Speaker:	**Privacy Commissioner John Edwards**
Organisation:	**SAS Users of NZ**
Date/Place:	**8 May 2019, Wellington**

Privacy 2020 - Understanding what's changing in data regulation and analytics

TORTURE THE DATA

- We've been thinking a lot about the risks that might be involved in extracting public or private value from data.
- Increasingly, CEOs, Ministers are asking "can we automate this? What can we learn from this dataset that will inform policy or business strategy?"
- The pressure to look to technology to provide answers to complex social problems is increasing, and is supported by consultancies, data scientists and software vendors.
- But as renowned British economist Ronald Coase warned: "If you torture the data long enough, it will confess to anything."

ALGORITHMS

- In 2014, Wired magazine warned '*Algorithms are great - but they can also ruin lives.'*
- The article pointed out that an algorithm may falsely profile an individual as a terrorist.
- This is something which in 2014 confronted about 1,500 unlucky airline travellers in the US each week.
- As the computer security expert Bruce Schneier noted:
  - *Finding terrorism plots is a needle-in-a-haystack problem and throwing more hay on the pile doesn't make that problem any easier.*
- Predictive risk modelling, algorithm-based decision making, machine learning techniques have been called "weapons of math destruction" (US mathematician Cathy O'Neil).
- Luke Dormehl (*The Formula: How Algorithms Solve All Our Problems and Create More*) echoed a similar line:
  - *A single human showing explicit bias can only ever affect a finite number of people.*

- *An algorithm, on the other hand, has the potential to impact the lives of exponentially more.*

BLACK BOXES

- Many algorithmic assessment tools operate as 'black boxes' with a lack of transparency or understanding over how they operate.
- This can lead to situations where decision-makers make bad decisions, and those subject to the decisions cannot appeal them – because of commercial sensitivity.
- Lack of transparency is compounded when private commercial developers claim trade secrecy over their proprietary algorithms.

SIX PRINCIPLES

- Principles supporting the safe and effective data and analytics
- Jointly developed by the Chief Government Data Steward and the Privacy Commissioner.

1. **Deliver a clear public benefit**
   - Take for example, the unconvincing use of facial recognition technology by the London Metropolitan Police.
   - According to information released under Britain's Freedom of Information laws, the technology gave a 98 percent false positive rate.
   - US research has also showed facial recognition technology is particularly inaccurate identifying minority ethnic women.
   - The study by MIT Media Lab researcher Joy Buolamwini on bias in facial recognition software showed gender was misidentified in up to 35 percent of darker skinned females.

2. **Ensure data is fit for purpose**
   - Algorithms can inherit biases. For example, data from a criminal justice system often involves elements of historic or systemic racism.
   - If you have a dataset of a "favoured" group of people and a "discriminated" group of people, you're deciding on an outcome that has historically been awarded to the favoured group more often.
   - The more accurately you rely on the historical data, the more the outcomes will discriminate against the disadvantaged group.

- The history of heightened scrutiny of black neighbourhoods in the US by police in what was known as broken windows policing made black people more likely to be arrested for a given crime.

3. Focus on people

- MSD individual client level data policy – my office gave the policy a fail.
- Social service providers that relied on government funding required to hand over information about individual clients.
- The policy could deter people from seeking support or assistance, making the them 'invisible' to policy makers and government – the opposite of a focus on people approach.
- MSD has since prioritised a piece of work called its *Privacy, Human Rights and Ethics Framework*.
- Application of this framework will ensure that any possible future operational predictive risk modelling carried out by MSD complies with the Privacy Act 1993 - and balances privacy rights with other objectives.

4. Maintain transparency

- Many algorithmic assessment tools operate as 'black boxes' with a lack of transparency or understanding over how they operate.
- This can lead to situations where decision-makers make bad decisions, and those subject to the decisions cannot appeal the decision.
- One such case is Compas - an algorithm developed by a US company Northpointe - which calculates the likelihood of someone reoffending.
- Electronic Privacy Information Centre President Marc Rotenberg has said "knowledge of the algorithm is a fundamental human right".
- Predictive risk modelling - Immigration New Zealand confirmed earlier this year it had scrapped data and predictive risk modelling work.
- The High-Harm pilot model was implemented in July 2014 with a focus on targeting over-stayers – leading to critics to claim it was a form of racial profiling.
- My office has said we will work with Immigration NZ if it developed technology or a similar initiative in future to ensure it was fit for purpose.
- Organisations can also apply for an OPC Privacy Trust Mark for a particular service or product – such as DIA's RealMe identity verification service.

5. **Understand the limitations**
    - From Cathy's book Weapons of Math Destruction - data scientists, like doctors, should pledge an equivalent of the Hippocratic Oath - one that focuses on the possible misuses and misinterpretations of their data models.
    - Two financial engineers, Emanual Derman and Paul Wilmott, drew up one such oath in the aftermath of the 2008 global financial crisis. It begins: *I will remember that I didn't make the world, and it doesn't satisfy my equations.*
    - And concludes: *I understand that my work may have enormous effects on society and the economy, many of them beyond my comprehension.*

6. **Retain human oversight**
    - GDPR Article 22 (1): "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her".
    - Article 22 (3): "The data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision".
    - These protections are significant in the context of international benchmark setting. They are an influential signpost to future regulatory settings on automated decision making for greater transparency.

DE-IDENTIFICATION/RE-IDENTIFICATION

- In 2016, Australia's Department of Health published the de-identified health data of 2.9 million people online.
- The data came from two medical benefits schemes and contained 1 billion lines of historical health data from 10 percent of the country's population.
- The de-identified data was published on the government's open data website for research purposes.
- But one month later, University of Melbourne researchers revealed the data could be re-identified.
- One of the researchers, Dr Vanessa Teague, looked for herself in the data.
- In her year of birth, more than 17,000 women in the dataset matched her.
- When the years of birth of two of her children were added, only 59 possible matches remained.

- Only 23 of these were in her home state of Victoria.
- When she added the specific days of her children's birth, there were no other matches – all that remained was her record.
- This 'anonymised' data had now unlocked her every interaction with the public health service.
- The researchers also matched unique anonymised patient records to seven well known Australians, including three former or current members of parliament and a top footballer.
- After a lengthy investigation, Australia's Privacy Commissioner concluded the department had breached the country's Privacy Act three times.
- The de-identification was bound to fail because it was trying to achieve two inconsistent aims: the protection of individual privacy and publication of detailed individual records.
- Last year, the Australian government announced plans to amend the Privacy Act to criminalise re-identification of published government data.

INTEGRATED DATA INFRASTRUCTURE

- Setting a high bar in the safe use of data.
- New Zealand's Integrated Data Infrastructure has eight broad categories – health, education, social services, justice, communities, population, income and work, housing.
- The IDI uses the 'Five Safes Framework'.
    - Safe people – researchers are vetted
    - Safe projects – researchers must demonstrate their project is in the public interest
    - Safe settings – privacy and security arrangements keep data safe
    - Safe data – identity is protected
    - Safe output – all information is checked to ensure it does not contain any identifying results.

PRIVACY BY DESIGN PRINCIPLES

- Proactive, not reactive – build in strong privacy controls to *prevent* issues rather than *fix* issues.
- Privacy as the default setting – build systems that automatically mitigate privacy risks, by reducing the amount of information you collect and limiting the circumstances where you disclose it.

- Privacy embedded in design – build privacy as an integral part of a system, rather than bolting it on after the fact.
- Full functionality – avoid false dichotomies, such as a "trade-off" of privacy against security. Full functionality allows you to build in privacy without trading off other, positive values. It's not a zero-sum game.
- End-to-end security – build in security at every point of the information life cycle. It's not sufficient to have strong security when you *hold* information but lax security when you *use* information (as I'll talk through shortly). Security needs to be part of every interaction with personal information.
- Visibility and transparency – let people know what information you have, how you will use it, and what your policies and procedures are around personal information.
- Respect for user privacy – keep the users at the centre of the relationship by ensuring you have their consent, you give them access to their information, and you let them correct information.

PRIVACY IMPACT ASSESSMENTS

- An essential tool for identifying potential risks.
- PIA are used to identify and assess the privacy risks arising from the collection, use or handling of personal information.
- A PIA will also propose ways to mitigate or minimise these risks.
- A PIA can be particularly useful when an agency is considering introducing a new policy or operating system, or when making changes to an existing process.

PRIVACY ENHANCING TECHNOLOGY

- Data minimisation
- Encryption to protect against hacking or data loss.
- Artificial intelligence to enhance privacy – e.g. surveillance cameras that record only when necessary.
- Automated systems that reduce opportunities for employee browsing.
- Use of algorithms to block unnecessary data collection.

PRIVACY BILL

- Mandatory data breach notification – an agency must notify my office of privacy breaches (defined as unauthorised or accidental access to, or disclosure of, personal information) that pose a risk of harm to people, and to affected individuals.

- Compliance notices – I will be able to issue compliance notices that require an agency to do something, or stop doing something, in order to comply with the law.
- New criminal offences – it will be an offence to mislead an agency in a way that affects someone else's information and to knowingly destroy documents containing personal information where a request has been made for it. The penalty is a fine up to $10,000.
- Binding decisions on access requests – and I will be able to make decisions on complaints relating to access to information, rather than the Human Rights Review Tribunal.
- I have recommended the Privacy Bill include a new principle to limit the harms arising from automated decision-making and to require "algorithmic transparency" in appropriate cases.

ORGANISATION CHECKLIST

- Organisational culture and awareness of good privacy practice:
    - an understanding of the way your organisation uses personal information
    - an understanding of the 'information lifecycle'
    - an appreciation of typical areas of legal risk
    - a willingness to consider mitigating strategies and effective remedies when things go wrong – do you have a data breach response plan?
- Adopt Privacy by Design principles wherever possible
- Carry out Privacy Impact Assessments
- Sensible, clear policies and privacy statements.
- Engaged privacy officers.
- A responsive and efficient complaints process.
- Active engagement with my office.
- Ensure privacy and information governance get a seat at the top table in management.
- Training! Training! Training!

OTHER WAYS WE CAN HELP

- Online tools to help agencies/individuals manage privacy requirements
    - Privacy ABC and other free online privacy training modules
    - Guidance on website
    - AskUs

AskUs

- OPC receives about 8,000 general enquiries and 300 media enquiries each year.
- We created AskUs as a search engine to answer privacy questions more efficiently.
- AskUs received nearly 19,000 enquiries last financial year
- Some examples of frequently asked questions include:
- Is my neighbour allowed to film our property with a security camera?
- Can I ask for information about a deceased relative?
- Are there any rules about where CCTV can be used?
- What are the rules for flying drones?
- Can I record someone without telling them?

IN CONCLUSION

- Privacy issues surround us every day - Facebook, drones, cameras, online shopping, smartphones, mobile apps, marketing cold calls, spam emails etc.
- There's a historic change coming - a new Privacy Act is on its way.
- It will better protect the privacy rights as individuals.
- It will give agencies clearer responsibilities and obligations.
- It will give my office greater enforcement powers.