14th June 2024

Privacy Amendment Bill

Privacy Foundation New Zealand Submission

Summary

1. Privacy Foundation New Zealand supports the intention of the Privacy Amendment Bill (the Bill) to better align Privacy Act 2020 with the European General Data Protection Regulation (the GDPR) in the area of indirect collection of personal information.
2. We understand that the reason for the Bill is to uphold the privileged status of our privacy laws as ensuring an adequate level of protection to the GDPR ('adequacy') which brings a number of economic and non-economic advantages to New Zealand.
3. Because of that need for adequacy, we propose to include additionally in the Bill the already existing under the GDPR protection against surreptitious indirect collection of personal information by means of web scraping.
4. We recognize that web scraping enables the creation of detailed profiles of large number of New Zealanders by harvesting their publicly accessible data like social media profiles, which can be used for a number of activities that create significant harms. Those activities are, for example: Artificial Intelligence-enabled deep fakes, scams and security intrusions based on stolen identity and mass surveillance tools enabling instant facial recognition (like Clearview AI).
5. We invite Parliament to enact law that both secures New Zealand adequacy and protects New Zealanders and New Zealand public and private sector institutions from those harms.
6. To achieve this, we invite the Committee to use the elements of the legislative proposal prepared by the Office of the Privacy Commissioner included in the Exposure Draft of the proposed Biometric Processing Privacy Code (April 2024), that are designed exactly to achieve that purpose.
7. We believe that such protection should be available as soon as possible for all personal information and enacting these changes is urgent in light of rising numbers of scams and security intrusions.

Detailed analysis

8. We support the intent of this bill. It better aligns the Privacy Act 2020 (the Act) in New Zealand with international best practice as epitomised by the European Privacy Regulation (Regulation 2016/679 on the General Data Protection Regulation [2016] OJ L119/1) (GDPR) Art 14 which requires individuals to be given certain information when information about them has been obtained indirectly. It will also align New Zealand law with the legislation in Australia, a country where many of New Zealand's businesses operate and which is a home of many New Zealand investors. Australian Privacy Principle (APP) 5 requires notification when personal information "about" an individual is collected, unlike current Information Privacy Principle (IPP) 3(1) in New Zealand which requires notification only when information is collected from individuals. The *Australian Privacy Principles guidelines* (2015) issued by The Office of the Australian Information Commissioner (OAIC), at [5.5] specifies that the notification requirement applies to all personal information collected about an individual either directly from the individual or from a third party.
9. We generally support all the clauses of the bill including the technical amendments in Part 2. However, we wish to address the remainder of this submission to clause 4 of Part 1, specifically the wording of the proposed new IPP 3A. Whilst we agree with the clause as worded, we think it does not go far enough in one respect and can be improved as outlined below:
10. Clause 4 adds a new IPP 3A to the Act requiring agencies to be transparent when they are collecting personal information from sources other than the individual concerned. However, IPP 3A(4) retains existing exceptions contained in IPP 2(2) that allow non-compliance for a range of circumstances including where it would not prejudice the interests of the individual, that compliance would prejudice the purpose of the collection, for law enforcement, health and safety and for other reasons. One such exception (IPP 3A(4)(b)) replicates that contained in IPP 2(2)(d) where "information is publicly available information". This term is defined in section 7 of the Act to mean "information that is contained in a publicly available publication" which in turn is defined as meaning a "publication… in printed or electronic form that is, or will be, generally available to members of the public free of charge or on payment of a fee". This is generally understood to include public registers as well as websites. We do not oppose the exception for publicly available information generally. However, we think that with the advent of digital technologies the mischief that has arisen is that of automated collection and processing of publicly available information on a large scale. This is known as web scraping. We therefore submit **that the exception for publicly available information should not be permitted to apply to indirect collection of personal information through web scraping.**
11. We note that Rule 2(3) contained in the Office of the Privacy Commissioner's Exposure Draft of Biometric Processing Privacy Code (April 2024) stipulates that the exception for publicly available information does not permit the collection of biometric samples by means of web scraping. Web scraping is further defined to mean "using automated tools to extract biometric information from publicly available online sources including websites and social media platforms". Although biometric techniques, such as facial recognition, undoubtedly pose greater risks for this type of

collection and processing, we believe the same risks arise for automated techniques of mass harvesting of other types of personal information.

12. Automatic collection of personal information from publicly accessible sources, like social media, enables the creation of detailed profiles of large number of New Zealanders, who might have posted much of that information by themselves, but they had not intended it to be automatically collected. Such detailed profiles could be further used to target individuals in many different ways which create significant harms. Many of the those harms stemming from use of artificial intelligence (AI) technologies, including generative AI, stem from the mass collection of personal information and its potential misuse through algorithmic processing. Those activities are, for example: Artificial Intelligence-enabled deep fakes, scams and security intrusions based on stolen identity.
13. New Zealand has already witnessed a huge increase of those harms. According to Consumer NZ (https://campaigns.consumer.org.nz/stamp-out-scams) in New Zealand, approximately:
    - $200 million was stolen from scam victims in 2023
    - 1 million households were targeted by scammers in the past year,
    - 185,000 households were scammed out of money in the past year,
    - 55,000 New Zealanders were repeat victims of fraud and cybercrime.

    Web scraping also enables mass surveillance tools enabling instant facial recognition (like Clearview AI).
14. We do not oppose indirect collection of publicly available information through manual means in individual instances. For instance, a business intending to enter a contract with an individual might wish to verify their identity by searching a public register, LinkedIn profile or even the website of their purported employer. Indeed, several of the other exceptions contained within IPP 3A may well also apply in these circumstances. On the other hand, we do not think the same should apply to commercial harvesters of this type of information and its subsequent use by data brokers and the like. If it is allowed, there should at least be a requirement for transparency through the new notification requirements of IPP 3A.
15. The risks posed using publicly available information may not be readily apparent. There is an implicit social license whenever such information is made available. When it is collected on a large scale through automated means and sophisticated technologies employed that allow profiling and linking of such information the results are qualitatively different from the rationales underpinning original publication of such information. Individuals cannot be assumed to have consented to these secondary purposes in respect of their information made publicly available.
16. In New Zealand, these risks are compounded by the fact that the IPPs do not differentiate ordinary personal information from sensitive personal information unlike the position in Australia. There, specific consent is needed for collection and use of sensitive information, including biometric data (see APP 3(a)). The lack of such consent has underpinned investigations conducted by privacy authorities there (see for example *Commissioner initiated investigation into Clearview AI, Inc (Privacy)* [2021] AICmr 54 (14 October 2021)).
17. Clearview has also been the subject of litigation in the United Kingdom, following enforcement action by the Information Commissioner (ICO) there (ICO fines facial

recognition database company Clearview AI Inc more than £7.5m and orders UK data to be deleted | ICO), and in the United States itself, pursuant to state legislation. Although there are substantial differences between the operating legislation in the United Kingdom (UK GDPR) and New Zealand a point of commonality is the ICO's concern as to the lack of transparency to the individuals whose images were scraped - a key goal also of the Privacy Amendment Bill. The onus of transparency should fall on the companies concerned and not the individuals. Another similarity is the requirement that information collection be for "lawful purpose" (IPP 1) and the manner of collection through "lawful means". Commentators have pointed out that data scraping often occurs in contravention of intellectual property rights of the platforms that are targeted as well as of their terms and conditions thereby contravening contract law. (See Slaughter and May "Data Scraping and Compliance – No Clearview – (Yet)?" https://my.slaughterandmay.com/insights/briefings/data-scraping-and-compliance-no-clearview-yet). Information privacy laws such as the Act as well as the GDPR empower individuals through automatically deeming these other legal contraventions to also trigger privacy rights.

18. Publicly available information often reveals sensitive information about individuals, including their sexual orientation, political views and health information. When harvested on a large scale such information can be used to create a detailed profile of the individuals and target them, exploiting their vulnerabilities, exploiting their businesses, friends and families in a plethora of different scam schemas, for political purposes, and other potentially harmful reasons. We believe these risks go further than the mischiefs addressed in the proposed Biometric Processing Privacy Code that has been referred to above and that the sudden increase of those risks mandates much stronger and earlier legislative action. Transparency surrounding the collection of such information would go towards mitigating these risks.
19. Finally, we note that privacy authorities around the globe have cautioned against the risks of web scraping generally see *Joint statement on data scraping and the protection of Privacy"* dated August 24, 2023. The statement was made by 12 Data Protection Authorities across the globe. (Office of the Privacy Commissioner | New Zealand part of global effort on data scraping). The Netherland's privacy authority has highlighted that under GDPR, scraping of websites is prohibited even where personal information appears to be publicly accessible although there are permitted exceptions, such as for law enforcement purposes (Netherlands' DPA issues guidance against web scraping | IAPP).
20. **To conclude: we support the addition of the following subclause to IPP 3A:** "**(5) Subclause 4 (4)(b) does not permit noncompliance with subclause (1) if the personal information is collected by means of web scraping". The term "web scraping" could either be added to the section 7 definitions in the Act or a further subclause added as follows: "(6) for the purposes of IPP 3A web scraping means using automated tools to extract personal information from publicly available online sources including websites and social media platforms."**

This submission was prepared by Privacy Foundation New Zealand Chair Dr Marcin Betkier, Committee Member Associate Professor Gehan Gunasekara, and Member Maria Tenorio.