Before the
Office of the Privacy Commissioner of Canada

Proposals for ensuring	)
appropriate regulation of artificial	)
intelligence	)

Comments of Privacy Researchers on Proposals 4 and 5

via e-mail to
OPC-CPVPconsult1@priv.gc.ca

March 13, 2020

Samuelson-Glushko Technology Law & Policy Clinic (TLPC) • Colorado Law

Kelsey A. Fayer • Student Attorney
Blake E. Reid • Director

blake.reid@colorado.edu

Samuelson-Glushko Canadian Internet Policy and Public Interest Clinic (CIPPIC)
University of Ottawa

Vivek Krishnamurthy • Director

vivek.krishnamurthy@uottawa.ca

*Affiliations included listed for identification purposes only

Margot E. Kaminski
Associate Professor of Law
Colorado Law
Director, Privacy Initiative, Silicon Flatirons Center

Vivek Krishnamurthy
Samuelson-Glushko Professor of Law
University of Ottawa

Dr. Solon Barocas
Assistant Professor
Department of Information Science
Cornell University

Dr. Reuben Binns
Researcher
Department of Computer Science
University of Oxford

Dr. Maja Brkan
Associate Professor in EU Law
Maastricht University

Kiel Brennan-Marquez
Associate Professor
William T. Golden Research Scholar
University of Connecticut
School of Law

Lee Andrew Bygrave
Professor of Law, Department of Private Law
University of Oslo

Danielle Keats Citron
Professor of Law
Boston University School of Law

Ignacio Cofone
Assistant Professor of Law
McGill University

Giovanni Comandé
Full Professor in Comparative Private Law
Director LIDER-Lab at Sant'Anna School of Advanced Studies of Pisa

Lilian Edwards
Professor of Law, Innovation & Society
Newcastle Law School

Jessica Fjeld
Assistant Director and Clinical Instructor, Cyberlaw Clinic and Lecturer on Law
Harvard Law School

Michael Geist
Professor of Law and Canada Research Chair in Internet and E-commerce Law
University of Ottawa

Woodrow Hartzog
Professor of Law and Computer Science
Northeastern University
School of Law

Mireille Hildebrandt
Research Professor
Vrije Universiteit Brussels (VUB)

Pauline Kim
Daniel Noyes Kirby Professor of Law
Washington University School of Law

Dr. Mark Latonero
Fellow, Carr Center for Human Rights Policy
Harvard Kennedy School

*Affiliations included listed for identification purposes only

Gianclaudio Malgieri
Doctoral Researcher in Law and Technology
Vrije Universiteit Brussels (VUB)
Lecturer
Sant'Anna School of Advanced Studies of Pisa/University of Pisa

Dr. Florian Martin-Bariteau
Assistant Professor & Director, Centre for Law, Technology and Society, University of Ottawa

Jason Millar
Canada Research Chair in the Ethical Engineering of Robotics and AI
School of Electrical Engineering and Computer Science,
University of Ottawa

Frank Pasquale
Piper & Marbury Professor of Law
University of Maryland Francis King Carey School of Law

Dr. Michael Santoro
Professor of Management and Entrepreneurship
Leavey School of Business
Santa Clara University
Co-Editor-in-Chief, Business and Human Rights Journal

Andrew Selbst
Assistant Professor of Law
UCLA School of Law

Kristen Thomasen
Assistant Professor of Law
University of Windsor

Dr. Michael Veale
Lecturer in Digital Rights and Regulation
University College London

Summary

We write to express our support for the proposals of the Office of the Privacy Commissioner of Canada (OPC) to amend the Personal Information Protection and Electronic Documents Act (PIPEDA) to "[p]rovide individuals with a right to explanation and increased transparency when they interact with, or are subject to, automated processing" (Proposal 4), and "[r]equire the application of Privacy by Design and Human Rights by Design in all phases of processing, including data collection" (Proposal 5).

In our view, a revised PIPEDA should include:

1. An individual right to an explanation of an algorithmic decision with significant effects on individuals;
2. Legal requirements for the application of Privacy and Human Rights by Design in all phases of data processing;
3. Legal requirements that algorithmic systems operating in the "human realm"[1] be tested for their privacy and other human rights impacts prior to their deployment; and
4. Ongoing systemic accountability tools, such as audits, and stakeholder oversight and engagement, to ensure expert oversight over algorithms and the natural and legal persons that build or use them.

We applaud the OPC for framing the question of algorithmic accountability as a matter of both individual rights and compliance. In our view, both elements are necessary to prevent an array of different kinds of potential harms. As we discuss below, the European Union's General Data Protection Regulation (GDPR) illustrates that a system of individual rights combined with compliance obligations is both necessary to provide adequate protections and can create positive feedback loops between individual and systemic forms of accountability.[2]

[1] By the "human realm," we mean the realm in which of uses of algorithmic systems concern individual human beings or the relationship between them. AI systems that predict recidivism or educational attainment are in the human realm, whereas systems that predict the weather or the behavior of subatomic particles are not. Likewise systems that predict diseases in people are within the human realm, whereas those that diagnose illnesses in animals are not.

[2] Margot E. Kaminski & Gianclaudio Malgieri, *Algorithmic Impact Assessments under the GDPR: Producing Multi-Layered Explanations*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3456224.

Table of Contents

Summary ..... **iv**

Discussion **1**

I. A revised PIPEDA should include an individual right to explanation of algorithmic decisions with significant effects on individuals. 1

II. A revised PIPEDA should require the application of Human Rights by Design in all phases of data processing. 6

III. A revised PIPEDA should require algorithmic systems operating in the "human realm" be evaluated for their privacy and other human rights impacts prior to their deployment. 9

IV. To ensure accountability, a revised PIPEDA should include ongoing systematic accountability tools such as audits and stakeholder oversight and engagement, in addition to enhanced transparency measures. 14

Discussion

I. A revised PIPEDA should include an individual right to explanation of algorithmic decisions with significant effects on individuals.

We here discuss the right to explanation outlined in the GDPR and analyze it as a potential model for a revised PIPEDA. Drawing on the GDPR, we recommend that a revised PIPEDA include at least:

1. An individual right to be notified that a decision has been made by an automated system;
2. A right to be notified on a systemic level about the logic of an automated decision-making system;
3. A right to challenge/contest an algorithmic decision with significant effects; and
4. An ex post right to explanation of an individual decision that would be simple enough but deep enough to enable an individual to contest such a decision.

We discuss in greater detail below what such disclosures might entail.

The GDPR governs algorithmic decision-making both through its provisions that apply to all data processing (such as the right to object, the right to rectification (correction), data protection by design and by default, and more) and four Articles that specifically address algorithmic decision-making (Arts. 22, 13, 14, & 15).[3] The GDPR contains both a set of uncontested algorithmic transparency rights (including a right to be notified of "meaningful information about the logic

---

[3] *See* Lillian Edwards & Michael Veale, *Slave to the Algorithm: Why a Right to an Explanation is Probably Not the Remedy You Are Looking for*¸ 16 DUKE L. & TECH. REV. 18, 19 (2017) (noting that the generally applicable provisions of the GDPR also play an important role in governing algorithmic decision-making) [hereinafter Edwards & Veale]; *see id.* (noting "other parts of the GDPR related (i) to the right to erasure ('right to be forgotten') and the right to data portability; and (ii) to privacy by design, Data Protection Impact Assessments and certification and privacy seals"), 23, 77; Bryan Casey, Ashkon Farhangi & Roland Vogl, *Rethinking Explainable Machines: the GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise*, 34 BERKELEY TECH. L. J. 145, 173-76 (2019) (discussing DPIA safeguards); ARTICLE 29 DATA PROTECTION WORKING PARTY, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING AND PROFILING FOR THE PURPOSES OF REGULATION 2016/679, 17/EN. WP 251rev.01 (Feb. 6, 2018) at 29 (discussing DPIA and data protection officer), 34 (discussing right to object) [hereinafter GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING]; *see also* Regulation (EU) 2016/679, of the European Parliament and the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1, arts. 22 13, 14, 15 [hereinafter GDPR].

involved" in an algorithmic decision-making system) and a more debated "right to explanation" of individual algorithmic decisions.[4] We understand the GDPR to create a right to explanation of individual decisions as a necessary element of other individual rights, such as the right to contest such decisions. We will discuss these in turn below.

Article 22. Article 22 of the GDPR states that individuals "have the right not to be subject to a decision based solely on automated processing."[5] According to the GDPR guidelines, Article 22 prohibits solely algorithmic decision-making, except under certain exceptions:[6] contract, explicit consent, or Member State law.

Article 22 applies, however, only when the decision is "based solely" on algorithmic decision-making.[7] The GDPR guidelines explain that unless human involvement is meaningful and carried out by someone who has the authority and competence to change the decision, a company cannot escape Article 22 by adding de minimis human involvement.[8] Additionally, Article 22 applies only when an algorithmic decision produces "legal effects" or "similarly significant" effects.

---

[4] For arguments that a right to explanation does exist, *see* Maja Brkan, *Do Algorithms Rule the World? Algorithmic Decision-making and Data Protection in the Framework of the GDPR and Beyond*, 27 INT'L J. OF INFO. TECH. 91 (2019); Bryce Goodman & Seth Flaxman, *European Union Regulations on Algorithmic Decision-Making and "a Right to Explanation"*, 38 AI MAG. 50, 55–56 (2017); Gianclaudio Malgieri & Giovanni Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, 7 INT'L DATA PRIVACY L. 243, 246 (2017); Isak Mendoza & Lee A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling*, *in* EU INTERNET LAW: REGULATION AND ENFORCEMENT 77, 84 (Tatiani Synodinou et al. eds., Springer, 2017); Andrew D. Selbst & Julia Powles, *Meaningful Information and the Right to Explanation*, 7 INT'L DATA PRIVACY L. 233, 235 (2017). *But see* Sandra Wachter, Brent Mittelstadt & Luciano Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 INT'L DATA PRIVACY L. 76 (2017) (arguing that a right to explanation does not exist).

[5] GDPR, *supra* note 3, at art. 22(1).

[6] GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 19.

[7] GDPR, *supra* note 3, at art. 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."). Both Recital 71 and the guidelines provide examples of decisions with significant effects.

[8] GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 21 ("The controller cannot avoid the Article 22 provisions by fabricating human involvement [, and] must ensure that any oversight of the decision is meaningful, rather than just a token gesture."); *id.* ("[I]f someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing."); *see id.* (noting that the controller "should consider all the relevant data" during analysis of the decision).

Suitable safeguards and the right to contestation. When a company uses algorithmic decision-making and determines that its conduct falls under the contractual exception, explicit consent exception, or an exception provided by Member State law, it still must implement "suitable measures to safeguard the data subject's rights and freedoms and legitimate interests[.]"[9] These safeguards must include "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."[10] This right to express one's view and contest the decision is, effectively, a right to an opportunity to be heard—effectively, the right to "technological due process" that several scholars have called for.[11]

Ex post right to explanation. Recital 71 states that "suitable safeguards . . . should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to *obtain an explanation of the decision reached* after such assessment and to challenge the decision."[12] The guidelines counsel that there is a need for a right to explanation because an individual can challenge a particular decision or express her view only if she actually understands "how it has been made and on what basis."[13] In other words, an individual has a right to explanation of an individual decision because that explanation is necessary for her to invoke her other rights, like the right to contestation or the right to correction, that are explicitly enumerated in the text of the GDPR.[14]

Individual notice and access. Beyond Article 22, Articles 13, 14, and 15 of the GDPR specifically address automated decision-making and include individual notification and access rights.[15] Article 13 requires that companies notify individuals when collecting information from them.[16] Article 14 requires a similar

---

[9] GDPR, *supra* note 3, at arts. 22(2)(b), 22(3).

[10] *Id*. at art. 22(3).

[11] Several U.S. scholars have called for algorithmic due process, mimicking procedural due process rights. *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249 (2008) [hereinafter Citron]; Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1 (2014) [hereinafter Citron & Pasquale]; *see* Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93 (2014).

[12] GDPR, *supra* note 3, Recital 71 (emphasis added).

[13] GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 27.

[14] *See* Isak Mendoza & Lee A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling*, in EU INTERNET LAW: REGULATION AND ENFORCEMENT 77,; see also Andrew D. Selbst & Julia Powles, Meaningful Information and the Right to Explanation, 7 INT'L DATA PRIVACY L. 233, 242 (2017).

[15] GDPR, *supra* note 3, at arts. 13, 14, 15.

[16] *Id.* at art. 13.

set of notices when a company collects information from third parties.[17] Article 15 creates an individual right of access to information.[18] All of these Articles require that companies disclose "the *existence* of automated decision-making, including profiling."[19] Additionally, companies must disclose "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject."[20]

What disclosures should contain. The GDPR's individual transparency provisions (notice, access, and explanation) are not meant to provide expert oversight. They are, as discussed, designed to enable individuals to invoke their other rights.[21] Therefore, individuals need not be provided with source code.[22] However, companies need to provide individuals with enough information so they can understand what they are agreeing to (if a company is relying on the explicit consent exception),[23] contest a decision,[24] and correct erroneous information, including erroneous inferences.[25]

To ensure that individuals can act on the explanation they are given, Article 12 requires companies to communicate clearly. Communications must be both comprehensible and actionable.[26] The goal of Article 12 is to prevent companies

---

[17] *Id.* at art. 14.

[18] *Id.* at art. 15. *See* GDPR, *supra* note 3, Recital 63 (described as "[r]ight of access").

[19] GDPR, *supra* note 3, at arts. 13(2)(f), 14(2)(g), 15(1)(h) (collectively, "meaningful information" provisions) (emphasis added).

[20] GDPR, *supra* note 3, at arts. 13(2)(f), 14(2)(g), 15(1)(h).

[21] GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 27 ("The controller should provide the data subject with general information . . . which is also useful for him or her to challenge the decision . . . . The data subject will only be able to challenge a decision or express their view if they fully understand how it has been made and on what basis.").

[22] *Id.* at 25 ("[N]ot necessarily a complex explanation of the algorithms used or disclosure of the full algorithm."), 31 ("Instead of providing a complex mathematical explanation about how algorithms or machine-learning work, the controller should consider using clear and comprehensive ways to deliver the information to the data subject.").

[23] *Id.* at 13 ("Controllers seeking to rely upon consent as a basis for profiling will need to show that data subjects understand exactly what they are consenting to.").

[24] *Id.* at 27.

[25] *Id.* at 17–18 ("Individuals may wish to challenge the accuracy of the data used and any grouping or category that has been applied to them. This rights to rectification and erasure apply to both the 'input personal data' (the personal data used to create a profile), and the 'output data' (the profile itself or 'score' assigned to the person)."), 31 ("Controllers providing data subjects with access to their profile in connection with their Article 15 rights should allow them the opportunity to update or amend any inaccuracies in the data or profile.").

[26] Malgieri & Comandé, *supra* note 4 (introducing the concept of legibility to this debate: "legibility is concerned with making data and analytics algorithms both transparent and

from overwhelming individuals with information that is not useful to them or is unnecessarily complicated.[27] This does not entitle individuals to all information about an algorithm, but it requires the companies to give the individual more than a counterfactual.[28] According to the GDPR guidelines, individuals should be told the categories of data used, why these categories are considered relevant,[29] the "factors taken into account for the decision-making process, and . . . their respective 'weight' on an aggregate level[.]"[30] Individuals should be told the sources of data in a profile,[31] and how that profile was built, "including any statistics used in the analysis[.]"[32] Companies should explain why a profile is relevant and how it is used in a decision.[33]

The GDPR's transparency rights are closely connected to its other underlying individual rights. One needs to see and understand errors to meaningfully invoke a right of correction. One needs to see what factors are used in a decision to meaningfully invoke a right against discrimination. Otherwise, information asymmetries render underlying rights ineffective.

---

comprehensible") (citing Richard Mortier, et al., *Human Data Interaction: The Human Face of the Data-Driven Society*, MIT TECH. REV. (2014); *see* Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1520 (2013) (discussing the related concept of interpretability) [hereinafter Zarsky].

[27] *See* Mike Ananny & Kate Crawford, *Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*, 20 NEW MEDIA & SOC. 973, 979 (2016) ("[S]trategic opacity—in which actors 'bound by transparency regulations' purposefully make so much information 'visible that unimportant pieces of information will take so much time and effort to sift through that receivers will be distracted from the central information the actor wishes to conceal.'")[hereinafter Ananny & Crawford]; Zarsky, *supra* note 26, at 1508 ("The process of merely flooding the public with facts and figures does not effectively promote transparency. It might even backfire."); *see also* Wendy E. Wagner, *Administrative Law, Filter Failure, and Information Capture*, 59 DUKE L. J. 1321, 1324–25 (2010); GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 31 ("Instead of providing a complex mathematical explanation . . . the controller should consider using clear and comprehensive ways to deliver the information to the data subject.").

[28] *But see* Sandra Wachter, Brent Mittelstadt & Chris Russell, *Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR*, 31 HARV. J. L. & TECH. 841 (2018)

[29] GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 31 (explaining good practice recommendations for data controllers).

[30] *Id.* at 27 ("[W]hich is also useful for him or her to challenge the decision.").

[31] *Id.*

[32] *Id.* at 31.

[33] *See id.*

II. A revised PIPEDA should require the application of Human Rights by Design in all phases of data processing.

Relying on a rights-based framework can be powerful, but also puts the onus of enforcement on the very people affected by algorithmic decisions, who will not necessarily see or understand the decisions without help, nor always have the capacity to invoke their rights. It also emphasizes enforcement after-the-fact, rather than preventing rights violations in the first instance. Any such framework of individual rights protection must therefore be bolstered by systemic efforts at both rights protection and accountability.

We support incorporating human rights by design as a legal requirement into a revised PIPEDA, as it is among the most efficient measures available to ensure that algorithmic systems do not adversely impact the full range of human rights protected by law. This is because it is always more effective to build systems that incorporate human rights considerations by design, than to try to fix them once they have caused adverse human rights impacts in the real world.[34]

Since privacy is a human right guaranteed by, *inter alia*, Article 12 of the Universal Declaration of Human Rights, and Article 17 of the International Covenant on Civil and Political Rights, our use of the term "human rights by design" in the discussion that follows should be understood as incorporating the concept of "privacy by design." Given, however, that algorithmic systems have a demonstrated capacity to impact the full range of human rights, and not just privacy,[35] we believe it is appropriate for a revised PIPEDA to mandate the application of "human rights by design," rather than focusing more narrowly on "privacy by design." This broader focus on human rights is consistent with the guidance issued by the OPC in 2018 suggesting that data processing activities on grounds contrary to human rights law are inappropriate under PIPEDA's "appropriate purpose" test.[36]

Our support for incorporating "human rights by design" as a legal requirement in PIPEDA is based on our view that accountability is a problem of human organizations, not just technology.[37] Algorithms are embedded in social systems,

---

[34] Woodrow Hartzog, *The Inadequate, Invaluable Fair Information Practices,* 76 MD. L. REV. 952, 979-81 (2017)

[35] FILIPPO A. RASO, HANNAH HILLIGOSS, VIVEK KRISHNAMURTHY, CHRISTOPHER BAVITZ & LEVIN KIM, ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS: OPPORTUNITIES AND RISKS 14-19 (2018) [hereinafter BKC REPORT]

[36] OFF. PRIVACY COMM'R CAN., *Guidance on inappropriate data practices: Interpretation and application of subsection 5(3)* (May 2018), https://www.priv.gc.ca/en/privacy-topics/collecting-personal-information/consent/gd_53_201805/.

[37] Citron, *supra* note 11, at 1271; Citron & Pasquale, *supra* note 11, at 20-27; Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 124–28 (2014); Ananny & Crawford, *supra* note 27, at 11.

express the values of human institutions (including programmers), and are shaped by the way they are used.[38] Many human choices go into building and deploying an algorithm. These include: (1) the decision on whether to automate processes or systems that had hitherto been operated by humans; (2) what factors or values the algorithm will be designed to optimize (among other design considerations); (3) the training data that is used in developing the algorithm (and what explicit and implicit biases it might reflect); and (4) decisions about the social circumstances in which the algorithm is deployed.[39]

Given the foregoing, algorithmic accountability is as much about making human systems accountable as it is about getting the technology right.[40] The question, then, is how a legal mandate requiring the application of human rights by design advances such accountability, and how such a mandate should be incorporated into the law.

Article 25 of the GDPR is instructive on both questions. As the OPC notes in framing Proposal 5, GDPR Article 25 requires data controllers to "implement appropriate technical and organizational measures . . . which are designed to implement data-protection principles . . . in an effective manner" in view of the risks that data processing activities may pose to the "rights and freedoms of natural persons."[41] Importantly, this provision arguably goes beyond the usual focus of the GDPR to protect not just individuals whose personal data are at risk ("data subjects"), but all natural persons whose wider range of rights and freedoms are put at risk. The "data protection principles," which are set forth in Chapter II of the GDPR, are wide-ranging, running the gamut from accuracy to data minimization to accountability.[42] Moreover, Article 37 of the GDPR specifies that data controllers and processers must implement the specific organizational measure of appointing a data protection officer if one among a number of threshold conditions are met.

Large international technology companies have already invested heavily in implementing such technical and organizational measures. For example, several member-companies of the Global Network Initiative—a multi-stakeholder initiative devoted to protecting the rights to free expression and privacy in the

---

[38] Jessica M. Eaglin, *Constructing Recidivism*, 67 EMORY L. J. 59, 63 (2017)

[39] BKC REPORT, *supra* note 35, at 15-16.

[40] Ananny & Crawford, *supra* note 27; *see also* ANDREW D. SELBST ET AL., *Fairness and Abstraction in Sociotechnical Systems*, *in* PROCEEDINGS OF THE CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, FAT* 59-68 (2019) [hereinafter Selbst].

[41] *Consultation on the OPC's Proposals for ensuring appropriate regulation of artificial intelligence*, OFF. PRIVACY COMM'R CAN., https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/consultation-ai/pos_ai_202001/ (last visited March 13, 2020).

[42] GDPR, *supra* note 3, at art. 5. Lee A. Bygrave, *Data Protection by Design and by Default: Deciphering the EU's Legislative Requirements,* 4 OSLO L. REV. 105 (2017).

digital sphere against unwarranted government interference—have integrated human rights specialists into their product teams to ensure that human rights considerations are reflected in product design, development, and deployment decisions.[43] Numerous companies have appointed executive-level Chief Privacy Officers to provide high-level direction and strategic oversight of their organization's efforts to protect and respect privacy, and in doing so, have "integrate[d] […] privacy into core firm values" and thus moved privacy "from a cost center to a functional concern on the level of product operability, manufacturing accuracy, and process effectiveness."[44] All of these measures are examples of human systems becoming more accountable to ensure that algorithmic systems do not adversely impact human rights by design.

While large companies have the scale, resources, and sophistication to develop systems, policies, and procedures to implement human rights by design, smaller companies generally do not. This creates significant problems when small companies scale up rapidly and begin serving large customer bases, with equally large adverse human rights impacts to boot. The recent controversies involving Clearview AI, a facial recognition start-up whose privacy practices are currently under investigation by the OPC, serves as a case-in-point of a small company that can have a disproportionately negative impact on the enjoyment of the right to privacy by hundreds of millions of people around the world.[45]

To ensure that even the smallest entities take human rights seriously when they are developing or deploying algorithmic systems, a legal mandate to apply human rights by design is necessary. Furthermore, Article 25 of the GDPR helps smaller companies comply with what might otherwise seem to be a vague and onerous regulatory requirement by permitting them to demonstrate their compliance with "Privacy by Design" principles by adhering to industry- and issue-specific Codes of Conduct.[46] The substantive provisions of such Codes of Conduct are subject to the approval of competent data protection authorities.[47] The compliance of companies

---

[43] *See generally* GLOBAL NETWORK INITIATIVE, PUBLIC REPORT ON THE 2015/2016 INDEPENDENT COMPANY ASSESSMENTS (2016), https://globalnetworkinitiative.org/wp-content/uploads/2018/02/Public-Report-2015-16-Independent-Company-Assessments.pdf.

[44] Kenneth A. Bamberger & Dierdre K. Mulligan, *New Governance, Chief Privacy Officers, and the Corporate Management of Information Privacy in the United States: An Initial Inquiry*, 33 LAW. & POL'Y 477, 478-79 (2011).

[45] Kashmir Hill, *The Secretive Company That Might End Privacy as We Know It*, N.Y. TIMES (Jan. 18, 2020), https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html.

[46] GDPR, *supra* note 3, at arts. 40-41.

[47] *Id.* at art. 40.

with such Codes is subject to monitoring by independent, expert entities who are also subject to the oversight of competent data protection authorities.[48]

We support the inclusion of similar provisions in a revised PIPEDA. Specifically, we support the inclusion of a legal mandate requiring all entities that engage in data processing activities that fall within the material scope of PIPEDA to implement appropriate technical and organizational measures to protect the right to privacy, and any other human rights that might reasonably be impacted by the data protection activities in question. We also support the incorporation of provisions modelled on Articles 40-42 of the GDPR that would permit the development of Codes of Conduct to assist companies in complying with this legal mandate. We suggest that, unlike the GDPR, external stakeholders should have the opportunity to review and comment upon draft Codes of Conduct before they are approved by the OPC, and that such stakeholders should play a role in the enforcement of such codes of conduct by independent, expert entities—subject to the supervision of the OPC.

III. A revised PIPEDA should require algorithmic systems operating in the "human realm" be evaluated for their privacy and other human rights impacts prior to their deployment.

A legal requirement that algorithmic systems be evaluated for their human rights impacts (including their impact on the right to privacy) prior to their commercial deployment goes hand in hand with a legal requirement that such systems incorporate human rights by design. Specifically, a testing requirement helps to ensure that systems that seek to respect human rights by design do so in practice.

Such a legal requirement is consistent with trends in the last decade that seek to hold corporations increasingly accountable for the adverse human rights impacts of their business operations and relationships. The United Nations Guiding Principles on Business and Human Rights (UNGPs)[49] has established that corporations have a responsibility to respect human rights. This responsibility, which applies to companies of all shapes and sizes regardless of where in the world they operate, first and foremost requires them to avoid causing or contributing to adverse human rights impacts through their own business activities, and to address such impacts when they occur.[50] Furthermore, the responsibility requires companies to "seek to prevent or mitigate adverse human

---

[48] *Id.* at art. 41.

[49] Human Rights Council, *Guiding Principles on Business and Human Rights: Implementing the United Nations "Protect, Respect and Remedy" Framework,* Human Rights Council, Guiding Principles on Business and Human Rights: Implementing the United Nations 'Protect, Respect and Remedy' Framework, U.N. Doc. A/HRC/17/31 (Mar. 21, 2011) [hereinafter UNGPs].

[50] UNGPs, *supra* note 49, at 14, Principle 13(a).

rights impacts that are directly linked to their operations, products or services by their business relationships, even if they have not contributed to those impacts."[51]

The most important way in which companies operationalize their responsibility to respect human rights is by conducting human rights due diligence ("HRDD") to "identify, prevent, [and] mitigate […] actual and potential human rights impacts."[52] Such due diligence should be undertaken "as early as possible in the development of a new [business] activity or relationship,"[53] and it should also be "ongoing, recognizing that the human rights risks may change over time as the business enterprise's operations and operating context evolve."[54]

Many of the world's largest technology companies have heeded the command of the UNGPs by conducting human rights impact assessments ("HRIAs") of their algorithmic products and services. For example, Microsoft completed a multi-year HRIA into its artificial intelligence technologies in 2018 to "[i]dentify potential risks related to the research and development (R&D) and sales of AI products and services" and "[c]ontribute to Microsoft's continuing efforts to meet its responsibility to respect human rights through its products, services and business activities and relationships."[55] Intel has undertaken a similar HRIA with regards to its portfolio of AI products and services,[56] while Google conducted a more detailed HRIA of a new application provider interface (API) it has developed to recognize the faces of celebrities.[57] Similarly, Waterfront Toronto has commissioned an HRIA to assess the human rights implications of the AI systems that Sidewalk Labs intends to deploy as part of its plan to build a "smart city" in Toronto's Quayside district.[58]

Such HRIAs permit the identification of actual and potential human rights risks using a variety of means and methodologies. They do not explicitly require that

---

[51] UNGPs, *supra* note 49, at 14, Principle 13(b).

[52] *Id.* at 16-17, Principle 17.

[53] *Id.* at comment to Principle 17.

[54] *Id.* at 16, Principle 17(c).

[55] MICROSOFT, HUMAN RIGHTS ANNUAL REPORT 21-22 (2018).

[56] ARTICLE ONE ADVISORS, Challenge: Intel is Committed to Maintaining and Improving Systems and Processes to Avoid Complicity in Human Rights Violations Related to its own Operation, Supply Chain, and Products. In 2016, Intel Decided to Undertake a Human Rights Impact Assessment (HRIA) to Refresh its Risk Profile, Identify Potential Gaps and Strengthen its Strategy www.articleoneadvisors.com/intel-hria (last visited Mar. 4, 2020).

[57] BSR, *Google Celebrity Recognition API Human Rights Assessment: Executive Summary* (October 2019), www.bsr.org/reports/BSR-Google-CR-API-HRIA-Executive-Summary.pdf.

[58] Emma Loewen, *Preliminary Human Rights Impact Assessment for Quayside Project*, WATERFRONT TORONTO (Jan. 17, 2020), http://blog.waterfrontoronto.ca/nbe/portal/wt/home/blog-home/posts/preliminary+human+rights+impact+assessment+for+quayside+project.

algorithmic systems be "tested" for their human rights impacts in a sandbox environment, although such testing may be one of several components of an HRIA in this sector.

The notion that companies should evaluate the human rights impacts of their activities prior to embarking on a new course of business conduct, and periodically throughout the product lifecycle, is beginning to harden into law. For example, the French "duty of vigilance" law requires large companies to develop, implement, and publish a vigilance plan to identify and prevent human rights risks linked to their business activities.[59] Similar legislation is under consideration in several European countries, along with efforts that are under way in the European Union itself.[60]

With regard to algorithmic systems specifically, a recent report by Element AI summarizing the outcome of a multi-stakeholder consultation held in October, 2019 suggests that "governments should begin a phased approach to making HRDD and HRIA a regulatory requirement."[61] Moreover, the data protection impact assessment ("DPIA") provisions of GDPR Article 35, when applied to algorithmic systems, institute a legal requirement upon entities that are subject to the GDPR to conduct what are, effectively, Algorithmic Impact Assessments (AIA) when the use of such systems "is likely to result in a high risk to the rights and freedoms of natural persons . . . ."[62] Again, the reference to "natural persons" rather than "data subjects" suggests that the goal of impact assessments is broader than traditional privacy concerns and applies to impacts well beyond the misuse of personal data.

---

[59] *Loi n° 2017-399 du 27 mars 2017 relative au devoir de vigilance des sociétés mères et des entreprises donneuses d'ordre*, J.O., 28 March 2017, no 1. [Law No. 2017-399 of 27 March, 2018 Regarding the Duty of Vigilance of Parent Companies.] (Fr.).

[60] *See* BUSINESS & HUMAN RIGHTS RESOURCE CENTRE, *National movements for mandatory human rights due diligence in European* Countries (Mar. 11, 2020), https://www.business-humanrights.org/en/national-movements-for-mandatory-human-rights-due-diligence-in-european-countries/ (noting that countries that are considering similar laws include Denmark, Finland, Germany, Italy, Ireland, Luxembourg, the Netherlands, Norway, Sweden, Switzerland, and the United Kingdom, while draft legislation to this effect has already been introduced in the Austrian and Swiss Legislatures.)

[61] PHILLIP DAWSON, ELEMENT AI, CLOSING THE HUMAN RIGHTS GAP IN AI GOVERNANCE 8 (2019) https://s3.amazonaws.com/element-ai-website-bucket/whitepaper-closing-the-human-rights-gap-in-ai-governance.pdf.

[62] GDPR, *supra* note 3, at art. 35.

The GDPR's version of AIAs, which are but one tool among many in a larger regulatory ecosystem,[63] can serve as a point of connection between compliance-oriented approaches to accountability and the protection of individual human rights.[64] These AIAs, in an ideal world, would cause companies and public sector entities to test a system in advance, provide ongoing oversight, and document decisions that are made so that when problems are discovered later, there is a way to trace them back to decisions that can be corrected.[65] Such AIAs could require disclosure of performance metrics on an ongoing basis to regulators or external experts or both.[66]

In our view, an Algorithmic Impact Assessment process should at least:

1. Involve civil society as a form of oversight and source of expertise;
2. Involve and engage impacted individuals through representative boards, before an algorithm is deployed;
3. Require companies, or regulators, to help fund the involvement of both of the above, and provide technical expertise or the resources for obtaining technical expertise;
4. Involve external experts in technology along with experts in law and ethics to help define what we mean by terms like "discrimination" or "bias;"[67]

---

[63] Edwards & Veale, *supra* note 3, at 77-80 (understanding this as they discuss the DPIA in the context of many other rights in the GDPR); *see also* Margot E. Kaminski, *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, 92 S. CAL. L. REV. 1529, 1596(2019).

[64] Only one proposal, to our knowledge, suggests using Impact Assessments to establish something resembling individual rights—a system of "enhanced due process mechanisms for affected individuals". Dillon Reisman et al., *Algorithm Impact Assessment: A Practical Framework for Public Agency Accountability*, AI NOW INST. 18 (Apr. 2018), https://ainowinstitute.org/aiareport2018.pdf [hereinafter *Public Agency Accountability AIA*]; *see also* Pauline Kim, Auditing Algorithms for Discrimination, 166 U. PA. L. REV. ONLINE 189 (2017).

[65] Alessandro Mantelero, *AI and Big Data: A Blueprint for a Human Rights, Social, and Ethical Impact Assessment*, 34 COMP. L. & SEC. REV. 754 (2018); L. EDWARDS, D. MCAULEY, & L. DIVER, *From Privacy Impact Assessment to Social Impact Assessment*, *in* 2016 IEEE SECURITY AND PRIVACY WORKSHOPS (SPW), 53-57 (2016); David Wright & Charles D. Raab, *Constructing a Surveillance Impact Assessment*, 28 COMP. L. & SEC. REV. 613 (2012) Charles Raab & David Wright, *Surveillance: Extending the Limits of Privacy Impact Assessment, in* LAW, GOVERNANCE AND TECHNOLOGY SERIES (2012).

[66] Edwards & Veale, *supra* note 3, at 80.

[67] For example, the COMPAS recidivism risk assessment algorithm led to a significant public discussion over different ways of defining discrimination. Julia Angwin et al., *Machine Bias*,

5. Assess algorithms as technological systems embedded in human systems that design and use them;[68]
6. Take place not just before deployment of the algorithm, but on an ongoing basis;
7. Be disclosed to the public (either in part, in summary, or ideally, in whole).[69]

AIAs so construed go far beyond the requirements of the Treasury Board Secretariat's recent Directive on Automated Decision-Making, whose impact assessment provisions do not provide for any meaningful participation by the general public or by independent experts other than designated peer reviewers.[70]

To ensure their effectiveness, the AIA provisions of a revised PIPEDA should include ongoing assessment and performance evaluation requirements, especially for those algorithms that change over time. To make such impact assessments more meaningful, regulatory bodies such as OPC should have the power to spot-check AIAs and take appropriate enforcement measures against entities that do not follow the legal requirements for such assessments, in order to prevent companies from creating impact assessments that serve only their own best interests. As with the "privacy by design" provisions of the GDPR, regulators might, over time, establish best practices and/or encourage sector-specific codes of conduct around algorithmic fairness.

Impact assessments should be coupled with substantive legal backstops to algorithmic decision-making, prohibiting decision-making based on particular factors, or that is discriminatory or biased. For example, Slovenia in its implementation of the GDPR's Article 22 couples a required impact assessment process with a substantive prohibition on discrimination.[71] Backing impact

---

PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (describing leading risk assessment tools for sentencing and corrections developed by Northpointe); Corbett-Davies et al., *A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not that Clear*., WASH. POST. (Oct. 17, 2016) https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-thanpropublicas.

[68] ANDREW D. SELBST ET AL., *Fairness and Abstraction in Sociotechnical Systems*, *in* PROCEEDINGS OF THE CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, FAT* 59-68 (2019) [hereinafter Selbst].

[69] *See* Kaminski & Malgieri, *supra* note 2.

[70] TREASURY BOARD SECRETARIAT, *Directive on Automated Decision-Making* (Feb. 2, 2019), https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592.

[71] Gianclaudio Malgieri, *Automated Decision-Making in the EU Member States: The Right to Explanation and other "Suitable Safeguards" for Algorithmic Decisions in the EU National Legislations*, at 27 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3233611.

assessments with substantive prohibitions may serve a crucial role in ensuring companies take impacted individuals and communities more seriously.

IV. To ensure accountability, a revised PIPEDA should include ongoing systematic accountability tools such as audits and stakeholder oversight and engagement, in addition to enhanced transparency measures.

We view enhanced transparency measures as essential to improving privacy protection, alongside more traditional accountability measures such as audits and enforcement actions.

Transparency in practice is not limited to public disclosure.[72] It includes internal company oversight, oversight by regulators, oversight by third parties, and communications to affected individuals—each of which may be disclosures of different depths and kinds. Frank Pasquale has talked of the importance of "qualified transparency": a system of targeted revelations of different degrees of depth and scope aimed at different recipients, as a manner of creating better governance/accountability for algorithmic systems.[73] A systemic transparency regime includes not just audits and individual transparency, but other tools such as expert input, impact assessments, and general government oversight powers.

Under the GDPR, most companies using algorithmic decision-making are subject to regulatory oversight, must set up internal oversight, and should subject themselves to some forms of third-party oversight such as audits and expert boards. As discussed, companies using algorithmic decision-making must perform an impact assessment.[74] If they fall in the category of companies that must have a data protection officer, that person must be provided information and have

---

[72] *See* Zarksy, *supra* note 26, at 1532 ("Intuitively, transparency is linked to merely one meaning—that the relevant information is disseminated broadly to (1) the *general public*" but "[f]ully understanding this concept, however, calls for distinguishing among the *recipients* of the information transparency policy provides."). *But see* Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 659-60, 662 (2017)(appearing to define transparency only as disclosure to the public).

[73] FRANK PASQUALE, THE BLACK BOX SOCIETY 140–88 (2015).

[74] *Id.* at art. 35(3)(a) (requiring a data protection impact assessment "in a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person"); GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 29–30 (explaining that this requirement "will apply in the case of decision-making including profiling with legal or similarly significant effects that is not wholly automated, as well as solely automated decision-making defined in Article 22(1)").

information-forcing abilities as well.[75] GDPR guidelines suggest that companies performing decision-making with a "high impact on individuals" should use independent third-party auditing provided with "all necessary information about how the algorithm or machine learning system works."[76] And the GDPR in general gives regulators the ability to obtain significant information about companies' practices, including through accessing companies' premises and technologies.[77]

Looking to the GDPR shows why both individual and systemic transparency is necessary. The GDPR's individual transparency rights largely occur after an algorithm has been developed and deployed, when it is difficult (if not impossible) to fix problems—but when an individual needs that particular information to invoke her other rights.[78] By contrast, the GDPR's ongoing, continuous[79] systemic accountability measures are envisioned as being implemented early on in an algorithm's design. This creates, in theory at least, oversight over the development of an algorithm from its inception, and better serves the purposes of correcting error, inaccuracy, and bias over time.

---

[75] *Id.* at art. 38(2) ("The controller and processor shall support the data protection officer in performing the tasks . . . by providing . . . access to personal data and processing operations . . . ."); GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 29–30.

[76] GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 32.

[77] GDPR, *supra* note 3, at art. 58, ("Each supervisory authority shall have all of the following investigative powers: (a) to order the controller and the processor, and, where applicable, the controller's or the processor's representative to provide any information it requires for the performance of its tasks;…(e) to obtain, from the controller and the processor, access to all personal data and to all information necessary for the performance of its tasks; (f) to obtain access to any premises of the controller and the processor, including to any data processing equipment and means, in accordance with Union or Member State procedural law.").

[78] Kroll et al., *supra* note 72; Deven R. Desai & Joshua A. Kroll, *Trust but Verify: A Guide to Algorithms and the Law*, 31 HARV. J. L. & TECH. 1, 39-42 (2017).

[79] *See, e.g.*, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING, *supra* note 3, at 28. Controllers should introduce appropriate procedures and measures to prevent errors, inaccuracies or discrimination on the basis of special category data. These measures should be used on a cyclical basis; not only at the design stage, but also continuously, as the profiling is applied to individuals. The outcome of such testing should feed back into the system design. *See, e.g.*, Ananny & Crawford, *supra* note 27, at 976.