CANON Canadian Anonymization Network

Charles Taillefer
Director, Privacy and Data Protection Directorate
Innovation, Science and Economic Development Canada
charles.taillefer@canada.ca

October 15, 2019

Submission re: ISED's "*Strengthening Privacy for the Digital Age*"

Dear Mr. Taillefer,

Thank you for the opportunity to provide comments on Innovation, Science and Economic Development (ISED) Canada's proposals to modernize the *Personal Information Protection and Electronic Documents Act (PIPEDA)*, as described in *Strengthening Privacy for the Digital Age* (the "White Paper").

We are writing on behalf of CANON, the Canadian Anonymization Network. Launched in May 2019, CANON is a not-for-profit organization whose members are comprised of large data custodians from across the public, private, and health sectors. The Network's primary purpose is to promote de-identification and anonymization in Canada as privacy-respectful means of supporting innovation and leveraging data for socially and economically beneficial purposes.

This submission is not intended to state a formal policy position on behalf of CANON or its members in respect of the entire White Paper. Rather, the present submission is limited to comments and questions raised about de-identification issues more specifically.

The members of CANON are encouraged to see that ISED has included a thoughtful discussion of the potential benefits of de-identification in its White Paper. In brief, we would like to voice our mutual interest in continuing to advance the policy debate on the important virtues of de-identification as a privacy-protective measure and offer some suggestions and points of clarity. To that end, we would be pleased to make ourselves available for any follow-up questions or discussion of the points raised below, or any other matter for which the collective experience and/or expertise of CANON members may prove to be helpful.

We have divided our submission into four main areas for consideration: defining terms; adopting a risk-based approach; clarifying the role of consent in relation to de-identification; and promoting codes of practice.

Defining Terms

In the White Paper, ISED proposes to add a definition of "de-identified information" to PIPEDA. We believe that defining "de-identified information" and related terms, such as "anonymized information", "pseudonymized information" and "aggregate information", will be critically important to clarify these concepts and ensure common meanings that are well understood by all affected stakeholders vis-à-vis one another. This includes organizations vis-à-vis their customers when describing their personal information management practices; controllers vis-à-vis processors when negotiating terms and conditions of a service agreement; and, organizations vis-à-vis regulators when demonstrating their accountability and data governance frameworks in the context of an investigation or audit. To this end, we offer the following suggestions, to the extent they may be helpful:

- **Consistency:** We believe that when defining terms, there is great benefit to looking broadly at other laws to ensure clarity and consistency of terms across Canada as well as globally. Recognizing the reality that significant amounts of data flows (and will continue to flow) between sectors and across borders, it would be in Canada's interest to develop definitions that are consistent, or at least interoperable, with those of other laws and other jurisdictions. This consistency would provide the level of predictability needed for organizations to design their current data practices, and to invest in new innovative ideas, with some level of assurance that they are grounded on well-established concepts. As one of its early deliverables, CANON has set out to develop a standard lexicon of terms which we would be pleased to share with ISED officials and publicly to the extent it may be helpful in facilitating common understanding.

- **Contextualization:** It has become clear, based on the collective experience of CANON members over many years, that the binary concept of personal information is no longer fit for purpose. As alluded to by ISED, complete anonymization for which there is virtually no risk of identifying an individual is becoming practically unattainable. For its part, de-identification is a relative concept that requires a contextual evaluation. It is therefore important that any definitions allow for consideration of contextual factors such as: the nature of the data involved; the reasonable expectations of potentially affected individual(s); the intended purposes for its use; the release environment; the availability of other linkable data; the likely incentives to re-identify the data; the costs and level of expertise required to re-identify data; and, the potential harm to individuals should an individual be re-identified.[1] Also to this end, CANON has undertaken to develop a series of different use cases to help demonstrate the instrumental importance of contextualization.

[1] See, for example, the "Five Safes Framework". http://www.fivesafes.org

- **Flexibility**: We agree with ISED that among PIPEDA's strengths is the fact that it is a flexible, principle-based and technology-neutral law. Thus, whatever definitions are used, we would recommend that PIPEDA retain these well-recognized properties of the current law. By not setting out highly specific criteria or prescriptive processes for how personal information is to be de-identified, but rather, focusing on the end state that must be achieved, organizations will be permitted to develop, and continually improve upon, robust and accountable processes using techniques which are rapidly evolving. This would avoid the risk of constraining innovation (or potentially lessening privacy protections) with definitions which at some point in the near future may become no longer relevant or fit for purpose.

Adopting a Risk-based Framework

The White Paper makes reference to a risk-based framework for de-identification, a concept which CANON members would generally support. We have seen such an approach successfully developed by many regulators, including the Ontario Information and Privacy Commissioner in its highly regarded guidance paper, *De-identification Guidelines for Structured Data.*[2] Risk-based frameworks have also been adopted as a globally accepted strategy within the statistical community.[3]

As part of this approach, CANON members also recommend that ISED consider the adoption of a spectrum of identifiability rather than the existing black or white approach in which information is either identifiable or non-identifiable -- completely in or out of PIPEDA's ambit -- respectively. For example, information that poses no serious risk of re-identification could remain outside of PIPEDA, while information with a low risk of re-identification could be covered by PIPEDA, potentially exempted from consent (see below), but subject to other fair information principles as appropriate, including accountability, safeguarding and transparency. Such risk gradations would be determined in accordance with developed guidelines or standards and could be achieved using a variety of different technical methods and appropriate governance models.

Such a risk-based, spectrum approach to de-identification allows for a broad range of innovative uses of data, while accounting for, and mitigating, a reasonable amount of residual risk. It also allows for greater flexibility by allowing the use of other, innovative transformations of data, helping to encourage development of new privacy-enhancing technologies and future-proofing the legislation.

Ultimately, this would be more privacy-protective than a regime in which reasonable people could disagree on the bright line between identifiable and non-identifiable data, and which provides no protection for data which may fall in the grey zone in between. Moreover, such a risk-based approach is more conducive to innovative uses of data subject to appropriate protections. For instance:

[2] https://www.ipc.on.ca/resource/de-identification-guidelines-for-structured-data/
[3] See, for example, *Statistical Disclosure Control for Microdata: A Practice Guide*, developed by a consortium led by the World Bank. https://sdcpractice.readthedocs.io

- A highly detailed dataset, which would serve a significant public benefit but for which re-identification cannot be completely ruled out in the context of a public release, might instead undergo sufficient transformation to allow it to made available within a secure environment in which re-identification is meaningfully prohibited, while retaining its analytical value.
- Rather than performing analytics on raw data, an organization might create a de-identified dataset to provide to its internal analytics team and establish organizational and technical controls sufficient to mitigate against re-identification.

Such an approach can also serve as a vehicle towards increased accountability, as it would require organizations to create a robust governance framework for this "middle category" of data, which could include measures for identifying, documenting and mitigating risks. At the same time, a risk-based approach can allow organizations to focus on, and mitigate against, *likely* or *significantly harmful* risks, rather than spending limited resources attempting to exhaustively catalogue and mitigate against very remote and relatively harmless risks.

Clarifying the Role of Consent in Relation to De-identification

In the White Paper, ISED raises the potential for an exception to consent for the use and disclosure of de-identified information or pseudonymized information for certain prescribed purposes. What would constitute "prescribed purposes" will of course be the subject of important discussions in which CANON members, among many other stakeholders, would be interested in actively participating. By way of example, we would suggest that ISED consider including within the prescribed purposes at the very least those practices which would fall under the proposed "standard business practices" exception to consent. However, in addition, there could be an even broader range of acceptable practices for which de-identified or pseudonymized data may be used or disclosed without consent by reason of the decreased risk associated with such data and the additional safeguarding measures associated therewith.

Alternatively, and in order to maintain greater flexibility, ISED might wish to consider establishing a set of criteria by which it could be determined whether de-identified or pseudonymized data could be used or disclosed, rather than attempting to prescribe all acceptable purposes. Such criteria may include consideration of socially- and economically-beneficial purposes that outweigh potential privacy harm resulting from possible re-identification.

We would further welcome clarification that the consent exception also applies to the initial creation of de-identified or pseudonymized data which will be used or disclosed for those purposes. We note on this point that, in many cases, an organization looking to create a de-identified dataset will not have a direct relationship with data subjects, nor have any means of contacting them to obtain consent to de-identify in the first place. Given the virtues of de-identification as a privacy protective measure which enables innovation, it would be a counter-productive result for organizations to have to collect and/or retain additional personal information for the purpose of protecting privacy in the future.

Codes of Practice

Lastly, the White Paper speaks to incentivizing the use of standards and codes, including the development of Codes of Practice. We believe that de-identification is an area which would benefit significantly from such a development. We have seen the creation of Codes and frameworks in other jurisdictions, including the United Kingdom[4] and Australia[5], and believe that a similar resource, developed through a stakeholder-initiated process and specific to the Canadian context, would be a highly effective way to promote the use of robust de-identification by Canadian organizations.

CANON was created with just such an effort in mind. It is for all the reasons above that CANON has identified as one of its goals the development of a risk-based framework of principles to promote a more realistic approach to de-identification, subject to a robust governance model. We believe this proactive initiative will help support both private sector organizations seeking to innovate for economic and social prosperity, as well as public and health-sector institutions seeking to leverage their data for public good.

In the past year, we have also seen the creation of data protection certification mechanisms, seals and marks, largely encouraged by the incorporation of these concepts into the EU *General Data Protection Regulation.*[6] The introduction of a similar system of voluntary certifications in Canada may further incentivize privacy-protective operations within organizations. Such a system could, of course, be implemented in combination with a Code of Practice which would set out related technical and operations standards.

Our members look forward to the opportunity to engage with ISED, technical experts, regulators, civil society, and/or groups such as the Standards Council of Canada. Collectively, we can work towards the creation of an accessible, operational resource that supports and enhances Canada's position as a leader in both privacy protection and innovation through the use of de-identification.

Conclusion

Once again, the members of CANON appreciate this opportunity to provide feedback on ISED's proposals for the modernization of PIPEDA, and look forward to have the opportunity to continue a discussion on the issues described above as well as the many benefits of de-identification. Please feel free to contact Patricia Kosseim or Vance Lockton, at info@deidentify.ca, any time.

Yours sincerely,

- The Canadian Anonymization Network

[4] United Kingdom Anonymisation Network (UKAN). "Anonymisation Decision-Making Framework" https://ukanon.net/ukan-resources/ukan-decision-making-framework/

[5] Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the Office of Australian Information Commissioner (OAIC). "The De-Identification Decision-Making Framework." https://data61.csiro.au/en/Our-Work/Safety-and-Security/Privacy-Preservation/De-identification-Decision-Making-Framework

[6] The General Data Protection Regulation (EU) 2016/679 (GDPR). Article 42.