the launch workshops of the 3A Institute. I discovered a really exciting group of people who are working not only on emerging technologies, but thinking about how we build a future we all might want to live in.

**Leslie Loble:** Genevieve, you've previously explained that technology is nothing new by giving an example of how fish traps created out in Brewarrina lasted some 40,000 years – pretty impressive for a piece of technology. But machine learning feels different. Why might that be, and what questions does it raise for us?

**Genevieve Bell:** Back in the 1950s, researchers in the United States proposed what they called the 'artificial intelligence agenda'. This was the notion that you could break down human tasks into small enough pieces that a machine could be made to enact them. Those researchers imagined that AI systems would be able to handle abstractions, understand speech and even learn for themselves.

In framing the problem that way, they were attempting to make machines seem like they were going to become more human. Of course, there's an argument that machinery and technology have always been learning systems. The idea that you could program a system, and you could teach it to do the same thing over and over is old – that's how we taught looms with punch cards nearly 300 years ago. Yet the notion that a



machine could acquire information, make sense of it and replicate itself feels a little bit different. That's what the 'artificial intelligence agenda' has aimed for for over 60 years. We now have systems around us that have various pieces of technology at their core that can take information, make sense of it, and act without reference to a human.

While we could talk about more 'science fiction-y' examples like autonomous vehicles, I'm equally interested in the systems around us that are more mundane. I gave a long lecture at the end of last year where I considered what would happen if you took something as mundane as a lift, and gave it a capacity to learn and act without having to wait. If you summon a non-AI lift, you press a button and the lift comes to you – the lift is

effectively a command and control infrastructure. In the next generation when lifts are running AI at their core, the lift has decided long before you press the button where it needed to be in the stack, and it is anticipating your actions because it has been tracking actions of people inside the building for months, possibly years. It knows that you exit the building at about noon and so do all your colleagues, and so it has arranged all the little lift carriages, anticipating your action long before you press the button. It's that capacity to determine action without a command – that's the thing that feels different.

Now of course as soon as you say that, two things happen. One, it triggers all kinds of science fiction fantasies that all of us grew up with. There are many different stories that we





have been told in movies, books and comics about what would happen when machines could do that kind of thing. Spoiler alert – it never ends well. And of course, it also triggers a set of regulatory and policy issues and principled questions about what it means to have machinery make decisions that have historically required human oversight. Will those decisions be right? How often will they be reviewed? What are the consequences of those decisions? Do you need to have a human somewhere in that conversation? If so, what would that look like? We need to spend our time in the space between the dystopian-utopian narratives that frame our thinking about AI, and the legislation that doesn't yet make sense of it.

**Leslie Loble:** Sticking with the lifts for a moment, we're told that part of the reason we need to move to this system is because it's more energy efficient – it has a good and positive purpose behind it. What are the sorts of things we need to do to shape AI for positive and widespread benefit? How do we even know if the positive benefit is being achieved, if what sits behind it is highly technical and specific knowledge around algorithms?

**Genevieve Bell:** Two good questions. I think the first caveat is that we need to be more precise in our language, about the difference between machine learning, an algorithm and AI. Those are different things. Some

of them are subsets of the other. There have been algorithms as long as there have been patents, even if they weren't always computational. If you've used a washing machine in the last 20 years, you've encountered something that had algorithms. As soon as you press the 'delicates' cycle, someone has made a decision about water temperature and spin, and ringing and agitation, and sequenced them accordingly. The idea of machine learning, for the most part, is just statistics. My colleagues in computer science don't always like it when I say that, but much of machine learning is the application of long-standing statistical methodologies: linear regressions, Bayesian analysis, notions about finding similarities and patterns that are based on older ideas. What's different here, like with algorithms, is the scale and speed at which we can now deploy them, and the tasks to which they're being deployed. AI is the combination of algorithms and machine learning, plus a sensor that will know the world around it and garner more information, as well as some ability to drive action.

The second piece, for me, is to think about what it means to make the objects around us smarter and to be really clear about why we're doing it. So you're right to ask the question about the lifts. What was the intention behind using artificial intelligence inside lifts? Well, part of it was to do a better job of rationalising where the lifts were at any given moment in time, because lifts use

considerable energy. Of course, the trade-off for saving energy is that the lift doesn't come quite as quickly as it used to. So what are the criteria for evaluating the machinery? Is it that you have to wait 30 seconds more? Or is it that it saves an enormous amount of energy over the lifetime of the building? Well, it turns out as humans, we're not really good at making that trade-off, and we're not good at articulating it. We're equally bad at asking the question about what the trade-off should be.

For me, thinking through how we evaluate what makes a technical system good means moving past some of the rhetoric about efficiencies and productivities, and considering other factors such as sustainability, energy use, safety, comfort, fairness and equitableness. Even if you can imagine the right metrics by which you would evaluate these new technical systems, you then need to consider who would do that evaluation. It's hard because much of what's going on inside these systems is deeply complicated.

Beyond that, these systems are now enmeshed in other systems and so evaluating one actually requires having a line of sight over many more. So if you think about the smart lifts, their ability to now talk to other systems within and beyond the building is unprecedented. Evaluating whether the lifts are working well may also mean asking whether the public transportation system that takes people to the building is running


**Thinking through what makes a technical system good means moving past some of the rhetoric about efficiencies and productivities, and considering other factors such as sustainability, energy use, safety, comfort, fairness and equitableness.**

effectively, or how the lifts are drawing from the electrical grid. It's a degree of complexity and systems thinking that can be hard. So that might make you wonder if we are ever going to fix any of it. The good news is that it requires a different set of skills. Amy and I, and the rest of our team, are thinking through what it will take to create a vocabulary and a set of skills that will handle both the questions about the intentionality and the questions about complexity.

**Leslie Loble:** Amy, what skills do people need? And in terms of education, what do students need to know?

**Amy McLennan:** I think there's an opportunity here to flip one of the questions we've been asking – how will technology disrupt education? Instead, we can ask: how could education disrupt, change or shape the technical systems that we're building? If we think about things in that way, then we start a different





**There's an opportunity here to ask ... how could education disrupt, change or shape the technical systems that we're building?**

conversation. If, for example, we wanted to see more sustainability built into AI systems, we can think about how we educate so that technology and sustainability science aren't in completely different silos. Thinking about how we educate for more systems-level thinking, imagining how we educate for a world where that is changing quite quickly, how do we give people the skills to remain on top of those systems regardless of what they are?

Here at 3A Institute, we're identifying the current building blocks of the technical systems around us, and thinking about how we can sharpen our understanding and talk about them accurately. We're learning to frame a series of questions around the way these systems come together and how they intersect with other areas beyond the computational. We're also considering how you go about working with people from different backgrounds, valuing what they bring to the table, and incorporating that into your thinking. While you might not be an expert in something, they might be.

As Bruce Pascoe recently wrote, it's important to teach children to embrace the kind of doubt that leads you to question, be curious and to seek to understand the world. I think beyond all of that, something that I have taken away from working in the 3A Institute is thinking about how to imagine a future we might want to be in, and cultivating that sense of imagination and possibility. That's something that scientific methods won't necessarily get us to, because they are much more focused on how to understand the world we're currently in. So we need to bring in imagination, new worlds, creation and creativity into the conversation as well.

**Leslie Loble:** We only need to look as recently as COVID-19 to realise that a range of skills, capacities, perspectives and backgrounds are essential to solve these problems, even in something as technical as the medical treatment of a pandemic. You're both making the case that young people must develop a sense of curiosity, but also a sense of agency, if they are to control and shape technology, and not think of it as a black box that only certain people have the keys to operate. How do we best develop and deploy that sense of agency in students?


**Genevieve Bell:** It is absolutely the case that our entire educational system is going to require an orientation to technology. But it also requires an orientation to critical thinking. It's not about what programming language you teach people, but about framing an introduction to the fact that there is a programming language.

I tend to think that this is a conversation to have not just with students, but their parents and their communities too. It's sometimes really tempting to say that the younger generation is in charge of the future, but the reality is we're all going to be living in it. For me, that requires thinking differently about a myriad of conversations. I think it's partly about what kind of conversations we want to have in our learning and teaching environments, but it's also about the kind of conversations we want to have in the broader community.

**Leslie Loble:** A fair point for sure. In the Education for a Changing World initiative at the department, we have certainly concentrated on thinking skills as being essential components. While technology pushes the boundaries of what is possible, the further we go as educators the more we come back to the core elements of education. This requires a strong foundation in content knowledge, as well as the capacity to think deeply about any aspect of the world and a sense of agency so that you can act where needed.

**Genevieve Bell:** At the 3A Institute, we often talk to large technology companies, and every one of them is thinking about what skills their next generation of workers will need. These tech companies are looking in particular for people with great critical thinking skills. They want people who know how to ask good questions, and how to work in teams of people who don't share their practices. They want people who know how to drive a conversation forward and also inhabit ambiguity. Of course it's a bonus if people also have tech knowledge – but thinking skills can be much harder for companies to develop than technological skills. So at the 3A Institute, we are thinking about what it means to teach critical thinking in combination with a set of technical skills. When I look at the arc of what will unfold in terms of advances in technology, we know what we have in 2020 will be completely different from what we have in 2025 and 2030. So the challenge is how we give people a way of making sense of new technologies that isn't beholden to the ones they grew up with.

**Leslie Loble:** Amy, part of the transition that Genevieve just referred to is a step into public policy, which you explicitly called out as being an important domain. What is the role of public policy in developing these notions of shaping technology and thinking critically about systems?



**Instead of – what problem is this technology trying to solve? [we should be asking] – what future are we trying to build?**

**Amy McLennan:** In policymaking and in government we could, where appropriate, rethink the way we are approaching a particular policy, regulation or piece of legislation. We have a tendency to think of them as solutions to problems, but there are some moments where it may be valuable to think of them as interventions in a system instead.

When we think of these instruments as interventions in systems, it raises a couple of interesting questions. What are the unintended consequences of what we are doing? How could this particular instrument create changes elsewhere in the system? Instead of – what problem are we trying to solve, what future are we trying to build? Now of course these questions aren't necessarily things that scientific skills will be best equipped to answer. This is where we need to consider who else to bring into the conversation, what other skill sets we might need. Our proposed solution could have economic, environmental, social, health or many other impacts, and we need to bring in people with

expertise in those areas. We also need to think about how we imagine possible futures and which voices are involved in creating those futures. This type of thinking really requires a different approach.

**Leslie Loble:** Another thing that COVID-19 has revealed to us is the importance of transparency. There are thousands of journal articles being written as people desperately search for treatments and solutions to the pandemic. Transparency is also often mentioned as a key component of the guardrails when it comes to technology. What other guardrails do we need for technology design and use, and is there a role for ethical reasoning in particular?

**Genevieve Bell:** Always a good question to ask an anthropologist, because I'm going to make that an even more complicated question than it first appeared. One of the things about transparency is that it's a cultural value – it's not necessarily a shared one. We talk about making things transparent as though that would make everything okay, but the reality is there is power in who gets to reveal and what is revealed in those contexts. I worry that when we talk about transparency, we forget sometimes that there are reasons why certain things are not shared, and the consequences of that. I worry when we talk about ideas that seem on the face of them

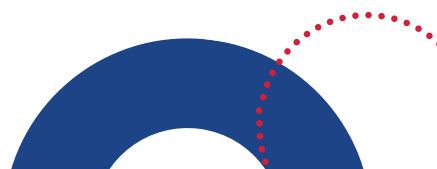


to be really sensible, that they are in fact deeply embedded in forms of cultural logic that don't necessarily scale and are deeply complicated and consequence driven. Which, for me, is the same when we talk about ethics. I think one of the more startling consequences for many people since the start of this pandemic is that a whole series of things that were invisible became visible, whether it was the supply chain of personal protective equipment and ventilators or the supply chain of toilet paper. I suspect for both policymakers and many citizens, making some of those things visible – and revealing their vulnerabilities – will make them very hard to ignore again.


Ethics is not just about morality – what is good and bad – it's about having ethical conversations that are contingent on context and considering

the diversity of perspectives. For example, we often talk about the notion of things being fair. I always wonder who gets to decide what's fair and for whom. Because we live in a society that is both multicultural and multilayered, we may need multiple ethical formulations; it's actually quite hard to articulate a single definition of what's fair. What is ethical also changes over time, despite the fact we might like to imagine that it's constant. Some of the ethical norms of Australia 50 years ago were quite different and would be considered unethical by today's standards.

Many of my colleagues in large tech companies have asked for an ethical framework because they'd like to build it into the machinery. But there isn't one set of timeless and universal ethical principles that we can hardwire into technology to solve all our








problems – it's actually an ongoing conversation that we need to have. We need to consider who is and is not involved in the conversation, who the technology may impact, and what the consequences could be of having an ethical framework in place, or not having one. It involves critically asking a set of complicated questions to make sure that we don't inadvertently build something that no one wants to live with. So, long before we think about ethical frameworks, we should be having public conversations about laws and policies and how we want to regulate AI design and its application. It's not just about designing for a hypothetical trolley problem, the question we need to answer is: how do we design for the daily lived experience of the broader community who are going to be affected by AI technologies?

**Leslie Loble:** Both of you are painting a powerful picture of complex, interactive, constantly changing systems and interrelationships at the human level, the system level and the technological level. We're often encouraged to see science and maths skills as the solution. What you're saying is that, in fact, there's tremendous need for many different bodies of knowledge, expertise and perspectives, which I think is quite exciting. When you think about it in educational context, computational thinking is incredibly important, but it has to be coupled with a much wider and diverse set of skills.



**There isn't one set of timeless and universal ethical principles that we can hardwire into technology to solve all our problems – it's actually an ongoing conversation that we need to have.**

**Genevieve Bell:** Oh, absolutely, and I watch with interest. In the United States, the big engineering and computer science schools are adding in humanities and social science subjects. Getting degrees in those fields now requires you to be in dialogue with these other disciplines because it is clear to the leading lights in various fields that they've been having conversations without all the right people in the room. So for me, as I watch the re-energising of relationships across the academic disciplines, I see those as being really hopeful things.

**Leslie Loble:** Amy, if you were speaking to a student and their parents at the kitchen table, what would you say about their future and what they should be doing with their learning?

**Amy McLennan:** I think it has to be a message about exploring widely. It's really easy, especially when you get to the pointy end of school, to be focused on choosing your subjects, and then your degree or career path –

which of course automatically narrows things down. I think potentially choosing one thing to learn that is the direct opposite of where you think you wanted to go is a neat way in. If you can talk about that over the kitchen table and debate it with your family, just for fun, then that can be a really simple and exciting way to bring a different way of thinking into your everyday studies.

**Leslie Loble:** Genevieve, what would you say to teachers? And indeed, what would you say to Mark Scott, myself or other leaders of an education system?

**Genevieve Bell:** I think the most powerful thing I heard Mark Scott say was two years ago at the Education for a Changing World symposium. He said that in 2018, he had to have an eye on 2030. I remember looking at him and thinking that it is the hardest job in the world to have to think about the distant future and the immediate present at the same time. So I think the hardest challenge that teachers and educational leaders have is that they are simultaneously in the immediate present and the distant future. They have to balance constantly the arc of twelve years of education and the realities of what has to happen today. I think that is really tricky, because it requires having an incredibly optimistic, yet flexible, view of the future, but also being aware of the things that need to be done today and tomorrow. While it's important to have a view about the future, you must ensure that it doesn't become so clearly decided as a destination that you don't have room to move.

**I think the hardest challenge that teachers and educational leaders have is ... to balance constantly the arc of twelve years of education and the realities of what has to happen today.**

I imagine that one of the interesting complexities of being an educator at the moment is that the last few months were never anticipated. In a matter of weeks, the education system in Australia had to move from classroom delivery to a kind of distance education. That must have both created significant difficulties, but also some unexpected green shoots. So for me, the challenge about managing something that is effectively a future industry is about how you balance between the immediate future, and how you have a vision for something that is more distant. I think it's important to acknowledge that it's actually really hard to sit between those two things.

**Leslie Loble:** That's a great way to end it. Thank you both for a terrific discussion. Your insights, knowledge, optimism and enthusiasm come across very powerfully.



# We asked NSW students

how they'd like technology to support their learning in the future...

I didn't have enough time to do all my work [when learning from home] because I was using my mum's computer. So, mine would be to ask Microsoft about making a certain app for Xbox for education. Because if, like, you don't have any devices because your parents were using the computer, you could use classrooms on the Xbox. It would make a big improvement with people.



**Male student, 9 years old**

In the future I'd like to use programs to create art online.



**Female student, 11 years old**

In the future, I'd like to be able to make things easily, using coding.



**Male student, 9 years old**

I think a technology that could help us learn is, let's say that people with disabilities that are blind, like some sort of app on the iPad that the teacher could upload stuff for home learning and the video could be read to them. Other people in class could post things onto this app and help out this person.



**Female student, 12 years old**

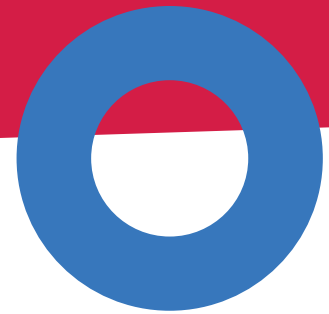




# Learning together:

using technology to support expert teaching practice

Mark Scott



Education Week is usually held in the middle of Term 3 and it's always a fantastic celebration of local schools and students.<sup>1</sup> It's a showcase of excellence and a reminder of the role public schools play as cornerstones of their communities.

This year the way we celebrated was a little different, of course. In the face of bushfires, floods and a pandemic we all needed to come together and celebrate education more than ever, but we knew that in mid-2020, in-person events across the state – in schools, communities and our corporate offices – were just not going to happen. So, at the start of August, Education Week went virtual, with the fitting theme 'learning together'. An exciting program of virtual events was developed and schools across the state shared videos of learning together with their own communities.

In such a challenging year, it was heartening to see over 2,200 schools come together to celebrate learning and community.

On the last day of Education Week, I was fortunate enough to be scheduled to meet with a group of students from across the state. They shared their experiences of remote learning and asked me plenty of tricky questions. One thing they all agreed on was that they now realised more than ever how much they value the work teachers do to support their learning. One student, Joseph, said:

---

1. Education Week 2020, our annual state-wide celebration of NSW public education, ran from Monday 3 August to Friday 7 August 2020: <https://education.nsw.gov.au/public-schools/education-week>.





“The way teachers communicate with students is incredibly important in my opinion ... just communicating via email or messages is not the same as being physically in the room with them.” Another student, Genoveva, appreciated her teachers checking in with her daily, and reflected: “I feel like it was a moment of realising how fortunate we are to be able to be going to school in person and having those face to face lessons. Because when we were forced to go online ... it was hard.”

The experience of COVID-19 and remote learning has reinforced for me that, as technology advances, the person at the front of the classroom becomes even more important. As our entire system switched to remote learning, it was our teachers and

**The experience of COVID-19 and remote learning has reinforced for me that, as technology advances, the person at the front of the classroom becomes even more important.**

school leaders who were delivering education to our students and checking on their wellbeing – not a computer. Even when it is used to connect in new ways, technology amplifies the need for personal connections in the profession and expert craft that is teaching. It’s hardly surprising our students missed seeing their teachers in person.

We've never seen technology used on such a scale in education as we have this year. New technology was adopted practically overnight.

While 2020 has been the most disruptive year of learning in several generations, it may also be the most transformative. At the start of the year, the main debate was around how much technology should be used in the classroom, and the limits of what a screen experience can provide. I think too much attention can be paid to devices, and not enough to the way we are actually using devices for teaching and learning. I've previously written about the risk that shiny new technologies can distract us and divert our attention from where it's most needed. The answer, of course, isn't to deny students access to digital technology, because for these 'digital natives' it is a tool that will be central to their post-school lives and careers. We must prepare them to use technology effectively and responsibly.

Taking advantage of digital tools was a core feature of the Education NSW response to COVID-19. We already knew these tools had the potential to transform how learning was delivered but this was an unprecedented opportunity to see that potential unfurl in the real world, in real time. How students experienced learning from home varied. For some it was a challenging period but we also heard students speak of being more motivated by the increased responsibility for their own learning, and we can take valuable lessons

from that feedback. Now, I believe, the debate around technology and education can move on from whether or not iPads should be in the classroom, to focus on how we can best use technology to enhance learning and teaching.

We hear a lot about technology and personalised learning – that advanced technology will usher in a new era of education tailored to the individual needs of each student. But it's vital to recognise that technology is a tool, and it will always be expert teachers who can best deliver rich learning experiences. Artificial intelligence (AI) algorithms can deliver content that we find engaging, but it is through human connections that we are inspired to try to master new domains and skills.

COVID-19 has reinforced what we already knew: having a teacher to support students in their learning is irreplaceable.

### Hindsight is 2020

The experience of COVID-19 was, of course, far from uniform across the state. Some communities were just starting to rebuild following last summer's devastating bushfires and floods when they had their lives profoundly disrupted yet again. Some schools in Sydney, where the concentration of the virus was greater, temporarily closed completely due to positive cases. All our schools offered remote learning from 24th March to 25th May this year. In many schools,





particularly those in rural and remote areas, a lack of access to reliable internet or devices presented equity challenges. These differences made clear that a one-size response would not fit all.

As a department, we've had a long experience with distance education and were able to use our expertise during the remote learning period. This included drawing on the knowledge and resources of Aurora College, a selective, virtual high school that provides students in rural and remote communities with face-to-face online classes and a challenging academic program. We established a Learning from Home Hub that contained more than 400 online teaching and learning resources, and also partnered with the ABC to provide students of all ages with access to educational content at home. The department implemented virtual state-wide

staffrooms, where staff could get advice, share resources, and support each other – with over 30,000 teachers joining them.

We worked quickly to put in place professional learning to upskill teachers on the use of digital platforms. Many of our teachers are now more prepared than ever to integrate technology effectively into their lessons. Dwayne Hopwood, Principal of Ashfield Boys High School, explained that teachers are “reporting that they will be able to use the digital skills they've learned to enhance their teaching going forward.” Across the system we've seen that staff are now more comfortable with new technologies and are considering how they can continue to use them to improve teaching and learning. This increased capability presents a golden opportunity for teachers to leverage the power of technology in the classroom.

**Technology is a great tool for learning, but it's only ever the right tool when carefully selected and managed by educators who lead the delivery and personalisation of learning.**

During COVID-19, some of our biggest challenges as a department were to do with equitable access to technology, and this is an area we are continuing to focus on. We worked fast to procure devices and internet dongles and get them out to students who needed them most, particularly prioritising Year 12 students, and many schools provided devices to their students. One thing I've learned is that, as a department, we hadn't fully understood the importance of having devices at home to supplement the learning at school. The COVID-19 experience has revealed that we may need new strategies to support the movement of devices between home and school, to make it easier for every child to access the wealth of resources available to them, wherever learning takes place.


Overall, I am extremely proud of the way our staff members and schools supported students in transitioning to virtual learning – and equally proud of how students responded. This year has presented many challenges, no doubt; however, with the capability uplift we have seen in staff and our continued focus on giving all students

a level technological playing field, we are now in a very strong position to make full use of new technologies to complement teaching and improve learning outcomes.

### Tailored solutions and technological transformation

Technology is a great tool for learning, but it's only ever the right tool when carefully selected and managed by educators who lead the delivery and personalisation of learning. For school leaders and teachers during the remote learning period this year, implementing effective remote learning solutions required knowing the school context, understanding the learning needs of the students and communicating effectively with families.

At Canley Vale Public School, relieving Principal Brad Lanham recognised early on that COVID-19 had the potential to cause significant disruption to his students' learning. As principal of a school where 97% of students have a non-English speaking background, ensuring equity during online learning was a key focus. The team at Canley Vale worked hard to ensure that every student had access to the technology they needed, and used technology to engage their community: the school included QR codes on letters that parents could scan to hear the important messages read in their preferred language. Parents reported that this made them feel much more confident in supporting their child's learning.



More than 600 kilometres north of Canley Vale, Rowena Public School is a small school of 28 students in a farming community in far north NSW. There is patchy, and at times non-existent, internet access across the community, coupled with inconsistent access to devices amongst the student population. Moreover, in March 2020 Rowena had just experienced its best rainfall in years and many students were working on family farms during the day to take advantage of the upcoming growing season. This meant that students were learning at irregular hours, and parents (like parents everywhere) were doing their very best to help with schoolwork while juggling other commitments.

Knowing the school and community well, Principal Paul Cecil and his teachers developed a personalised learning schedule for each student, included in a hardcopy lesson pack, which reduced the need to rely on internet access. The school also loaned devices to students who needed them, provided links to pre-recorded lessons so that they could access them when it was convenient, and set up easy ways of contacting teachers for feedback. Like Canley Vale, with deep care and considered judgement, educators at Rowena were able to effectively deliver remote learning for their school community.

We also saw a constant drive among school leaders and teachers to improve their remote learning approaches. Wisemans Ferry Public School had experienced bushfires and floods in the months prior to COVID-19, which meant that the school's teachers and students were already experienced in learning from home. Principal Deirdre Dorbis said: "It was just a matter of refining our practice, learning from the floods to ensure that we could do it a little bit better." Wisemans Ferry's experience shows how passionate schools are about improving their practice, including through the use of appropriate technology, to ensure students are getting the best outcomes.

While the transition to remote learning was certainly challenging, there have been plenty of positive outcomes. Many students flourished in the remote learning environment, and school leaders and teachers have told me that it has been one of the most significant professional learning experiences of their careers, particularly in terms of digital uplift.

One school where remote learning has had a transformative (and no doubt lasting) impact is Northern Beaches Secondary College Freshwater Campus, which caters to Year 11 and 12 students. Principal Frank Pikardt said that "in terms of using technology in the classroom, I think it's accelerated

us five to 10 years.” The school continued to use technology after the students returned to their classrooms, and it has improved engagement in class, and supported teachers to further tailor their teaching to students’ learning needs.

Likewise, at Warilla North Public School in Wollongong, learning from home gave families an opportunity to be more involved in their child’s learning and support the school’s high expectations culture at home. As a result, students were submitting more work than ever and the community became more connected. Principal Nicole Riley explains that “technology has given us a chance to get ‘back in the game’ and we are experiencing real success in student growth across literacy and numeracy which is giving the whole school community a real boost.”

### NSW AI Strategy

Our system’s rapid embrace of digital technology to facilitate learning during COVID-19 puts us on track to address wider trends in technology and education, including advances in AI. We already interact with AI on a daily basis – whether we are asking Siri for the weather forecast or checking our pre-filtered emails. In education, AI is developing rapidly, with a raft of AI-based technologies now available. These include AI-powered learning assistants, which can adapt learning content to best suit student needs and provide tailored feedback and other recommendations, such as the best time of day to study.

**Embracing AI in education means establishing the right guardrails to protect our students.**

In addition to AI, other forms of advanced technology – such as the use of predictive analytics to forecast student learning outcomes – are having an increasing influence on education globally. While such technology could potentially support teachers in their work, without strong ethical guardrails there are also substantial risks to students, as Evan Selinger and Amelia Vance discuss in this issue of Future EDge.

The increasing power of smart technologies make human expertise and judgment more – not less – important. We know that rigorous and holistic assessment of a student’s learning needs is informed by analysis of data. AI can provide teachers with real-time insights into student progress, learning processes and potential barriers to learning. This quick information might inform a teacher’s decision-making to support student learning. If used appropriately, AI-based applications could free up teachers’ time to focus on what really matters: quality teaching.

As AI becomes more sophisticated, it will provide further opportunities to personalise learning for students and reduce the administrative burden for teachers. We often hear about





personalised learning as a way of addressing the diverse learning needs found in every classroom. But we must remember that AI cannot replace a teacher – it can merely assist the teacher to better understand the needs of the student. The teacher, not the technology, will be supporting students to reach their learning goals.

Teaching requires innovation and creative thinking, an ability to build rapport and inspire. An expert teacher supports and motivates learning in ways that technology will never be able to. Teachers and school leaders will of course take on greater responsibility for the use of AI and other technologies, as they do for any other educational approach they use. Our expert educators are best placed to judge when the use of technology is helpful, appropriate and safe. As a department, our role is to ensure

that they have the support and skills they need to enhance their teaching practice, so that students can benefit from AI and be protected from its risks.

The potential for risk raises important questions about how to navigate the appropriate use of smart technologies in the classroom, so students can enjoy the potential benefits of these technologies in a safe way. AI is created by humans, and can be imbued with our biases and blind spots. Embracing AI in education means establishing the right guardrails to protect our students. We also need to ensure that education is in conversation with tech companies, so we can contribute to building technologies that are fit for purpose and include diverse perspectives in their design.

The recently announced NSW Government AI Strategy will help to ensure that AI solutions used in education are clearly focused on student needs, trusted by our communities and meet the highest ethical standards.<sup>2</sup> The strategy is organised around five key principles:


1. **Community benefit** - In education, this means that the needs of the student and school community must always come first in determining whether AI is the most appropriate tool.
2. **Fairness** - AI applications used in schools must be high quality, fit for purpose and (crucially) include safeguards to manage data bias or data quality risks.
3. **Privacy and security** - All data collected must be for the purpose of improving student learning, and be stored securely with sufficient privacy protections in place.
4. **Transparency** - Students, and/or their parents and carers, must be able to access the data that has informed a decision or approach, and be able to challenge those decisions.
5. **Accountability** - The department, schools, school leaders and teachers are responsible for decision-making that uses data gained from AI applications.

We are already using AI effectively across the NSW government. For example, NSW hospitals are piloting AI technology to better detect patients with sepsis in hospital emergency department waiting rooms. The tool analyses information entered into a patient's electronic medical record, identifying the more difficult to identify signs of sepsis, and providing emergency room clinicians with a report of patients most at risk. Patients can clearly benefit from faster detection of a potentially deadly infection, and doctors can use this information to provide the best care. Similarly, as a teacher, AI can help you identify a learning need, but your expertise is required to implement the most effective intervention for the student.

At the same time, it's not enough to have access to the tools. We need to make sure that teachers and students have the support and skills they need to use them. That means continuing to build on progress we have already made as a system in improving digital literacy and efficacy. To this end, we are also implementing a Schools Digital Strategy to ensure that we build digital equity and capability in every school. Our goal is to empower schools to shape their own digital journey, and leverage new and emerging technologies to deliver rich, engaging, personalised learning in every classroom, for every student.

---

2. The NSW AI Strategy was released on 4 September 2020, and can be accessed via [www.digital.nsw.gov.au/policy/artificial-intelligence-ai/ai-strategy](http://www.digital.nsw.gov.au/policy/artificial-intelligence-ai/ai-strategy)



**As a teacher, AI can help you identify a learning need, but your expertise is required to implement the most effective intervention for the student.**

### Concluding thoughts

While COVID-19 and the bushfires rightly consumed much of our attention in 2020, it has also been a landmark year for Australian education reform. In June, the final report of the NSW Curriculum Review was released, and the NSW government committed to reforming the curriculum by the middle of this decade. A review of the Australian Curriculum is underway, and reviews of Senior Secondary Pathways and NAPLAN were also released this year.

The NSW Curriculum Review included recommendations from Professor Geoff Masters on how we can build a curriculum that ensures ‘every student leaves school well-prepared for a lifetime of on-going learning and informed and active citizenship’. Professor Masters notes that rapid advances in technology mean that the ‘work of the future will require employees with deep understandings of subject matter who can think

critically about problems and use their understandings to create new solutions.’ It’s more important than ever that we have in place a curriculum that ensures our students gain both the core content knowledge and the key skills they’ll need to thrive in a complex world. And a smart, effective and safe approach to the use of smart technologies in the classroom will free up teachers’ time to focus on teaching core content and skills, and provide them with the data they need to ensure that all students can meaningfully access the curriculum.

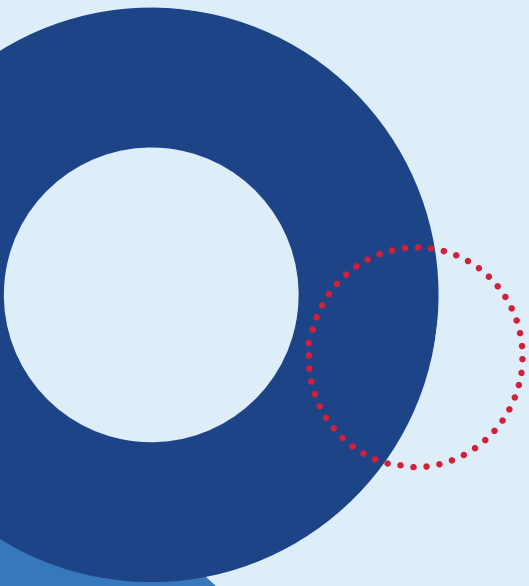
By learning from our experience during COVID-19, and developing and implementing policies such as the NSW AI strategy, we can ensure that students are getting the maximum possible benefit from technology. One of the department’s key commitments is to ensure that every student and every school improves every year. Harnessed well, AI and other smart technologies will help us towards this goal; but ultimately, as this year has reinforced once again, it will be the judgment, expertise and dedication of every teacher in every classroom that makes all the difference.











© State of NSW (Department of Education) 2020

