

University of New South Wales Law Research Series

HELPING FUTURE CITIZENS NAVIGATE AN AUTOMATED, DATAFIED WORLD

LYRIA BENNETT MOSES

[2019] *UNSWLRS* 28

UNSW Law
UNSW Sydney NSW 2052 Australia

EDUCATION: FUTURE FRONTIERS

OCCASIONAL PAPER SERIES

Helping future citizens navigate an automated, datafied world

Lyria Bennett Moses



ABOUT THE AUTHOR

Lyria Bennett Moses is a Professor and Director of the Allens Hub for Technology, Law and Innovation at UNSW Law. She leads the UNSW Grand Challenge on Living with 21st Century Technology, and is a Key Researcher and Project Leader on the Data to Decisions CRC.



© Lyria Bennett Moses, UNSW-Sydney and the State of New South Wales (Department of Education), 2019. EDUCATION: FUTURE FRONTIERS is an initiative of the NSW Department of Education exploring the implications of developments in AI and automation for education. As part of the Education: Future Frontiers Occasional Paper series, the Department has commissioned essays by distinguished authors to stimulate debate and discussion about AI, education and 21st century skill needs. The views expressed in these essays are solely those of the authors.

INTRODUCTION

If there is one thing everyone has an opinion on, it is education. Everyone has been to school and many people go on to have opinions on the experiences of their children. I am therefore conscious that I am neither a school principal nor a teacher, but rather stand at a unique point in the broader education system. In particular, I teach tertiary students enrolled in a law degree and conduct research at the intersection of law and technology. One subject I teach is practical – how to design expert systems and run a small technical project for a not for profit to improve access to justice. This requires students to think about technical solutions to assist in resolving a practical challenge in access to justice (navigating complex information, under-resourced centres, understanding community needs). The other subject is more theoretical – what legal issues arise as technology changes, how does law deal with disruptive innovation, and how can we retain core protections in the face of developments such as automated data-driven decision making. I am therefore accustomed to encouraging students to think outside the disciplinary box to which they are otherwise assigned.

Because of my role, I do have some insights into the challenges of teaching in an interdisciplinary context. I am also researching legal and ethical challenges that society is facing relating to data-driven inferencing and decision-making. Many of the new tools being developed are described as artificial intelligence, and I will discuss the relationship among these ideas below. Developing technologies including those associated with artificial intelligence, automation and machine learning, are changing the world that young people need to learn to navigate. When tools are used by those who don't understand them, there

can be individual and community harms. There are thus important questions being asked about how the education system needs to take account of these new technological developments.

This essay articulates a vision or idea about how education could best equip students to navigate an increasingly automated and datafied world. It also develops a broader argument about cross-disciplinary learning that can be applied to other important modern challenges that cannot be solved within a single discipline.

ARTIFICIAL INTELLIGENCE, AUTOMATION AND DATA-DRIVEN INFERENCING

The broad ambitions of artificial intelligence are evident in the ways that it is defined as mirroring human capability. Alan Turing first suggested that artificial intelligence could be confirmed if a human interrogating both a computer and another human (providing responses from behind a closed door) was unable to determine which was the machine. More commonly, artificial intelligence refers broadly to computers and other devices or applications that function as if possessing human intelligence.¹ However, the definition is not consistently applied.² There are many machines that perform mundane tasks that used to require human intelligence (such as scientific calculators) that are rarely described as artificial intelligence. Instead, the term focusses on more recent, awe-inspiring systems that can perform more complex tasks. This leaves the definitional boundaries unclear, and perpetually on the cusp of an innovation wave.

The development of technologies that can replicate human intelligence raises questions about future employment prospects for today's students. There is an ongoing challenge of "occupational obsolescence", the tendency of

¹ Alan Turing, 'Computing Machinery and Intelligence', *Mind* 59 (236), 1950, 23.

² Macquarie Dictionary (online at 27 February 2019).

³ Andries De Grip and Jasper Van Loo, 'The Economics of Skills Obsolescence: A Review', *Research in Labour Economics* 21, 2002, 211-226.

IT IS ALREADY CLEAR THAT MACHINES CAN PERFORM SOME TASKS AT OR ABOVE HUMAN CAPABILITY, AND THAT THIS MAY WELL INCREASE. FINDING AREAS WHERE HUMANS CAN MAKE A UNIQUE CONTRIBUTION IS THUS ESSENTIAL TO FUTURE EMPLOYABILITY.

particular ways of doing and roles for humans to change as technology evolves.³ Predicting technological paths is never easy, but it is already clear that machines can perform some tasks at or above human capability, and that this may well increase. Finding areas where humans can make a unique contribution is thus essential to future employability.

Examples of artificial intelligence continue to capture the human imagination. Artificial intelligence is a common theme in science fiction, from *Terminator* to *Her*, with hypothetical future robots taking over the world or falling in love and abandoning humanity. There are, of course, very real achievements such as victories against human champions in games such as Chess and Go. These developments, both real and imaginary, lead to an important question – how do we prepare future citizens for a world where they will work and play alongside increasingly sophisticated machines?

Alongside the exciting possibilities presented by artificial intelligence research are a range of often-flawed or misapplied systems that

replace human decision-making in practice. A well-known example in Australia that involved relatively simple computer programming is the the so-called Robodebt system. Here, letters were sent to individuals concerning a debt owed to the Commonwealth where the calculation was performed by a system that compared (annual) reported taxable income with (fortnightly) earnings. The flaw here was not one of artificial intelligence but rather one where a generalisation was applied to everyone without sufficient recognition of individual circumstances (or sufficient mechanisms to allow affected individuals to understand and correct the record). Thus those whose income was uneven over the course of a year, and thus whose fortnightly income could not be directly calculated from an annual figure, were sent debt notices for amounts that they did not, in fact, owe.

While Robodebt is an example of automation being used to treat different circumstances as if they were identical, artificial intelligence commonly relies on machine learning techniques that have the opposite effect. In particular, it allows a computer to “learn” how to correctly classify something based on measurable differences. Machine learning involves many different techniques which share some features. They rely on a large quantity of sample or historical data from which the machine will learn to draw inferences (such as which properties of a thing are associated with which classification of that thing). Systems relying on machine learning act intelligently in the sense that they learn over time and can thus be responsive to feedback, and also in the sense that the inferences drawn can be useful and may have been difficult for a human to deduce. For example, machine learning could

It is important to note the difference between predicting a decision that will be made based on a machine learning model and making a decision that will have a real-world effect. This is obvious if you consider a system that can predict exam results. One would not want to cancel an exam and treat the outputs of the system as the student's final result. Instead, the impact of such a system would likely be indirect, for example, enhanced surveillance within the school community, the opportunity to select students for additional support and/or the provision of better targeted careers advice to Year 12s.

As simple as that distinction seems, it is often the case that systems designed for prediction become part of a decision-making process. As a system for predicting individuals who are likely to owe a debt, the Robodebt system would have performed reasonably well. But its logic was insufficiently certain to justify instantaneous implementation in the form of a debt notice.

An example of a system that makes predictions with real world effect, somewhat problematically, is predictive policing.⁴ Predictive policing software draws on data to predict where, probabilistically, crime is most likely to take place. Police deployments are then based on system outputs, so future police patrols are greater where crime is most likely to take place. While this may help prevent crime, it also feeds back into the system. This is because patrolling police are more likely to record crimes they encounter in more highly patrolled areas. This feedback loop makes it more likely that an initial prediction of a neighbourhood as high risk will be perpetuated.

The still hypothetical notion of judges being replaced with machine learning systems would be an even more problematic conflation of prediction with decision-making. Such systems may be able to predict and thus mimic the judicial decision-making process, at least probabilistically, but they would not be exercising judgment.

be used to predict whether a student is likely to pass an exam based on variables such as previous marks, level of interest in subject and amount of time spent studying. Where the data to be classified is unstructured, as in the case of systems that identify objects in images, so-called deep learning techniques like neural networks are used.

LIVING IN AN AUTOMATED, DATAFIED WORLD

Current school students are growing up in a world where these techniques are increasingly being applied in ways that affect what information or texts they see, what choices are available to them and what decisions are made about them. As such, they need to understand something about these techniques to navigate the world around them.

Consider something as simple as the texts and ideas to which people are exposed. This has always been affected by both individual preference and social context. For example, people in particular suburbs may be more likely to read a particular daily newspaper and the choice of what books to read (outside of school) may be influenced by friends. What has changed is the complexity and opaqueness of the process by which information is presented to us. On social media and on commercial platforms such as Kindle, algorithms classify us in various ways (based on data such as past choices and online activities) and then suggest books, news articles, posts, accounts to follow, and so forth. Even search engines, many people's primary entry point to the wealth of material on the World Wide Web, sometimes individualise search results or associated advertising based on what they know about us from previous searches, previous choices among search results and known or inferred characteristics such as location or gender. This individualisation of content has implications for those individuals, as where gender impacts what advertisements a person sees

⁴ Lyria Bennett Moses and Janet Chan, 'Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability', *Policing and Society* 28, 2018, 806.

INFERENCES BASED ON DATA CAN ALSO HAVE A ... DIRECT IMPACT ON AN INDIVIDUAL'S LIFE.

as they browse the Web.⁵ It also has social impacts, including the creation of filter bubbles that polarise political debate – where the information to which each person is exposed differs, it is hard to find common ground from which to debate questions of policy. Whether or not there is foreign interference or fake news campaigns, the structure of social media means it is rare to see fully developed arguments or fully articulated reasoning. It is difficult to create opportunities for rational and structured argument when each side of a debate can base its analysis on diverse sources that the other side is unlikely to access.

Inferences based on patterns in data can also have a more direct impact on an individual's life. In a commercial context, pricing is increasingly based not on direct bargaining (as in a traditional market) or standardised pricing (as in a supermarket) but rather individualised pricing based on inferences drawn from data. In particular, pricing algorithms can deduce relative wealth and predisposition to discounts from information such as operating system used, location, past purchasing patterns, word search terms, and so forth. As consumers, it is no longer simply a question of deciding whether a 1L or 2L bottle offers better value, but rather how to construct a digital identity that maximises the likelihood of lower price offers.⁶

In addition, inferences drawn about individuals through the use of data analytics are sometimes used by the state to draw inferences with real impact on citizens' lives. An extreme example of this is the use of risk assessment tools in the context of sentencing offenders in the United States. Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is a case management and decision-support tool that is used in some US jurisdictions to assess recidivism risk for the purposes of bail, parole, sentencing and/or prisoner management decisions. Its use by judges in the context of determining non-parole periods in sentencing is particularly controversial but has been upheld in Wisconsin.⁷ There is, in particular, controversy about the tool's racial bias.⁸ On the one hand, it has been found that there was a higher false positive rate for African American compared to white populations when the tool was tested in one community.⁹ On the other hand, the company producing the tool claims that the probability of an individual reoffending, given a high-risk score, is similar for both racial groups.¹⁰ As has been pointed out, it is impossible to meet both fairness standards simultaneously.¹¹ The challenge is that a debate about the appropriate statistic to represent fairness is not a purely mathematical question, but includes an understanding of the decision-making context (here, sentencing), the democratic and human rights commitments of a state, and philosophical questions about equal treatment as an ideal.¹² The different ways of measuring unfairness here do not answer the question, but rather are relevant facts in a broader debate.

⁵ Dina Fine Maron, 'Science Career Ads are Disproportionately Seen by Men', *Scientific American* (25 July 2018), available at <https://www.scientificamerican.com/article/science-career-ads-are-disproportionately-seen-by-men/>.

⁶ Price discrimination is, in principle, legal in Australia unless it is evidence of other misconduct (such as anti-competitive arrangements, misuse of market power or breach of privacy laws). An earlier prohibition in *Trade Practices Act 1974* (Cth) s 49 was repealed by the *Competition Policy Reform Act 1995* (Cth).

⁷ *State of Wisconsin v. Loomis*, 881 N.W.2d 749 (Wis. 2016). The United States Supreme Court denied certiorari on 26 June 2017.

⁸ One could of course also critique racial bias in human decision-making, which is also an important topic.

⁹ J. Angwin et al., 'Machine Bias', *ProPublica* (23 May 2016), available at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (last accessed 16 August 2018).

¹⁰ Northpointe, *COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity* (8 July 2016).

¹¹ Sam Corbett-Davies, Emma Pierson, Avi Feller and Sharad Goel, 'A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear', *The Washington Post* (17 October 2016), available at https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/?utm_term=.5b592875af58.

¹² See generally Andrew D Selbst et al., 'Fairness and Abstraction in Sociotechnical Systems', ACM Conference on Fairness, Accountability, and Transparency (FAT*) 1(1), forthcoming.

While this particular tool is not currently being used in Australia, we have similar issues with classifications of individuals being assessed as high risk in a racially biased way. In particular, the New South Wales Police Force has used a Suspect Targeting Management Plan (STMP) that seeks to identify and target individuals who are assessed as likely to commit future crime. While the consequences for targeted individuals under the NSW scheme are different from those in the US, there are real negative consequences and young people targeted report perceptions of oppressive, unjust policing.¹³ As in the US, there is a racial bias in that Aboriginal and Torres Strait Islander people are disproportionately affected.¹⁴ The statistical questions that have been asked about the STMP are different from those analysed in the US, but the same question remains crucial, namely what might a metric for “fairness” in this context look like? Again, the answer requires an understanding of statistics, but not in isolation from the broader social and historical context, including the troubled history of interactions between indigenous people and agencies of colonial, then Australian, powers.

The role of humans in these systems is crucial, as they are often the principle point for deciding how a system will be designed, what questions will be asked about its acquisition or use, what regulatory requirements apply, and so forth. These questions will not necessarily be answered by those with deep disciplinary knowledge of the techniques but also by the managers, lawyers, investors, administrators and consumers who interact with systems. The

understanding that humans have of those systems will be crucial in ensuring that applications are safe, appropriate, fair and effective (as required in the specific context).

This is the world of the immediate future; the world of the distant future is likely alien to even the best forecasters. But encouraging critical thinking about what it means to engage in this world – as citizen, as consumer, as policy-maker, as designer, as user – is essential for everyone. There is an opportunity for the education system to play a role here, a chance (and arguably a responsibility) to equip future citizens with the ability to thoughtfully navigate an increasingly automated and datafied world.

DATA LITERACY AND COMPUTATIONAL THINKING

There is a need to distinguish between the expertise needed to design systems that automate decision-making processes and the expertise needed to use such systems well. Everyone should have a sufficient understanding to be an effective user of such systems and participate in democratic decision-making about the appropriate deployment of such systems. However, it is not realistic to expect high school students (let alone every high school student) to be able to code a machine learning algorithm, let alone build a robot capable of deep learning. Advanced courses offered at tertiary level will continue to be the training ground for future data scientists, statisticians, roboticists, programmers, systems designers and AI researchers.

¹³ Vicki Sentas and Camilla Pandolfini, *Policing Young People in NSW: A study of the Suspect Targeting Management Plan* (Sydney: Youth Justice Coalition NSW, 2017).

¹⁴ Ibid.

¹⁵ Toby Walsh, 'The AI Revolution', in Leslie Loble, Tish Creenaune and Jacki Hayes (eds.), *Future Frontiers: Education for a Changing World* (Melbourne: Melbourne University Publishing, 2017), 1.

¹⁶ Resolution 7 In Support of the Guiding Principles on Using Risk and Needs Assessment Information in the Sentencing Process, adopted as proposed by the CCJ/COSCA Criminal Justice Committee at the 2011 Annual Meeting on 3 August 2011, <http://ccj.ncsc.org/-/media/Microsites/Files/CCJ/Resolutions/08032011-Support-Guiding-Principles-Using-Risk-Needs-Assessment-Information-Sentencing-Process.ashx>.

THERE IS AN OPPORTUNITY FOR THE EDUCATION SYSTEM ... TO EQUIP FUTURE CITIZENS WITH THE ABILITY TO THOUGHTFULLY NAVIGATE AN INCREASINGLY AUTOMATED AND DATAFIED WORLD.

So, what does everyone need to know about the embedding of increasingly complex technical systems into society? As Toby Walsh writes, for those who do not understand the fundamental principles of computation, “much technology will simply be magic”.¹⁵ For judges and legislators adopting the COMPAS tool in sentencing, being presented with a score representing risk would likely have felt like the magic solution to a challenging problem. If one sees a “risk assessment score” that is backed up by “research”, it is easy to support its use without asking what precisely has been demonstrated by research.¹⁶ We would hope that judges would ask more specific questions about the source of data (and potential systemic biases), choice of variables from which inferences are drawn, whether “research” includes measurement of differential impact on marginalised communities, and so forth. Choice of variables alone is not enough as, in fact, race is not an explicit variable in the COMPAS system but correlates with multiple input variables. But this is only likely if they have basic data literacy skills. Judges do not need a degree in data science; they do not need to design or build a machine learning tool. But they do need to be able to properly assess the applicability of a particular tool to the kinds of decisions that they make. Thus, while judges do not need to be expert,

they do need to know what questions to ask about these systems.

What is true for judges is equally true for people who go into different careers not normally associated with post-high school study of mathematics or computer science including police, business, welfare, and public policy. All of these domains are adopting or considering the adoption of tools such as predictive policing software, marketing analytics, automatic targeting tools, and data-driven policy modelling. Everyone will need to understand how data is analysed and feeds into different kinds of decision-making. Even if such people employ specialist statisticians or data scientists, they will need sufficient understanding in order to ask the right questions and interpret inferences in the domain of their specialist knowledge.

The same point applies to those who use systems as consumers or citizens. Everyone needs to be aware of how data about them is used to select information with which they are presented, fix prices, assess entitlements and detect non-compliance. There will always be an imbalance between how well companies like Facebook or Google understand their algorithm and the understanding of ordinary consumers, but we can and should reduce that gap. Again, we can encourage critical thinking about the right questions, such as how to avoid prejudging an issue based on a personalised information feed.

This basic training in skills needed by everyone is the work that schools have been doing for generations. While not everyone becomes a statistician, all students learn how to interpret a graph and how to recognise misleading ways in which information might be presented. For example, students learn about the impact of data collection methods on the conclusions that might be drawn. They also learn that starting the y-axis at a value other than zero

EVERYONE WILL NEED TO UNDERSTAND HOW DATA IS ANALYSED AND FEEDS INTO DIFFERENT KINDS OF DECISION-MAKING.

or drawing columns in three dimensions can alter perceptions of data. These skills remain crucial in assessing more advanced techniques such as machine learning - systemic biases in the data used to “train” the system will result in flawed inferences or predictions. For example, if members of one group within the population are more intensively policed, then police arrest data may not be representative of relative rates of actual offending. School curricula also cover misleading data displays, the impact of outliers, sampling (including inappropriate sampling) and statistical bias - all of which are crucial elements in understanding the circumstances in which inferences drawn from data may be misleading.

However, the standard curriculum assumes that when we are misled by representations of data, it is due to relatively standard statistical flaws and tricks. Advances in techniques have meant that misrepresentations of data have become more complex. Arguments about what statistic best represents fairness across racial groups in the context of COMPAS or STMP requires a deeper understanding not only of statistical inference but also of the intersections of this with philosophy and social science. Further, understanding how behaviour might feed into how each individual is represented to corporate and government bodies making human or automated decisions that affect that individual, or how such representations affect the availability of online information, could require exposure to some basic machine learning concepts.

In addition to understanding how data can be analysed to draw inferences, students need to link this with an understanding of automation and computing. Recognising the flaws identified in Robodebt requires an understanding of the implications of applying a standard formula, applicable in many cases, to a large group of people at scale. The limitations of repeated, unmonitored automation are also evident in stories like the existence of a book that was priced at \$23,698,655.93 on Amazon¹⁷ or the “flash crash” of the stock market in 2010 and 2018.¹⁸ One example raised earlier is where a particular population is flagged as “high risk” and targeted for closer law enforcement scrutiny, leading to more crimes committed by that population being recorded, feeding back into higher risk scores. Another example is where an automated system is used to select candidates to interview for a job category where there is existing bias. Such a program may recognise as desirable candidates that conform to existing demographic trends for that job category, leading to further embedding of historical biases. Exposure to such examples in high school will help inoculate the population against some of the risks of inappropriate applications of automation in a similar way to exposure to misleading graphs.

Most of the “bad” decisions made about applications of data-driven decision-making can be explained by a limited understanding of the purchasers and users of particular tools. More broadly, it is clear that public debate on the use of these tools by citizens is mired in poor understanding. Data literacy and computational thinking are core skills that citizens will need to understand and navigate a world where decisions are increasingly automated, often through processing of large volumes of data.

¹⁷ John D Sutter, ‘Amazon seller lists book at \$23,698,655.93 - plus shipping’, *CNN* (25 April 2011), available at <http://edition.cnn.com/2011/TECH/web/04/25/amazon.price.algorithm/index.html>.

¹⁸ Drew Harwell, ‘A down day on the markets? Analysts say blame the machines’, *Washington Post* (6 February 2018), available at <https://www.washingtonpost.com/news/the-switch/wp/2018/02/06/algorithms-just-made-a-couple-crazy-trading-days-that-much-crazier>.

Students can experiment with these issues themselves, particularly where they can use their own devices (laptops or mobile phones). For example, students can be directed to the same websites and observe whether different advertising is presented to them. Students can be encouraged to enter similar requests into search fields for different search engines (where relevant, with different settings) and observe differences in ranked and advertised results, including on different devices or in different browsers. An interesting variation of this exercise could encourage students to ask the same question in different languages or conduct an image search for a term such as “baby” and observe the relative lack of racial diversity compared to global human populations. The precise activities for which this will work will change over time as data practices evolve, so it could also be worth tracking results of this experiment (and changes in those results) over time and making historical data available to each new generation of students.

DISCIPLINARY EDUCATION

Disciplinary knowledge is the core foundation of our education system, at least from secondary school. Most of a student’s time is spent understanding core concepts in a variety of subjects including English, Mathematics, Science, and Humanities and Social Sciences (represented by History and Geography). This ensures students gain essential core skills that can help them understand the world around them, argue logically (including mathematically and rhetorically) and extract meaning from texts, images, graphs and diagrams.

The difficulty comes when all learning takes place within defined domains of knowledge. To

understand the problem across the education system, consider the very rough and simplistic image shown in figure 1. The white circle represents the total sum of human knowledge. In primary school and early secondary school, education looks like the blue circle in the middle – education that covers most domains of knowledge but does not (yet) extend out very far. As one progresses through high school, it looks more like the green layer – many domains are still covered (as required by the Australian Curriculum) but students start to specialise. By the HSC, for example, some students will no longer study chemistry. In university, one typically chooses a single discipline (shown in red), further specialising as one progresses through one’s studies. Finally, at the stage of a PhD (if a student gets that far), one goes all the way to the boundaries of knowledge on a particular topic, and marginally expands the realm of human knowledge.

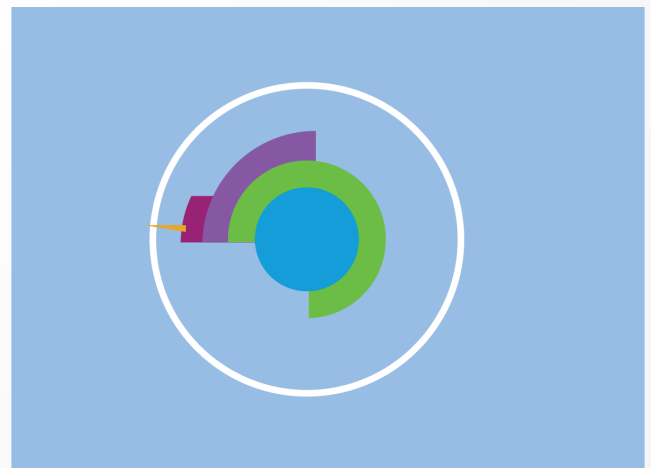


Figure 1: A conceptualisation of learning within domains of knowledge

The difficulty with this picture is twofold. The first is that many real problems require people to draw on more than one set of disciplinary knowledge.

As will be discussed below, understanding the impact of automation through data-driven artificial intelligence technologies on national politics or disadvantaged communities (for example) requires some understanding of data science and statistics, computer science, history and social science, and philosophy and ethics, among other ways of understanding and reasoning. Second, repeatable applications of clearly defined methodologies are exactly the sphere in which humans will struggle to compete with emerging artificial intelligence applications. For example, a computer program may be able to score well on an HSC mathematics examination, but we do not yet have systems that can apply ethical reasoning to diverse applications of machine learning and statistics. Students who can combine different ways of thinking and forms of reasoning will be in a better position to compete against the forces of automation.

I mentioned at the start that the model was simplistic. For example, the Australian curriculum includes general capabilities, such as critical and creative thinking, that are taught across different subjects. The ways in which this cross-disciplinary learning is made explicit to students may vary. For example, students may (or may not) have teachers explain to them how the same kinds of out of the box thinking and playing with different approaches that they use to develop creative projects in art can help in approaching an unfamiliar mathematics problem. One cross-curriculum initiative that is built into the Australian Curriculum is the idea of cross-curriculum priorities. These are three key areas that fit within rather than outside learning areas (being the specific disciplinary approaches taught). For example, the area of Sustainability “develops the knowledge, skills, values and world views necessary for people to act in ways that contribute to more sustainable patterns of living.”¹⁹ Some learning areas, like Science,

teach substantive concepts essential to this key idea (for example, learning OI.1, “The biosphere is a dynamic system providing conditions that sustain life on earth”). Specific learning areas provide background skills, as where English and Science provide investigative, analytic and communication skills to understand the problem, analyse proposed policies and arguments, and advocate for particular actions. English also supports students to analyse the language and rhetoric of policy and arguments in order to understand how they are positioned as an audience. However, the cross-curriculum initiative, while important, is still structured in a way that focuses on each separate learning area (or discipline) and its contribution to providing the kinds of knowledge, skills and approaches that students will need in this area.

Another difference from the above diagram is the idea of dual degrees at the tertiary level – students are not required to study only within a discipline but are encouraged (and, in the case of Law, often required) to study more than one discipline at tertiary level. While dual degrees are a fantastic way to broaden students’ education, they are not well supported by universities and take longer to complete (thus increasing costs). In particular, universities rarely offer structured opportunities for students to solve problems or tackle issues at the intersection between the disciplines they are studying. Rather, like cross-curriculum priorities, the different knowledges and skills from different disciplines are usually taught separately. Where students do activities or assessments at the intersections of their degrees, this is generally self-initiated (for example, bringing their knowledge from one discipline into an essay submitted for the other) and not always supported.

¹⁹ Australian Curriculum, Assessment and Reporting Authority, Sustainability, <https://www.australiancurriculum.edu.au/f-10-curriculum/cross-curriculum-priorities/sustainability/>

HOW SHOULD SOCIETY MAKE DECISIONS ABOUT APPROPRIATE TECHNOLOGIES GIVEN THE HIGH LEVELS OF EXPERTISE REQUIRED TO UNDERSTAND THE CHOICES INVOLVED?

THE ROLE OF APPLIED AND MULTI-DISCIPLINARY LEARNING

To truly live meaningfully and well in a world where decisions made affecting us are increasingly based on analysis of data held about us and others, we need more than data literacy and computational thinking. While STEM subjects are crucial, they alone are an insufficient basis for future citizens to be trained in order to ask the right questions and understand what is at stake.

Many of the questions presented above are ethical, not just as individuals but as a society (in which case they are also political). For example, is predictive accuracy the only metric for assessing a data-driven decision-support tool used on citizens, or do we need to consider differential impact on marginalised groups? What are the responsibilities of those designing, using or regulating particular tools – what questions need to be asked? What responsibilities do citizens engaging in political debate have to ensure they are sufficiently informed given the challenges of individually targeted search engines and social media feeds? How should society make decisions about appropriate technologies given the high levels of expertise required to understand the choices involved – should these be delegated to experts, can

we run (informed) citizen juries or similar, should politicians and other decision-makers be required to obtain the necessary expertise directly? How do we choose to engage with different technologies as humans, consumers and citizens?

Tools are always contextual and a good understanding of the contexts in which tools may be applied is essential to assessing the appropriateness of the tool. For example, the COMPAS tool is best assessed by someone who not only understands the way the tool uses data to make inferences, but also understands the criminal justice system and the various purposes of sentencing. While students cannot cover all or even most contexts in which data-driven decision-making operates, conversations about specific examples need to move beyond any single subject-area. Students need to bring together their statistical knowledge with their knowledge of particular domains such as history or legal studies in order to engage with the issues raised.

In other words, what is required is more than another cross-discipline topic introduced into the Australian Curriculum. Cross-discipline topics allow teaching of a common theme across the curriculum but do not envisage that all relevant disciplines would be brought into the same conversation. Understanding how to navigate the world of data-driven decision-making requires a contextual understanding of each application. For educational purposes, it requires students to consider the mathematical, computational, ethical, political and other questions raised simultaneously. What is required therefore is not so much new content to be added to the curriculum, but rather new interactions among teachers and students that stretch both beyond the confines of a single way of knowing to explore new problems from multiple perspectives.

²⁰ See Ruby Lohman, Navigating the AI Revolution, *Australia Unlimited* (19 April 2018), available at <https://www.australiaunlimited.com/society/genevieve-bell>.

WHAT IS REQUIRED
THEREFORE IS NOT SO MUCH
NEW CONTENT TO BE ADDED
TO THE CURRICULUM, BUT
RATHER NEW INTERACTIONS
AMONG TEACHERS AND
STUDENTS THAT STRETCH
BOTH BEYOND THE CONFINES
OF A SINGLE WAY OF
KNOWING TO EXPLORE NEW
PROBLEMS FROM MULTIPLE
PERSPECTIVES.

While one could create a new discipline to deal with any particular intersection (such as understanding data practices),²⁰ that is insufficient as a broader learning exercise. This is because there will be many fields in which students will need to cross disciplines to gain a holistic understanding of or develop solutions to a problem. The same interdisciplinary pedagogy can be applied to many complex problems. Consider sustainability, already a cross-discipline topic in the Australian curriculum. To decrease Australia's reliance on fossil fuels requires improved energy efficiency and electricity generation from renewable sources (science, technology), population distribution (geography), incentives and remove of dis-incentives (law) and political will to create such incentives (science, politics including observations from history), consumer behaviour (commerce), potential negative consequences of policy change (economics), and good decision-making (ethics). While aspects of these can and should be analysed in the context of the relevant

disciplines, students should also be given an opportunity to bring their knowledge together for a holistic analysis. This, after all, is what we expect from good policy makers and politicians, which is more likely if we have educated, thoughtful, critically engaged citizens.

STRUCTURING INTERDISCIPLINARY SPACES IN THE CURRICULUM

Of course, not all students will be on the same level for all disciplines that may be relevant in discussing a social problem. This mirrors the real world – where we do not know something ourselves, we work with those who have the relevant expertise. Groups can solve problems when individuals contribute disciplinary expertise to engage in conversations that cut across those disciplines. This has the additional benefit of turning students into teachers of their peers. A student studying economics, for example, can explain to their peers key concepts or facts relevant to a particular discussion. The group can then consider how this feeds into the broader discussion. In the context of data-driven decision-making, for example, an economics student might observe that Australia is often a buyer of products designed elsewhere, with little ability to influence the market.

How does one expand on interdisciplinary experiences in the high school curriculum? One option is to set aside time, either a regularly scheduled class or an occasional day, to explore a range of themes that cut across the curriculum, such as occurs when exploring themes around multiculturalism on Harmony Day. Another option, which would build on what some schools are already doing, could draw subject experts together to design an integrated model of teaching and assessment that cuts across traditional subject and

organisational divides. In either case, within this created space, students could focus on a specific issue drawing on their understanding of diverse disciplines. Where core disciplinary topics are covered, it makes sense for the interdisciplinary discussion to be scheduled after the disciplinary learning has occurred. For example, discussion of the discriminatory potential of data-driven decision-making should come after the relevant statistical and other discipline-based concepts have been explored.

Creating a space for interdisciplinary conversations requires a lot of teachers as well as of students. One way to do this would be to create a training program for specific teachers charged with implementing this program, where delegate teachers interact with their peers to confirm the timing of relevant topics. A better way would be to bring teachers together for cross-disciplinary planning, that ensures that topics and related assessments are structured in ways that allows for logical progression towards broader cross-disciplinary topics. Either way, support material on specific cross-disciplinary topics will need to be created and distributed to relevant teachers and additional training to those facilitating the sessions may be required. While students often engage in collaborative learning with their peers, the format here contemplates students explaining concepts in areas that their peers may not have studied themselves. Because the questions posed are complex, the point would not be to find a single solution, but rather to encourage critical, interdisciplinary thinking on questions they may face in future as citizens or consumers or in a professional capacity. In that sense, the proposal also supports existing learning in subjects such as HSIE and English.

WHILE ASPECTS OF THESE CAN AND SHOULD BE ANALYSED IN THE CONTEXT OF THE RELEVANT DISCIPLINES, STUDENTS SHOULD ALSO BE GIVEN AN OPPORTUNITY TO BRING THEIR KNOWLEDGE TOGETHER FOR A HOLISTIC ANALYSIS.

Because I teach in a Law school which requires all undergraduate students to study a dual degree, it is relatively easy to integrate different disciplinary knowledge into a discussion. This is often simply a question of recognising that there will be students familiar with, for example, particular technologies relevant to a discussion. I can encourage students, for example, to explain what genetic engineering is and how it works in a discussion about its regulation. While school students obviously do not have tertiary disciplinary knowledge, they can be encouraged to draw on what they are learning in one subject in the context of a discussion in another.

As an example of teaching across disciplines, I teach a course at UNSW Sydney called Designing Technology Solutions for Access to Justice. The course is available to law students, teaching them how to build an “expert system” which uses a pre-programmed logic to solve an access to justice challenge. Students learn computational thinking as well as how to use some of the available software products. As well as the content, it is always wonderful to see how their confidence increases over the semester as they realise they are able

²⁰ See Ruby Lohman, Navigating the AI Revolution, *Australia Unlimited* (19 April 2018), available at <https://www.australiaunlimited.com/society/genevieve-bell>.

to learn something initially unfamiliar. Students also draw on their legal knowledge, both in terms of understanding the law in a particular domain and their ability to understand client problems and organisational constraints. Bringing these capabilities together, they work in teams to build a system that can assist a not-for-profit organisation to provide targeted legal information or deliver legal services more efficiently. As well as core legal skills and learnt technical skills, each student in a team brings diverse capabilities relevant to solving a problem, including language skills, personal experience, and course content in fields such as design, business and computer science.

Assessment in the course includes presentations including a final presentation of the application built by the student team), the application built by each team (including its appropriateness and usefulness for the context in which it will be deployed), and contribution to the team (this is an individual mark based on (often diverse) intellectual and practical contributions to the group project and is assessed by themselves, their peers and myself).

CONCLUSION

My intention is not to propose a significant change to the way students are taught most of the time – that may be desirable but is an argument for another day. My proposal is relatively modest. First, I propose updating curricula on statistics and modelling to take account of more recent examples of data processing. In addition to basic statistics, students should be taught at a very general level about different ways in which data is analysed and used to draw inferences or make predictions. Examples could draw on student experiences with search

engines, social media and educational software analytics. Education about cyber safety could be supplemented with examples of how even ordinary consumer behaviour online can impact on how you are treated by government and the private sector, as where using a different operating system changes the prices at which goods are offered. Alongside broader computational thinking skills, students could be encouraged to think about examples where poor programming or implementation results in harm or error.

Beyond adjustments in what is taught in particular subjects, setting aside space, even if small, for problem-based interdisciplinary thinking on important contemporary issues is crucial. One example would be ethical considerations for data-driven inferencing and decision-making using an example like sentencing, policing or even national assessments of student capabilities. But there could be others including evaluating policy proposals for managing climate change or even imagining, designing and developing marketing for a more environmentally friendly product. The point is to encourage students to understand how disciplinary knowledge feeds in and combines to address important contemporary problems.

This is not only a question for schools, but also for universities. In Australia, students enrol in degree programs that build towards specialisation more than they encourage breadth of thinking. There are some,²¹ but few, opportunities for genuine interdisciplinary engagement, even where students enrol in a dual degree. While disciplinary expertise is important, in the real world, few people work in a context where the relevant skills are entirely disciplinary. More frequently, disciplinary skills are adapted to new contexts, as where general empirical skills are used to evaluate a project. Given this is the

²¹ See for example: 'Built Environment Interdisciplinary Learning (BEIL)', UNSW Sydney, <https://www.be.unsw.edu.au/future-students/interdisciplinary-learning/>; 'ENGG1200 - Undergraduate Special Projects', UNSW Sydney, <https://www.handbook.unsw.edu.au/undergraduate/courses/2019/ENGG1200/>; 'Interdisciplinary Projects', University of Sydney, <https://sydney.edu.au/study/study-options/undergraduate-courses/interdisciplinary-projects.html>.

case, should universities consider offering a program (as opposed to courses within a disciplinary program) that takes a problem-orientation rather than a discipline-orientation? Should we be encouraging *all students* to draw on diverse ways of knowing to solve a problem?

Large scale changes obviously require more consideration than has been possible here. It is worth bearing in mind that artificial intelligence as it is currently being built is good at single tasks (like playing Go or solving equations) and less good at tasks that require drawing domains together in creative ways. Creativity and cross-disciplinary skills will continue to be the domain of humans for longer than many of the more concrete, specific tasks. It is thus a fruitful domain in which to train future humans.

As an initial experiment, dedicating time for students to think broadly, purposefully and rigorously about how to navigate an increasingly automated and datafied world, drawing on diverse knowledge and skills, is not a bad place to start. After all, it will be today's students making future decisions that, collectively, can ensure thoughtful designs, appropriate applications and considered use of technology.

EXAMPLE OF PROPOSED ACTIVITY AND ASSESSMENT:

Risk assessment in criminal justice

As an example of an activity, students could spend a day discussing the application of data-driven risk assessment tools in the criminal justice system. Below I have sketched out background disciplinary content, pre-readings, guided discussion points, and possible assessment.

Disciplinary content (not all students need to have covered all of it, but it helps if some students have covered each point in class or through video or interactive material provided by experts):

- Statistical inference
- Historical or modern examples of racialized disadvantage or violence (such as the treatment of Indigenous Australians, the Holocaust, American slavery and the Civil War)
- Different kinds of ethical reasoning (in particular consequentialist versus values-based)
- Sentencing practices, including kinds of evidence used in sentencing in Australia

Pre-readings:

- ProPublica
- COMPAS questionnaire, available at <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html>.
- Sam Corbett-Davies, Emma Pierson, Avi Feller and Sharad Goel, 'A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear', *Washington Post*, 17 October 2016

Guided discussion:

- Why might a risk assessment tool be useful in sentencing? How does the use of this tool in the US differ from sentencing practices in Australia?
- What kinds of fairness are important in a society like Australia (both generally and in the specific context of sentencing)? What are our measures of a fair society?
- Take a look at the questionnaire that shows the variables used by the tool. What are some of the reasons why there might be more higher risk scores for African Americans compared to whites? Are there historical reasons why different groups might have different answers on a questionnaire such as this? What groups in Australia might have suffered similar historical disadvantage?
- What other issues exist with data collection that might distort the analysis in this example?
- Discussion of different ways that fairness might be measured - what is the difference between ProPublica's and Northpointe's approaches? Why is it mathematically impossible to ensure both kinds of fairness?
- Which measure of fairness should people be most concerned with, in particular from the perspective of (1) a judge using the tool, (2) a defendant being sentenced, (3) broader society?
- If we want to protect the public from harm while maintaining fairness in the criminal justice system, should there be a role for risk assessment tools? If so, what conditions should be put on their use?
- Could a tool such as this be designed in ways that (1) reduce unfairness, and (2) are more transparent to judges and defendants?

Assessment (any of the options below):

- Exposition arguing for or against the adoption of the COMPAS tool to assist judges in sentencing in Australia. The evaluation criteria could include standard writing skills in addition to critical thinking, ethical thinking, ability to draw on diverse material. For example, students would be expected to link mathematical/statistical analysis with historical explanations of bias.
- Verbal debate, with teams arguing for opposite propositions, along similar lines.
- Participation in class discussions, including willingness to bring personal experience or disciplinary knowledge into the broader debate.
- (Group project) Assume you have been asked to build a risk-based assessment tool to help judges predict how dangerous it might be to release a defendant into the community (on bail or parole). Looking at the COMPAS questionnaire, should any of these variables be excluded from analysis? How should the system be evaluated (for example, should there be a metric used to measure fairness)? What else might you build into the system's design? What warnings might you place on its outputs?

