Horizon Scanning Series

The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing

The Right to Privacy and Freedom from Surveillance

This input paper was prepared by Joy Liddicoat and Vanessa Blackwood

Suggested Citation

Liddicoat, J and Blackwood, V (2018). The Right to Privacy and Freedom from Surveillance. Input paper for the Horizon Scanning Project "The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing" on behalf of the Australian Council of Learned Academies, www.acola.org.

Privacy and Surveillance

Input paper prepared by Joy Liddicoat, Barrister and Solicitor, Wellington, New Zealand, and Vanessa Blackwood, Office of the Privacy Commissioner, Wellington, New Zealand at the request of the Working Group commissioned by the Australian Council of Learned Academics, the Commonwealth Science Council and the New Zealand Royal Society to produce an Horizon Scanning Report on Artificial Intelligence.

1 Regulation and the right to privacy

The theoretical and regulatory framework for the right to information privacy is well settled in New Zealand and Australia. In both countries, privacy laws regulate in a "technology neutral" manner, with standards for collection, use, storage and deletion of personal information applying regardless of the nature of technology that collects and uses personal information about individuals.

In general information privacy laws in Australia and New Zealand have stood the technological test of time. In an environment of rapid technological change questions about gaps inevitably arise, but before moving to fill these it is important to understand how current regulations apply. In New Zealand, there is no general legislative framework established to directly govern or regulate AI or algorithmic tools including automated decision-making (Edwards). Aspects of the current regulatory framework do apply to AI in New Zealand, including the information privacy principles of the Privacy Act 1993 and other human rights obligations which apply to private and State actions involving the personal information of individuals.

Caution is needed not to regulate too quickly, nor too late, and at the same time to keep pace with the field and emerging norms. In context of AI, there are challenges and opportunities for regulatory frameworks. Some of these are not new, as has been seen with the emergence of other new technologies. Lessons can be learned from experiences with transparency reporting and regulating copyright with regard to illegal file sharing online as well as the recently-proposed EU Copyright Directive.

All laws need regular review to ensure they are reflecting societal values and remain clear. In New Zealand, reform of the Privacy Act is underway, with a Privacy Bill introduced in early 2018. The Privacy Commissioner noted in his submission on that Bill that:

> "the [information privacy principles] do not directly – or arguably very effectively – address the particular risks and issues created by automated decision-making processes. Nor do they require specific mitigations such as algorithmic transparency."

In general, Australians and New Zealanders have good Internet and related technology uptake. To enable this for AI use, trust is a key issue: for AI to succeed in the private home sphere, individuals "need to know that their privacy is respected and maintained" (Kelly). New Zealanders are generally concerned with their individual privacy; two-thirds of respondents to the most recent privacy survey commissioned by the Privacy Commissioner declared concern about individual privacy, while more than half of New Zealanders are more concerned with their individual privacy now than they were in the last few years (OPC 2018).

An Office of the Australian Information Commissioner survey in 2017 found similar concerns, with 69% of Australians more concerned about their privacy than five years ago. Further, 83% believe there are more privacy risks dealing with an online organisation than an offline one, 79% are uncomfortable with a business sharing their personal information and 58% have decided not to share with an organisation because of privacy concerns (OAIC 2017).

Nevertheless, a recent New Zealand survey by Samsung found that despite concerns over data security, 38% of New Zealand respondents agreed they would feel more secure if they used smart technology to monitor their home and 61% considered the task they would find most appealing to be automated would be a security system that detects when they have left the house and sets alarms and locks. However, 62% were scared that their devices could be used without their knowledge (Parades).

2 Privacy and AI

AI tools may include algorithmic or automated decision-making and predictive analytics. These tools can enhance and improve services and productivity by increasing efficiencies compared to manual decision-making. These tools also risk causing significant interferences with privacy, and can perpetuate or worsen discrimination.

There is a tendency to treat algorithmic tools and AI as "a kind of twenty-first century Delphic oracle that seemingly makes unchallengeable and authoritative pronouncements divorced from human agency" (Cannataci). Assumptions that AI will be impartial and free from human fallibility such as racism, sexism or other institutional forms of bias ignore that AI is built by people, trained and implemented using historical data and involving policy decisions implemented by human actors (AI Now).

Researchers note that "bias in automated decision systems can arise as much from human choices on how to design or train the system as it can from human errors in judgment when interpreting or acting on the outputs" (AI Now). While many researchers point to this risk of implicit or explicit bias in algorithmic decision-making and machine learning, human actors can be just as if not more biased than the AI replacing them, and can replicate these structures of bias or discrimination both in training or programming the AI and also in making 'final decisions' involving human oversight.

The New Zealand Privacy Commissioner has identified two significant privacy risks from data analytics related to AI and automated decision-making: lack of transparency and meaningful accountability. The Commissioner notes that:

> "… systems may appear objective and yet be subject to in-built bias leading to discrimination. Many algorithmic assessment tools operate as 'black boxes' without transparency. This lack of transparency is compounded when private commercial interests claim trade secrecy over proprietary algorithms so that even the agencies using the tools may have little understanding over how they operate."

Accountability for decisions made using AI raises complexities as some decision-making techniques are more amenable to explanation than others. The result is an emerging field of 'explainable AI', where methods for explanation capability are being developed (AI Forum).

3 Surveillance

Multiple researchers evaluating how increased perception of surveillance might impact on people's behaviour have found that people alter the way they think and act even when faced with only the possibility of being under surveillance. This can include people avoiding talking or writing about sensitive or controversial issues, which not only has a "corrosive effect on intellectual curiosity and free speech" but inhibits the kind of democratic discussion necessary for a free society (Munn, 2016).

Recent research indicates that people may be less concerned about government surveillance stifling public criticism of government, and about governments gathering personal information in general. In the 2018 survey on individual privacy, for example, sixty-two percent of New Zealanders said they trust government organisations with their personal information, while only around a third of New Zealanders trusted private companies with that same information. In addition, public discourse on privacy and security led to significant reforms of intelligence laws in New Zealand. The Intelligence and Security Act 2017 contained the most significant reforms of intelligence agencies in New Zealand's history including increased transparency surveillance practices and the operation of intelligence agencies. The reforms may in part explain the greater levels of public comfort with government surveillance and the shift in public discourse from scrutiny of government actions to scrutiny of corporate information collection and surveillance.

The rise of increasingly invasive corporate data surveillance, including embedded tracking in computing and smart devices, raises new privacy and surveillance issues. In 2016 the Office of the Australian Information Commissioner and 24 other privacy enforcement authorities across the world evaluated 'Internet of Things' devices, finding that 71% of devices did not provide a privacy policy which adequately explained how personal information was being collected and managed (OAIC).

IoT devices which allow or facilitate the pervasive collection of personal information means private companies can increasingly use aggregate surveillance data to profile, predict, and manipulate customer behaviour. AI which supports this predictive analysis will increase the scope and availability of tools to evaluate and 'correct' individuals into their preferred course of action - which may be to increase profit and for the benefit of corporate interests rather than for a societal 'good'.

Private sector predictive data analytics also increasingly provide support for and are embedded into government agency functions, including law enforcement, healthcare, and public policy. In these situations, personal information collected with the coercive surveillance power of the state can be used to inform those privately developed analytical tools. Privacy experts warn that these new practices need to be monitored closely and, where appropriate, new ethics or regulatory practice developed.

4 Emerging regulation, ethics and regulatory practice

The rapid development of AI in diverse fields has prompted a range of regulatory and ethical responses. These can generally be separated into three broad areas:

- Frameworks and assessments useful during AI development;
- Tools and developments for ethical AI implementation and use; and
- Checks and balances once an AI is in use.

This section sets out examples of developments in four areas: algorithmic transparency, development of the right to erasure, algorithmic impact assessments, and new or emerging ethical standards.

Algorithmic transparency

Algorithmic transparency means having visibility over the inputs and decision-making processes of tools relying on algorithms, programming or AI, or being able to explain the rules and calculations use by AI if these are challenged.

Both the General Data Protection Regulation (GDPR) and the modernised Council of Europe Data Protection Convention 108 have legislated for aspects of algorithmic transparency. Article 13 of the GDPR, for example, imposes transparency obligations for automated decision-making, including profiling, and Convention 108 provides for the right to obtain, on request, knowledge of the reasoning underlying data processing.

The UK House of Commons Science and Technology Committee recommended transparency for government use of algorithms on the basis that the 'right to explanation' is a key part of accountability. The Committee recommended the default position be to publish explanations of the way algorithms work when the algorithms in question affect the rights and liberties of individuals.

The New Zealand Privacy Commissioner has recommended new measures be included in the Privacy Bill to better safeguard the interests of individuals, including a new privacy principle setting the high level expectations of fair practice and requiring algorithmic transparency in appropriate cases.

The right to erasure

The right to erasure is provided for to a certain extent through the GDPR and Convention 108. The New Zealand Privacy Commissioner recommended a new privacy principle on the right to erasure of personal information, recognising that:

> "the current rights and protections available to New Zealanders are gradually weakening as technology develops. In particular, the requirement in principle 9 for information to be kept no longer than is necessary is rendered meaningless in the context of advanced algorithms and artificial intelligence. For example, the thirst of artificial intelligence systems for data will mean that agencies will want to keep all of the data that is available for increasing periods of time."

Providing individuals with a right to erasure shifts the decision-making onus from agencies, who are incentivised to collect and retain information, to individuals who can then exert control over their own information.

The right to erasure raises issues in the context of the development of AI systems using individual information for machine learning and algorithmic development and training. It remains unclear whether the right to erasure, or the related right to data portability, will create obligations on an AI developer to delete personal information from the AI training database or to what extent the intellectual property in the AI is linked to or reliant on that personal information.

Algorithmic Impact Assessment

In both Australia and New Zealand a key tool for identifying and managing privacy risks is the privacy impact assessment. Building on this work, AI researchers have developed "practical framework" for an Algorithmic Impact Assessment (AIA), similar to impact assessment frameworks already used in data protection, privacy, and human rights policy domains. They note that "AIAs will not solve all of the problems that automated decision systems might raise, but they do provide an important mechanism to inform the public and to engage policymakers and researchers in productive conversation" (AI Now).

AI Stocktakes

The United Kingdom House of Commons Science and Technology Committee report on algorithms in decision-making contains recommendations to ensure oversight of machine

learning-driven algorithms, including producing, publishing and maintaining a list of where algorithms with significant impacts are being used within central government.

Similar work is being done in New Zealand, with a stocktake of algorithms in the public sector announced in 2018 by Government Ministers Clare Curran (Minister for Government Digital Services) and James Shaw (Minister of Statistics). The first phase of this assessment is underway and will focus on operational algorithms that result in, or materially inform, decisions which impact significantly on individuals or groups.

New Ethical Issues and Emerging Codes

There has been a surge in the creation of ethical codes or calls for such codes in the last five years. In the United Kingdom, for example, the House of Lords recommended the government introduce a statutory code of practice for the use of personal information in political campaigns, applicable to political parties and campaigns, online platforms, analytics organisations and others engaged with such processes. The Committee also announced it would produce draft guidance quickly in order for the code to be "fully operational" before the next UK general election.

However, while a variety of ethical standards are being developed, these do not appear related to each other. This may give rise to more difficulties if, for example, different ethical standards are applied to the same technology across its application or development in different sectors. In addition, some have called for these ethical codes of conduct and principles to be more closely tied to the everyday practice of AI design and development (AI Now).

In New Zealand, the AI Forum has noted there are no laws requiring AI developers to design a system so that it can explain its decisions, nor any clear guidelines on when AI systems should transfer control back to humans to prevent harm. The Forum has called for new ethical discussions that include rights, duties, conflicting values and other factors that may need to be taken into account in the particular context. The Forum recommended a working group be established to advocate for and provide expertise in applying principle based ethics to AI to assist end-user companies, government and not for profit organisations. The Forum also recommended sector specific regulators develop understanding of how AI applies to their fields.

As a practical step, in New Zealand the Privacy Commissioner and the Government Chief Data Steward have jointly developed six draft principles to support safe and effective data analytics, including algorithmic decision-making, across the public sector. These principles include: that the use must deliver clear public benefit; maintain transparency; have well understood limitations; and retain human oversight. These principles are intended to underpin public sector work involving data analytics, and to inform further development of guidance to support government agencies.

Whichever ethical framework is developed, experts emphasise that it is important to link ethical standards to strong oversight and accountability mechanisms and to ensure these are multi-stakeholder (Access Now and Amnesty International). The *Toronto Declaration* is one example of a multi-stakeholder statement on the human rights approach to machine learning systems, including AI. The Declaration signatories emphasise that while the ethics discourse is gaining ground, ethics cannot replace the centrality of universal, binding and actionable human rights law and standards, which exist within a well-developed framework for remedies for harms from human rights violations (Access Now and Amnesty International).

Conclusion

Misinformation about the functioning of AI tools, including algorithmic decision-making and predictive analytics, can significantly shape public discourse and understanding. Whether or not media reporting on AI tools intends to increase public unease or distrust, this shaping of the discourse results in increased difficulty explaining analytics and reaching a level of public understanding. Nevertheless the general public is engaged in the debate about privacy interests and AI and their wider democratic right to know when and how AI tools are being used. More work is needed to ensure that personal information laws are able to secure the right to privacy in the context of AI and to engage in the developing ethical standards.

References

Access Now and others *The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning system* 16 May 2018, San Francisco, United States of America.

Australian Community Attitudes to Privacy Study (2017) Office of the Australian Information Commissioner, Sydney, Australia.

Privacy shortcomings of Internet of Things businesses revealed (2016) Office of the Australian Information Commissioner, Sydney, Australia.

Cannataci, J *Report of the Special Rapporteur on the right to privacy* (2017) 72nd session of the United Nations General Assembly.

Chambers, C "The psychology of mass government surveillance: How do the public respond and is it changing our behaviour?" The Guardian, 18 March 2015 https://www.theguardian.com/science/head-quarters/2015/mar/18/the-psychology-of-mass-government-surveillance-how-do-the-public-respond-and-is-it-changing-our-behaviour

Edwards, J "Submission to the Justice and Law Select Committee on the Privacy Bill" (May 2018), Office of the Privacy Commissioner, New Zealand.

Gal, U "Data surveillance is all around us, and it's going to change our behaviour" The Conversation, 11 October 2016 http://theconversation.com/data-surveillance-is-all-around-us-and-its-going-to-change-our-behaviour-65323

House of Lords *AI in the UK: ready, willing and able?* Report of the House of Lords Committee on Artificial Intelligence, available at https://social.shorthand.com/LordsAICom/32KXpihQLj/ai-in-the-uk

Human Rights Commission *Privacy, Data and Technology: Human Rights Challenges in the Digital Age* Issues Paper (2018) Wellington, New Zealand.

Munn, N "How mass surveillance harms societies and individuals – and what you can do about it" Canadian Journalists for Free Expression, 8 November 2016 https://www.cjfe.org/how_mass_surveillance_harms_societies_and_individuals_and_what_you_can_do_about_it

Office of the Australian Information Commissioner "Submission on the Consultation Paper: The digital economy, opening up the conversation" (2017), Sydney, Australia.

Paredes, D "AI welcome in NZ homes, but privacy remains prime concern: survey" CIO New Zealand, 9 July 2018 https://www.cio.co.nz/article/643508/ai-welcome-nz-homes-privacy-remains-prime-concern-survey/

Reisman d, Schultz J, Crawford K and ors, *Algorithmic Impact Assessments: a practical framework for public agency accountability* (2018) AI Now Institute, New York University, United States of America.

The Artificial Intelligence Forum of New Zealand *Artificial Intelligence Shaping a Future New Zealand* (2018), Wellington, New Zealand.

Access Now and Amnesty International *The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning system* 16 May 2018, San Francisco, United States of America.

World Internet Project International Report 6th edition “The Internet in Australia” and “The Internet in New Zealand (2018), World Internet Project, United States of America.