Laboratoire de
CYBERJUSTICE
Laboratory

Rapport du Canada sur l'intelligence artificielle et l'administration de la justice

(XXIème Congrès de l'Association internationale de droit pénal)

Karim Benyekhlef
Gabriel Lefebvre

Document de travail n°34

Avril 2022

RAPPORT DU CANADA

SUR L'IA ET L'ADMINISTRATION DE LA JUSTICE :

I. POLICE PRÉDICTIVE - II. JUSTICE PRÉDICTIVE - III. DROIT DE LA PREUVE

En vue du XXIème Congrès de l'Association internationale de droit pénal

Version finale (AVRIL 2022)

Pr. Karim Benyekhlef[*], Ad.E. et Gabriel Lefebvre[**]

[*] Professeur titulaire, directeur du Laboratoire de cyberjustice, titulaire de la Chaire LexUM en information juridique, Centre de recherche en droit public (CRDP), Faculté de droit, Université de Montréal.

[**] Agent de recherche, Laboratoire de cyberjustice, CRDP, Faculté de droit, Université de Montréal et doctorant à la Faculté de droit de l'Université McGill.

INTRODUCTION GÉNÉRALE AU RAPPORT DU CANADA

Afin de familiariser les lecteurs étrangers avec certaines particularités du système juridique canadien, nous proposons en guise d'introduction générale de ce rapport de faire un bref exposé du découpage constitutionnel **(i)** en matière d'organisation des forces policières, **(ii)** en matière de droit criminel et **(iii)** en matière de droit à la vie privée. Nous préciserons également la manière dont nous avons organisé les différentes parties du présent rapport de manière à faciliter la compréhension des enjeux tout en respectant le plus possible l'ordre proposé par le questionnaire soumis par l'AIDP.

La Confédération canadienne comporte dix provinces, chacune possède l'autorité législative pour adopter des lois dans chacun des champs de compétence mentionnés à l'art. 92 de la *Loi constitutionnelle de 1867*. Il existe également un gouvernement central – le gouvernement fédéral – qui tire ses pouvoirs législatifs de l'art. 91 de la *Loi constitutionnelle de 1867*.

Organisation des forces policières. Le corps de police national au Canada est la *Gendarmerie Royale du Canada* (GRC). L'art. 92(14) de la *Loi constitutionnelle de 1867* prévoit que chaque province a la compétence législative de « l'administration de la justice dans la province ». En vertu de ce pouvoir, les provinces ont la possibilité de constituer un service de police provincial. Seulement trois provinces ont constitué leur propre service de police provinciale soit l'Ontario (*Police provinciale de l'Ontario*), le Québec (*Sûreté du Québec*) et la province de Terre-neuve-et-Labrador (*Force constabulaire royale de Terre-Neuve*). Ces polices ont juridiction partout à travers la province, sauf dans les municipalités qui ont constitué leur propre service de police. Les autres provinces n'ayant pas constitué de police provinciale sont protégées par la GRC qui offre par contrat un service de police dans ces provinces. La GRC a juridiction pour enquêter sur certaines affaires de nature fédérale, elle offre des services de police à chaque province qui n'ont pas de police provinciale et en offre également à certaines communautés autochtones et aux territoires fédéraux. Comme nous le verrons, les grandes villes canadiennes ont généralement choisi de constituer leur propre service de police.

Droit criminel. Au Canada, l'autorité législative en matière criminelle est exclusivement réservée au Parlement fédéral suivant l'art. 91(27) de *Loi constitutionnelle de 1867*. Cette manière d'approcher la loi criminelle fait l'économie d'un morcellement régional des normes en matière criminelle. Elle vise à assurer une uniformité et une constance dans les normes fondamentales qui garantissent le maintien de l'ordre public au pays. Contrairement aux États-Unis où il existe des « federal crimes » et des « state crimes », puisque chaque état peut édicter ses propres lois en matière criminelle, la *loi criminelle* et la *procédure criminelle* est la même à travers tout le Canada.

Droit à la vie privée. En matière de droit à la vie privée, le fédéral et les provinces ont compétence à l'égard des organisations qui sont sous leur juridiction. Il existe donc des lois différentes en matière de protection des renseignements personnels pour les organisations publiques et privées dépendamment si elles sont fédérales ou provinciales. Or, les principes organisant le droit à la vie privée sont sensiblement les mêmes au niveau fédéral et dans les provinces. La particularité des lois en matière de renseignement à la vie privée s'appliquant au secteur privé est que c'est la *Loi* fédérale *sur la protection des renseignements personnels et les documents électroniques*, L.C. 2000, ch. 5 qui s'applique à toutes les provinces qui n'ont pas édicté une loi « essentiellement similaire » à celle-ci. À ce jour, seuls le Québec, la Colombie-Britannique et

l'Alberta ont adopté une loi pour régir le secteur privé qui a été jugée essentiellement similaire à la loi fédérale. Nous reviendrons plus en détail sur ces lois dans la sous-partie 3.1.1.

Afin d'organiser notre Rapport et d'assurer une certaine conformité entre les rapports des différents pays, nous avons choisi de respecter les titres des Parties (I. Police prédictive – II. Justice prédictive – III. Droit de la preuve) et des sous-parties (*Pratiques nationales* - *Cadre normatif – Principes généraux du droit*) proposés par le Questionnaire fourni par l'AIDP. Lorsque nous avons choisi de déroger de l'ordre et de l'organisation proposés, nous avons précisé les motifs justifiant ces changements. Le lecteur s'en apercevra.

PARTIE I. POLICE PRÉDICTIVE

1. Pratiques nationales

En guise d'introduction à notre présentation des outils d'IA utilisés par la police au Canada, , nous chercherons à définir et à recontextualiser cette approche vers la « police prédictive ». L'usage de l'expression « police prédictive » ne semble pas à ce jour répandu dans les communications officielles des instances gouvernementales ou des départements de police au Canada. Aux États-Unis, le *National Institute of justice* (NIJ) avait dès 2013 proposé une première définition : « Predictive policing is the application of analytical techniques - particularly quantitative techniques - to identify likely targets for police intervention and prevent crime or solve past crimes by making statistical predictions. »[1] Au Canada, nous avons pu identifier une première définition dans un rapport sommaire de la *Division de la recherche et de la statistique* (DRS) du *Ministère de la Justice* datant de 2018 : « Predictive policing : when law enforcement identifies criminal activity using mathematical, predictive, and analytical techniques. »[2] Cette définition n'est toutefois pas celle du gouvernement, elle provient d'un spécialiste externe mandaté par la DRS.

Pour comprendre cette approche policière, le premier défi est d'établir une définition qui se situe en dehors de sa promesse commerciale. L'expression de « police prédictive » soutient l'idée *marketing* que la criminalité puisse effectivement être *prédite* à l'aide d'un traitement statistique algorithmique. La définition appropriée devrait traduire l'idée que la « police prédictive » se limite à un *traitement statistique* opéré par des algorithmes, effectué à partir de faits *quantifiables*, et qui offre une *suggestion* sur les lieux, les moments ou les personnes à *risque*. Cette approche est d'autant plus limitée qu'elle ne s'intéresse qu'aux facettes du « quand », du « où » et du « qui » de la criminalité. Comprendre et anticiper véritablement le phénomène criminel nécessite toujours l'interprétation et l'expérience humaine; le « pourquoi » et le « comment » constituent des données indispensables à une véritable compréhension de la criminalité. En raison de ce malaise définitionnel, les chercheurs du Laboratoire de recherche *Citizen Lab* et ceux de l'*Université de Toronto*, derrière le premier Rapport au Canada qui fait état de la « police prédictive » (2020), ont choisi de définir cette approche policière en dehors de sa connotation commerciale et d'une manière large afin d'y inclure les autres techniques de surveillance policière réalisée à partir d'un algorithme. On y définit alors la « police algorithmique » de cette manière-ci : « the use of algorithms by police services for the pre-emptive monitoring and forecasting of potential crime before any crime has occurred. »[3]

L'autre défi que rencontrera quiconque s'aventure à appréhender ce tournant vers la « police prédictive », est que cette approche policière peut difficilement être comprise distinctivement et isolément, car elle se confond en réalité avec une série encore plus large de mouvements de réforme policière ayant cours depuis

[1] Walter L. PERRY, Brian MCINNIS, Carter C. PRICE, Susan C. SMITH et John S. HOLLYWOOD, "Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations", 2013, p. 1, en ligne : https://www.ojp.gov/ncjrs/virtual-library/abstracts/predictive-policing-role-crime-forecasting-law-enforcement

[2] Dennis D. DRAEGER, "Justice Trends 2: Automated Justice Get the Gist of the future for technology in justice", June 2018, en ligne : https://www.justice.gc.ca/eng/rp-pr/jr/jt2-tmj2/index.html; Son auteur est un représentant de *Shaping Tomorrow* – une compagnie qui offre des services de recherche, d'analyse, de stratégie et de planification grâce à un outil d'IA qui serait d' « anticiper les tendances » au bénéfice de ses clients du secteur privé et public. Site Web de Shaping Tomorrow : https://www.shapingtomorrow.com/webtext/10

[3] CITIZEN LAB et INTERNATIONAL HUMAN RIGHTS PROGRAM, UNIVERSITY OF TORONTO, Kate Robertson, Cynthia Khoo et Yolanda Song, "To Surveil and Predict: A Human Rights Analysis of Algorithmic Policing in Canada", 2020, p. 29. [ci-après "Citizen Lab"]

les années 90 au Canada et aux États-Unis : la « predictive policing » s'entremêle avec le « community policing », le « hot spot policing », le « problem-oriented policing » et l' « intelligence-led policing ». Pour l'auteur Bilel Benbouzid, ces réformes ont comme caractéristique commune de chercher à « rendre la police plus proactive et vigilante que réactive et urgentiste – une police plus engagée dans la production de la sécurité que dans la répression des criminels »[4]. La police prédictive au Canada, comme nous le verrons dans la prochaine sous-partie, partage plusieurs caractéristiques avec les approches policières susmentionnées : **(i)** la collaboration avec d'autres acteurs dans la communauté, **(ii)** une intervention motivée par l'analyse et le traitement de renseignements personnels et **(iii)** une action préventive fondée sur le *risque* de victimisation. Aux États-Unis, le National Institute of Justice (NIJ) reconnait également cette transformation substantielle dans le travail quotidien des officiers de police : « Today more than ever, law enforcement work is also proactive. In proactive policing, law enforcement uses data and analyzes patterns to understand the nature of a problem. Officers devise strategies and tactics to prevent or mitigate future harm. »[5] Ce changement d'approche plus large vers la *prévention* du risque de criminalité est également présent au Canada; celui-ci se serait d'ailleurs *accéléré* suite aux attaques du 11 septembre 2001. Au tournant du nouveau millénaire, les départements de police se seraient retrouvés, d'un côté, pressés d'agir par une population qui demandait davantage de « rendement » dans le renforcement de la sécurité et qui réclamait des preuves tangibles de ce rendement (*imputabilité* des policiers) mais, d'un autre côté, ceux-ci devaient également opérer avec des ressources de plus en plus limitées en raison d'un contexte de compression budgétaire : « As a result, Canadian police services are turning to information technologies and innovations as a means 'to create smart, efficient processes and to leverage technology to move away from reactive to proactive policing' (Police Chief 2011, Ontario Association of Law Enforcement Planners Meetings). »[6] La collecte, l'agrégation et l'analyse de renseignements (*intelligence-led policing*) auraient alors permis aux policiers de modifier « substantiellement » leur approche en la réorientant vers une surveillance pro-active et ciblée, une gestion efficace des risques de criminalité et un renforcement préventif de la sécurité[7]. Le recours à l'IA s'inscrirait justement dans cette quête de rendement en offrant une mesure quantifiable à l'eurs intervention des policiers tout en permettant d'économiser leurs ressources[8]. Nous tâchons ici de rappeler qu'il existe toutefois une limite inhérente à ce qui peut être « mesuré » et « quantifié » en termes de « production de sécurité ». La sécurité, une fois comprise à travers les notions plus larges d'harmonie et de paix sociale, serait difficilement quantifiable, voire irréconciliable avec la surveillance pro-active et la vigilance policière *intensifiée* et *hyper-ciblée* suggérées par ces outils de prédiction. Nous ne pouvons réduire la sécurité, une fois comprise au sens de paix sociale, à un quelconque taux de rendement qui découlerait de l'application de la loi par les policiers.

Ce changement d'approche vers la police *préventive* est observable partout au Canada. Tout d'abord, dans la mission déclarée des différentes agences de renseignements policiers intégrées aux départements de police[9]. Nous l'observons également dans l'adoption de nouvelles politiques en matière de sécurité publique

---

[4] B. BENBOUZID, "Quand prédire, c'est gérer, La police prédictive aux États-Unis", *Réseaux*, vol. 211, no. 5, 2018, p. 223. *Cf.* également NATIONAL INSTITUTE OF JUSTICE, "Overview of Predictive Policing,", 2014, en ligne: https://nij.ojp.gov/topics/articles/overview-predictive-policing. ; INTERNATIONAL ASSOCIATION OF CHIEFS OF POLICE, en ligne : https://www.policechiefmagazine.org/changing-the-face-crime-prevention/

[5] Id., *National Institute of Justice*

[6] Carrie B. SANDERS, Crystal WESTON et Nicole SCHOTT, "police innovations, 'secret squirrels' and accountability: empirically studying intelligence-led policing in Canada", *The British Journal of Criminology*, Volume 55, Issue 4, July 2015, p. 711-712.

[7] *Id.*

[8] B. BENBOUZID, « Quand prédire, c'est gérer – La police prédictive aux États-Unis », *Réseaux*, vol. 211, no. 5, 2018, p. 240.

[9] Par exemple, Le *Provincial Operations Intelligence Bureau* de l'*Ontario Provincial Police*, en ligne : http://www.opp.ca/index.php?lng=en&id=115&entryid=576bf77e8f94ace216355e0f : « Their goal is to anticipate, prevent and

axées sur la prévention, comme la *Politique ministérielle en prévention de la criminalité* de la *Sûreté du Québec* (SQ) adoptée en 2001 et son pendant fédéral, la *Stratégie nationale sur la sécurité communautaire et la prévention du crime*, qui officialisent la mission préventive de la police[10]. Même s'il est précisé dans ces stratégies que l'intervention policière « préventive » devrait s'accompagner de moyens non répressifs, nous pouvons craindre que ces approches n'aient finalement permis l'extension du domaine pénal en faisant entrer, dans la mire des policiers, des gens tout à fait innocents mais socialement vulnérables.

En somme, la volonté d'intervenir *avant même* la survenance du crime date d'avant l'opportunité de recourir à l'IA afin de prédire la survenance d'un crime. Dès 2004, cette quête de prédictivité et ce sentiment de devoir faire preuve d'innovation pour lutter contre les nouvelles formes de criminalité, facilitées par les nouvelles technologies, habitaient déjà les chercheurs. La *DRS* publiait alors un rapport sommaire dans lequel l'auteur recommandait le financement par le gouvernement fédéral d'un groupe de recherche multi-sectoriel visant la « cartographisation », la réduction et la prévention de la criminalité (« examines and maps crime trends, forecasts future crime rates and patterns, and estimates the impact of crime (i.e., costs) for both the present and the future »). Ce groupe de recherche impliquerait les acteurs du secteur privé, comme les ingénieurs informatiques, les services de télécommunications et les fournisseurs d'Internet[11]. Au tournant des années 2000, la crainte de l'augmentation de la criminalité en raison de l'arrivée des nouvelles technologies commandait déjà une approche plus *innovante* dans la lutte contre la criminalité. C'est dans ce contexte que s'inscrit le tournant vers la *police prédictive* au Canada.

Dans ce contexte, où les services de police semblent adopter une approche axée sur le renseignement, l'innovation, la prévention et la gestion efficace des ressources, nous pouvons comprendre comment l'outil d'IA constitue un artéfact technologique attrayant pour les policiers en charge de la sécurité publique. Actuellement, il existe toujours un intérêt certain au Canada pour le développement d'outils d'IA aux fins de police prédictive. Par exemple, le récent *Plan ministériel 2020-2021* contient une nouvelle *Stratégie de services de police numériques* qui vise à rendre la GRC « connectée ». Cette *Stratégie* prévoit le développement massif de nouvelles technologies pour prévenir la criminalité : « l'avenir de la GRC reposera sur la mobilité et Internet. » L'un de ses objectifs est de « mieux utiliser les données pour prédire et prévenir le crime et lutter contre celui-ci ».[12] Du reste, une utilisation de ces outils d'IA a déjà eu lieu au Canada. On peut dès lors raisonnablement penser que leur utilisation sera continue et croissante.

Dans la prochaine partie, nous présenterons le recensement des outils d'IA utilisés actuellement au Canada qui a été effectué par les rapporteurs du *Citizen Lab* en 2020, que nous complèterons de nos propres constatations. Nous verrons d'abord que les différentes motivations qui justifient le recours à ces outils algorithmiques de prédiction et de surveillance trouvent leurs assises dans une quête de rendement dans la lutte contre la criminalité et de mesure de ce rendement afin de pouvoir démontrer à la population que la sécurité est efficacement renforcée (meilleure gestion des ressources et des fonds publics, meilleur service aux citoyens). Nous verrons aussi que cette quête pousse les policiers à faire preuve *d'innovation* et à développer leur secteur technologique et de renseignements et que cela a pour conséquence de modifier leur

monitor criminal activity in Ontario. The members of this bureau collect, assess and share intelligence data within the OPP and with other law enforcement agencies. »

[10] SÛRETÉ DU QUÉBEC, *Politique ministérielle en prévention de la criminalité*, 2001, p. 10 et à la p. 15, en ligne : https://www.securitepublique.gouv.qc.ca/police/publications-et-statistiques/politique-prevention-criminalite.html

[11] Stephen SCHNEIDER, « Predicting Crime: The review of Research », Report prepared for the Department of Justice Canada, 2004, p. 30 et à la p. 2-3 pour les craintes exprimées concernant les nouvelles possibilités offertes aux criminels par les technologies.

[12] GENDARMERIE ROYALE DU CANADA, « La GRC branchée », en ligne : https://www.rcmp-grc.gc.ca/fr/grc-branchee?wbdisable=true

approche et de les amener à agir préventivement (*ex ante*), c'est-à-dire *en amont* comparativement au schème traditionnel de la lutte contre la criminalité (*ex post*). Nous verrons enfin que cela les amène finalement à intervenir de manière ciblée, à la suite d'une analyse portant sur le « risque », auprès des personnes évaluées comme étant susceptibles de subir ou de commettre un crime. Nous soumettons donc l'idée que cette quête *prédictive* motivée par l'IA constitue une menace à l'effet de « justice » initialement souhaité par le droit criminel (pacification, harmonisation, sentiment de sécurité et de justice « rendue ») ; *effet de justice*, qui lui, ne peut résulter que (i) d'une application « juste » de la loi criminelle, c'est-à-dire d'une application faite avec parcimonie et retenue de la force étatique, (ii) d'une application humaine de la loi, qui résulte d'un rapport interpersonnel et d'une délibération ou d'un jugement fondé sur l'expérience humaine, et (iii) d'une application qui soit respectueuse de la primauté du droit, de nos droits et libertés constitutionnels. Nous partageons le constat des auteurs français Antoine Garapon et Jean Lassègue, que nous jugeons tout à fait à propos de rapporter :

> « Lorsque la justice a en charge la gestion en temps réel des infractions qu'elle traite sous forme de flux et de *process*, voire lorsqu'elle cherche à incriminer des actes avant qu'ils ne soient commis comme en matière de terrorisme, le principe de présomption d'innocence est menacé. Le fait risque de vaincre la fiction, et c'est le droit qui perdra; et avec lui, nos garanties. (…) Le monde qui s'annonce est cognitif, et non pas normatif, ce qui implique qu'il évolue dans le domaine du fait et pas dans celui des idéalités sous-tendant le droit. »[13]

1.1. *Volonté de rendement et d'innovation* : panorama des outils algorithmiques

En septembre 2020, l'*Université de Toronto* et le *Citizen Lab* publiaient une enquête-phare concernant le recours par les services policiers canadiens à des technologies algorithmiques à des fins de prédiction de la criminalité[14]. Il s'agit du premier et du plus récent recensement exhaustif des technologies d'IA utilisées par la police au Canada. Nous présenterons ensuite un panorama des différents outils de *surveillance* fonctionnant à l'aide d'un algorithme.

Vancouver – *Geodash APS*[15]. Après un essai pilote de 6 mois en 2016, la *Vancouver Police Department* (VPD) est devenue en 2017 le premier département de police au Canada à intégrer à sa pratique quotidienne l'usage d'une technologie algorithmique afin de guider et coordonner efficacement les actions des policiers sur le territoire[16]. Il s'agit d'une application visant à prédire les endroits et les moments où une infraction d'entrée par effraction est *susceptible* de survenir[17]. Déjà en 2017, un chef de police de la VPD évoquait l'idée d'étendre cette application afin de prédire les vols de voiture et les vols commis à l'aide d'une voiture[18]. Le système cumule des données historiques traitées en fonction du type de crime, des coordonnées géographiques ainsi que la date et l'heure. Ce traitement est effectué à chaque 24h et indique au service de police, selon le temps de la journée (par blocs de 2h), des périmètres « à haut risque » (aussi précis que 100 m² ou 500m²). Les patrouilleurs sont ensuite répartis à travers la ville de Vancouver en fonction de cette prédiction dans le but de prévenir les passages à l'acte, par leur simple présence, tout en menant une

[13] Antoine GARAPON et Jean LASSÈGUE, *Justice digitale*, PUF, Paris, 2018, p. 249.
[14] CITIZEN LAB et INTERNATIONAL HUMAN RIGHTS PROGRAM, UNIVERSITY OF TORONTO, Kate Robertson, Cynthia Khoo et Yolanda Song, "To Surveil and Predict: A Human Rights Analysis of Algorithmic Policing in Canada", 2020.
[15] Citizen Lab, p. 42-44.
[16] VANCOUVER POLICE DEPARTMENT, site web : Vancouver Police Adopt New Technology to Predict Property Crime - Vancouver Police Department (vpd.ca)
[17] Citizen Lab, p.42.
[18] VANCOUVER POLICE DEPARTMENT, vidéo à 10:28, en ligne : Vancouver Police Adopt New Technology to Predict Property Crime - Vancouver Police Department (vpd.ca)

surveillance « proactive »[19]. Le *GeoDASH algorithmic policing system* est le fruit d'un partenariat public-privé entre la VPD, l'entreprise *Latitude Geographics/Geocortex* et des chercheurs universitaires. Cette technologie est mue par une volonté d'innovation et de rendement dans la lutte contre la criminalité et par l'espoir de pouvoir court-circuiter la criminalité : pour le Constable en Chef, il s'agit de développer des stratégies *nouvelles* et *innovantes* afin de prévenir le crime et d'intervenir *avant même* qu'il ne survienne[20].

Toronto[21]. Dans la même famille d'outil algorithmique, le *Toronto Police Service* (TPS) se dit intéressé et considère utiliser un outil algorithmique capable d'identifier des zones où il existe un « haut risque » qu'un crime contre la propriété ou un crime par arme à feu ne survienne. Cet outil fournirait également une suggestion quant au nombre de patrouilleurs à déployer dans ces zones à risque pour les 12 prochains mois. Cet outil est le fruit d'un partenariat qui a débuté vers 2016 entre la police et une firme privée, *Environics Analytics*, qui offre des services d'analyse de données pour les entreprises. Ces prédictions prendraient en compte plusieurs facteurs dont le taux de criminalité de la dernière année, l'âge, le revenu et le type de logement des délinquants dans chaque quartier[22]. Un meilleur rendement dans la sécurité publique, une meilleure gestion des ressources policières ainsi que la nécessité d'améliorer les services offerts aux citoyens sont avancés comme étant les principales motivations du recours à cette technologie[23].

Edmonton – *Community solutions accelerator.* Inspiré par le modèle entrepreneurial des « business accelerator », le *Community Solutions Accelerator* (CSA), mis en place par la *Edmonton police service* (EPS) en 2020, est un laboratoire d'innovations policières réunissant plusieurs acteurs privés et publics qui développeront dans le futur des solutions technologiques aux problématiques touchant cette communauté. Parmi les acteurs derrière cette initiative, on compte la *Edmonton Police Foundation*, et des partenaires privés comme l'*Université d'Alberta*, et des entreprises comme *ATB Financial*, *TELUS* et *Motorola Solutions Canada*[24]. Les partenaires corporatifs fourniront des ressources techniques comme des espaces de travail, une infrastructure informatique et de l'expertise[25]. Ce laboratoire permettrait de développer des applications susceptibles de combiner des *données* provenant de diverses sources et qui pourraient

[19] Citizen Lab, p. 43 et s.
[20] Site Web de la VPD : Vancouver Police Adopt New Technology to Predict Property Crime - Vancouver Police Department (vpd.ca)
[21] Citizen Lab, p. 44-45.
[22] Citizen Lab, p. 44-45. Richard BOIRE, " Data-Driven Decisions for Law Enforcement in Toronto", *Machine Learning Times*, 2018, en ligne : Data-Driven Decisions for Law Enforcement in Toronto « Machine Learning Times (predictiveanalyticsworld.com)
[23] Citizen Lab, p. 44-45. Site Web d'Environics Analytics : Toronto Police Service is 2016 Client of the Year | News | Environics Analytics.
[24] Caley RAMSAY et Vinesh PRATAP, "Edmonton police use data, artificial intelligence to combat crime", *Global News*, February 12, 2020, en ligne : https://globalnews.ca/news/6535688/edmonton-police-data-ai-community-solutions-accelerator/. En septembre 2021, la *Edmonton Police Foundation* et ses partenaires se sont associés avec l'accélérateur d'entreprises de la Silicon Valley *Alchemist* pour lancer un nouvel accélérateur de "gestion des problématiques sociales" la *TELUS Community Safety & Wellness Accelerator*, voir en ligne : https://edmontonpolicefoundation.com/csa. Dans le communiqué de presse de la Edmonton Police Foundation du 23 septembre 2021 disponible dans l'hyperlien qui precede, on donne des exemples d'outils qui pourraient en émerger : « Predicting domestic violence earlier, for early intervention; empowering homeless people with tools that predict needs and match solutions; technology-based addiction management/reduction solutions; solving cold cases on missing people; gamified platform to provide racial bias awareness and corrective solutions; proactive mental health and wellness platforms for individuals and businesses/entities; predictive tool to enable law enforcement to help offenders of certain crimes go through rehab instead of putting them through the criminal justice system. » Ce nouvel accélérateur vise à "targeting ventures that apply technology solutions, especially artificial intelligence, machine learning and advanced analytics, to community safety & wellness. » On définit « safety challenges » ainsi : « Solutions that increase safety in the community (e.g., theft reduction, improved road safety, food safety, etc.)», voir en ligne : https://cswaccelerator.com/our-why/
[25] Caley RAMSAY et Vinesh PRATAP, "Edmonton police use data, artificial intelligence to combat crime", préc., note 24.

fonctionner à l'aide d'un système d'IA fonctionnant par apprentissage-machine (*machine-learning*)[26]. Ces technologies pourraient éventuellement être commercialisées. À titre d'exemple, l'un des premiers projets annoncé – financé par *Alcanna Inc*. – est le développement d'une technologie visant à prévenir le vol dans les magasins d'alcool.

Lors d'une conférence de presse en février 2020, le Chef de police de l'EPS souhaitait également utiliser ces innovations technologiques afin d'analyser le lien entre la criminalité et la consommation de métamphétamine afin de mieux cibler les personnes vulnérables susceptibles de consommer ces drogues et d'intervenir préventivement auprès d'elles pour les diriger vers le système de santé[27]. Ce laboratoire s'inspire alors du modèle HUB, adopté par certaines provinces, dont l'objectif est de résoudre des problématiques sociales, comme la toxicomanie, par l'identification des personnes les plus « à risque » à la suite d'un partenariat et d'une concertation entre différents organismes[28]. L'approche intersectorielle adoptée par la EPS semble permettre alors la mise en commun de renseignements provenant d'une variété de sources; on pense notamment au système de santé, aux services sociaux, au système de protection de l'enfance et à la police[29]. Même si le Chef de police affirme que la plupart des données qui seront utilisées par le CSA sont déjà accessibles par ces agences, que les futures technologies feront l'objet d'une analyse de l'impact de leur utilisation (*Privacy and Impact Assesment*) et qu'ils comptent travailler avec le *Commissariat à la vie privée* pour réguler leurs pratiques[30], il n'en demeure pas moins que cette collaboration soulève plusieurs craintes liées aux échanges de renseignements personnels entre organismes[31].

Cette initiative est mue, ici aussi, par un désir d'innovation dans la lutte contre la criminalité. Dans le cas du projet sur la toxicomanie, celle-ci serait justifiée au nom de la vulnérabilité de ces personnes[32]. La nécessité d'offrir un meilleur rendement dans le renforcement de la sécurité est également évoquée pour justifier une telle approche; on fait alors référence aux besoins d'affecter efficacement les ressources, en raison de leur quantité limitée, afin de sauver temps et argent et d'alléger la pression sur le système de santé, les policiers et le système de justice qui sont décrits comme étant surchargés[33].

---

[26] Kelly CRYDERMAN, "Edmonton police create Community Solutions Accelerator with aim to reduce crime", *Globe and Mail*, february 28, 2020, en ligne : Edmonton police create Community Solutions Accelerator with aim to reduce crime - The Globe and Mail

[27] EDMONTON JOURNAL, "Community Solutions Accelerator to fight crime", vidéo en ligne : https://www.youtube.com/watch?v=GqeBnDXR9bI&t=10s

[28] Citizen Lab, p. 55. Pour en apprendre davantage sur le modèle HUB en général : SÉCURITÉ PUBLIQUE CANADA, « The Hub Model / Situation Table », en ligne : Crime Prevention Inventory (publicsafety.gc.ca)

[29] Site Web de Motorola solutions : https://newsroom.motorolasolutions.com/news/partnering-with-technology-to-fight-crime-and-improve-public-safety.htm ; Kelly CRYDERMAN, "Edmonton police create Community Solutions Accelerator with aim to reduce crime", *Globe and Mail*, february 28, 2020, en ligne : Edmonton police create Community Solutions Accelerator with aim to reduce crime - The Globe and Mail

[30] Anna JUNKER, "Edmonton police launch Community Solutions Accelerator, using data to reduce crime"*, Edmonton journal*, 11 février 2020, en ligne : https://edmontonjournal.com/news/local-news/edmonton-police-launch-community-solutions-accelerator-using-data-to-reduce-crime . Les renseignements personnels seront gérés par le EPS et le transfert de données à l'Edmonton Foundation Police qui, elle, est chargée de transmettre les données aux participants ("Challenge contestants"), se limiterait à des données qui ne permettent pas d'identifier les personnes, voir la Charte du CSA disponible sur le site Web de la Edmonton Police Foundation, en ligne : http://truebluefriendyeg.com/wp-content/uploads/2020/07/Final-Signed-CSA-Charter-Document.pdf

[31] Kelly CRYDERMAN, "Edmonton police create Community Solutions Accelerator with aim to reduce crime", *Globe and Mail*, february 28, 2020, en ligne : Edmonton police create Community Solutions Accelerator with aim to reduce crime - The Globe and Mail.

[32] Site Web de Motorola solutions : https://newsroom.motorolasolutions.com/news/partnering-with-technology-to-fight-crime-and-improve-public-safety.htm

[33] Kelly CRYDERMAN, "Edmonton police create Community Solutions Accelerator with aim to reduce crime", *Globe and Mail*, february 28, 2020, en ligne : Edmonton police create Community Solutions Accelerator with aim to reduce crime - The Globe and

Saskatchewan – *Saskatchewan police predictive analytics lab*[34]. Un laboratoire d'innovations technologiques a également été mis en place en 2015 en Saskatchewan; y collabore activement la *Saskatoon Police Service* (SPS), l'*Université de la Saskatchewan*, le gouvernement de la Saskatchewan et les services sociaux de la province. On y a développé une technologie algorithmique qui permettrait de prédire et de cibler les personnes qui sont « à risque » d'être victime d'un crime. Cette technologie est utilisée afin de guider les interventions policières. Le modèle algorithmique développé par le *Saskatchewan police predictive analytics lab* (SPPAL) permettrait d'identifier les enfants et les jeunes susceptibles de faire l'objet d'un enlèvement. Le SPPAL a également l'intention d'utiliser cette technologie pour intervenir préventivement auprès des délinquants récidivistes, des personnes vivant avec des problèmes de toxicomanie ou atteints de troubles mentaux et pour prévenir la violence domestique. Cette technologie fonctionne actuellement grâce aux données de la SPS, mais on compterait y incorporer les données de tous les services de police municipaux de la province et de la Division « F » de la GRC, qui est la division de la GRC associée à la province de Saskatchewan[35]. Le SPPAL a également l'intention d'intégrer éventuellement dans le développement de ses modèles algorithmiques des données provenant des médias sociaux[36]. Même si ce modèle ne semble pas fonctionner actuellement grâce à un partage massif de données entre les services sociaux de la province et la police, il en demeure pas moins que cette approche – décrite comme une « extension du modèle HUB » déjà implanté en Saskatchewan[37] - pourrait intégrer ces données à son algorithme comme le suggérait le *Citizen Lab* : « the potential use of Hub model data in algorithmic policing methods was recognized by Public Safety Canada in 2015 when it reported that "[i]ntegrated health, social services, education and criminal justice data analysis will help to identify and plan predictive risk patterns at local, regional and provincial levels". »[38] Cette approche innovante est ici aussi justifiée par une volonté d'intervenir plus efficacement auprès des populations jugées *vulnérables* ou « à risque » afin d'assurer leur sécurité, ce qui par le fait même, devrait assurer la sécurité de toute la communauté[39].

Quelle est la situation au Québec? Le rapport du *Citizen Lab* ne traite pas du potentiel usage d'outils d'IA par les services de police au Québec. Selon la *Ligue des droits et Libertés* (LDL), il y aurait toutefois « de bonnes raisons de penser qu'en tant que deuxième corps de police municipal en importance au Canada, le *Service de Police de la Ville de Montréal* (SPVM) pourrait avoir intégré des outils statistiques de prédiction du crime à ses stratégies de lutte contre la criminalité »[40]. En novembre 2019, dans le cadre d'une séance publique de la *Commission de la sécurité publique*, le SPVM a refusé de confirmer cette possibilité évoquée par la LDL sous prétexte qu'il s'agissait là de simples « technicalités d'enquêtes policières »[41].

Mail . La *Edmonton Police Foundation* affirme que les principaux objectifs de la CSA est de "Diminishing harm to individuals - Disrupting, mitigating, and decreasing crime and disorder. - Creating new opportunities for social and economic prosperity including better healthcare outcomes for our most vulnerable. » On mentionne notamment les principes suivants : « Principle 1: Above all, our focus will be on Community Safety and how best to maximize this for all Albertans - Principle 2: Work to create a better experience for Albertans most in need through human-centred design and innovation. - Principle 4: Create new opportunities for social and economic prosperity for Albertans most in need. », Voir Site Web de la Edmonton Police Foundation, en ligne : https://edmontonpolicefoundation.com/csa

[34] Citizen Lab, 51-52.

[35] Citizen Lab, p. 51.

[36] *Id*., p. 51-52

[37] Pour en apprendre davantage sur le modèle HUB en Saskatchewan, *Cf*. SÉCURITÉ PUBLIQUE CANADA, en ligne : Crime Prevention Inventory (publicsafety.gc.ca) et Crime Prevention Inventory (publicsafety.gc.ca)

[38] Citizen Lab, p. 55. *Cf*. également SÉCURITÉ PUBLIQUE CANADA, "Economics of policing and community safety. Policy Makers' Dialogue on Privacy and Information Sharing", Workshop Report, 2015, p. 13.

[39] Citizen Lab, p. 52.

[40] LIGUE DES DROITS ET LIBERTÉS, en ligne : https://liguedesdroits.ca/memoire-reconnaissance-faciale-lapi-csp-montreal-2020/

[41] *Id*.

Les outils de surveillance algorithmique. Toujours selon le Rapport du *Citizen Lab*, plusieurs outils de *surveillance*, dont le fonctionnement est assuré par des algorithmes, sont utilisés par les policiers au Canada. Nous les présenterons ici de manière succincte. Tout d'abord, des technologies automatisées de lecture de plaque d'immatriculation sont utilisées par les services policiers de l'Ontario, de la Colombie-Britannique, de la Saskatchewan, de l'Alberta, de la Nouvelle-Écosse, du Québec et de l'Île-du-Prince-Édouard[42]. Le SPVM a reconnu utiliser des systèmes de reconnaissance de plaques d'immatriculation qui comprenaient des lecteurs automatiques[43]. Ce système vise à s'assurer de la conformité des automobilistes au *Code de la sécurité routière*, il vérifie notamment si les conducteurs ont payé leur permis et leur plaque d'immatriculation. Il peut aussi être utilisé pour rechercher et retrouver des véhicules volés ou dans le cadre d'une alerte AMBER (personne disparue)[44]. Ces vérifications sont effectuées à partir de banques de données qui sont mises à jour par la *Société de l'assurance automobile du Québec*, par le *Centre d'information de la police canadienne* (CIPC) et, dans le cadre d'enquêtes précises, par le SPVM.

La *Calgary Police Service* (CPS), la TPS et la GRC utilisent ou ont utilisé des systèmes algorithmiques pour effectuer de la surveillance sur les médias sociaux. Toujours selon le Rapport du *Citizen Lab*, la TPS a eu recours à une technologie d'analyse des médias sociaux fonctionnant grâce à l'IA de l'entreprise *Sysomos*/*Meltwater*[45]. La GRC se serait également adonnée au monitorage des médiaux sociaux à l'aide d'un système nommé *Social Studio* des entreprises *Carahsoft* et *Salesforce*. La GRC a récemment passé un contrat avec une compagnie américaine pour effectuer du « social media monitoring » sur diverses communautés en ligne à l'aide de son logiciel fonctionnant par IA, plus précisément : « the software analyzes relationships between the content and its senders, translates content into hundreds of languages, and filters it based on geographic areas and expressed sentiments. »[46] La *Ontario Provincial Police* (OPP) aurait également développé et utilisé un outil de type « Chat Room Scraping » nommé *ICAC Child On-line Protection System*[47]. Selon les rapporteurs du Citizen Lab, cette technologie serait en mesure de scanner des salons de clavardage en ligne. Grâce à un mécanisme automatisé, elle retient le contenu à la fin de la discussion, le télécharge puis le conserve dans une banque de données avec un moteur de recherche accessible aux policiers.

Les outils de RF bénéficient également d'un certain engouement auprès des services de police au Canada. Les services de police d'Edmonton, de Calgary, de Vancouver, de Toronto et d'Halifax auraient confirmé utiliser ou avoir utilisé cette technologie[48]. La CPS utiliserait aussi un logiciel de RF développé par la *NEC Corporation* nommé *NeoFace Reveal*. Ce logiciel permettrait d'associer des photographies ou des portraits-robots de suspects non-identifiés avec des photographies d'identité judiciaire déjà possédées ou nouvellement prises. Toujours selon *Citizen Lab*, la TPS utiliserait également un logiciel de RF, tandis que

---

[42] Citizen Lab, p. 57.

[43] COMMISSION DE LA SÉCURITÉ PUBLIQUE DE MONTRÉAL, « Rapport sur l'Utilisation par le SPVM de technologies de reconnaissance faciale et de systèmes de reconnaissance de plaques d'immatriculation », Ville de Montréal, Juin 2021

[44] Site Web du SPVM, « Processus d'utilisation du Système de reconnaissance de plaque d'immatriculation (SRPI) », en ligne : Processus d'utilisation du Système de reconnaissance de plaque d'immatriculation (SRPI) - Service de police de la Ville de Montréal (SPVM)

[45] Citizen Lab, p. 58-59.

[46] Anastasia KONINA, "The Privatization of Law Enforcement: Promoting Human Rights through Procurement Contracts", *McGill GLSA Research Series*, *1*(1), 1-36. p. 14. cf. PUBLIC WORKS AND GOVERNMENT SERVICES CANADA, "Request For a Standing Offer M7594-184225/B" (14 April 2020).

[47] Citizen Lab, p. 60-61.

[48] Céline CASTETS-RENARD, Émilie GUIRAUD et Jacinthe AVRIL-GAGNON, « Cadre juridique applicable à l'utilisation de la reconnaissance faciale par les forces de police dans l'espace public au Québec et au Canada - Éléments de comparaison avec les États-Unis et l'Europe », Observatoire international sur les impacts sociétaux de l'IA et du numérique, Chaire de recherche I.A. responsable à l'échelle mondiale, 2020, p. 12.

la *York Regional Police* et la *Peel Regional Police Service* ont entrepris des démarches pour s'en procurer un[49]. Si le SPVM affirme ne pas avoir utilisé cette technologie, il se disait prêt en 2020 à recourir à des tiers qui posséderaient déjà cette technologie dans le cadre d'enquêtes d'envergure[50]. De son côté, la SQ a conclu un contrat en août 2020 avec l'entreprise française *Idemia*[51] pour acquérir une technologie de RF et d'empreintes digitales capables d'associer automatiquement et en temps réel des plaques d'immatriculation ou des personnes ainsi que leurs tatouages, à partir, notamment, de la *Banque centrale provinciale d'empreinte digitale et de photographies signalétiques*; cette technologie serait utilisée dans le cadre d'enquête criminelle précise[52].

Développements récents d'outils d'IA au Canada. En Pennsylvanie, l'Institut MILA, l'*Université McGill* et l'*École de science informatique* de l'*Université de Carnegie Mellon* auraient développé un nouvel outil d'IA dans le but d'aider les services policiers à identifier des victimes potentielles ou des participants dans la traite des personnes sur Internet et sur les réseaux sociaux[53]. L'algorithme appelé *Infoshield* se veut une réponse à la *Stratégie nationale de la lutte contre la traite des personnes 2019-2024* développée par le gouvernement canadien et la GRC. Cette stratégie faisait un appel au développement de nouvelles technologies qui permettraient de lutter contre les nouvelles formes d'exploitation sexuelle[54].

Des chercheurs de l'*Université de Colombie-Britannique* auraient mis sur pied un logiciel d'IA capable de prédire quelles nouvelles drogues de synthèse ont le plus de chances d'être mis en circulation sur le marché[55]. Afin de contourner les règlementations sur les drogues, des laboratoires clandestins s'affairent à modifier certaines molécules de drogues bien connues afin d'éviter qu'elles soient identifiées par les services policiers. La société d'état *Radio-Canada* expliquait que « pour aider les agences gouvernementales à identifier ces nouvelles substances psychoactives potentiellement dangereuses, [ces] chercheurs ont entraîné un algorithme d'intelligence artificielle à partir d'une base de données de 1800 drogues de synthèse. À partir de la structure moléculaire de ces 1800 substances, l'algorithme de type réseau neuronal a généré presque 8,9 millions de drogues synthétiques potentielles. »[56] Le modèle développé par l'*Université de la Colombie-Britannique* serait d'ailleurs déjà utilisé par la *US Drug Enforcement Agency*, l'*Office des Nations unies contre les drogues et le crime*, l'*Observatoire européen des drogues et des toxicomanies* et la *Federal Criminal Police Office of Germany*[57].

Les outils algorithmiques rejetés par les services de police. Certaines technologies d'IA ont été utilisées par nos services de police avant d'être abandonnées. Après avoir testé l'application de RF *NeoFace Reveal*

---

[49] Citizen Lab, p. 62.
[50] COMMISSION DE SÉCURITÉ PUBLIQUE DE MONTRÉAL « Rapport sur l'Utilisation par le SPVM de technologies de reconnaissance faciale et de systèmes de reconnaissance de plaques d'immatriculation », Ville de Montréal, JUIN 2021, ANNEXE 3, p. 20.
[51] *Id*., p.9
[52] Céline CASTETS-RENARD, Émilie GUIRAUD et Jacinthe AVRIL-GAGNON, « Cadre juridique applicable à l'utilisation de la reconnaissance faciale par les forces de police dans l'espace public au Québec et au Canada », Observatoire international sur les impacts sociétaux de l'IA et du numérique, Chaire de recherche I.A. responsable à l'échelle mondiale, 2020, p. 12.
[53] Pascal ROBIDAS, «Un algorithme de conception québécoise contre l'exploitation sexuelle en ligne », Radio-Canada, 2021, en ligne : https://ici.radio-canada.ca/nouvelle/1800791/mila-intelligence-artificielle-algorithme-police-exploitation-sexuelle
[54] Site Web de l'institut MILA, en ligne : https://mila.quebec/des-chercheurs-de-mila-participent-au-developpement-dun-outil-pour-lutter-contre-le-trafic-de-personnes-et-lexploitation-sexuelle-en-ligne/
[55] M.A., SKINNIDER, F., WANG, D. PASIN et al., *"*A deep generative model enables automated structure elucidation of novel psychoactive substances", *Nat Mach Intell* 3, 973–984, 2021
[56] Site Web de Radio-Canada, Émission Les années Lumières, « Des algorithmes pour prévoir les nouvelles drogues de synthèse », en ligne : Des algorithmes pour prévoir les nouvelles drogues de synthèse (radio-canada.ca)
[57] Site Web de l'Université de Colombie-Britannique, « UBC researchers train computers to predict the next designer drugs », 15 novembre 2021, en ligne : https://www.med.ubc.ca/news/ubc-researchers-train-computers-to-predict-the-next-designer-drugs/

pendant trois mois, le service de police d'Ottawa indiquait « ne pas vouloir l'implanter sans consulter la communauté pour assurer le respect de la vie privée et des droits de la personne. »[58]

À la suite du dépôt du rapport de l'*Enquête sur Clearview AI, Inc.* en 2021, plusieurs services de police ont cessé d'utiliser la technologie de RF offerte par *Clearview AI, Inc*. Les différents commissariats à la vie privée[59] à travers le pays avaient alors recommandé à l'entreprise de cesser de rendre disponible son outil au Canada. Plusieurs services de police, dont la GRC, avaient alors recours à cette technologie. Selon l'enquête, cette technologie d'IA recueillait des images à partir des médias sociaux pour constituer une banque de données biométriques. Les commissariats ont alors jugé que cette pratique nécessitait l'obtention par *Clearview AI inc*. d'un consentement explicite de la part de la personne faisant l'objet de la collecte, ce qui n'avait pas été fait[60]. En décembre 2021, la *Commission d'accès à l'information* du Québec (CAI) envoyait une ordonnance à *Clearview AI*, *Inc*. afin qu'elle détruise, dans un délai de 90 jours, toutes les photos de Québécois qu'elle détenait[61].

Finalement, il est possible d'identifier d'autres technologies qui ont été abandonnées pour des raisons légales ou pour des raisons techniques. La TPS a dû abandonner en 2019 un système de détection automatique de coups de feu, nommé *ShotSpotter*, à la suite de craintes exprimées dans la société civile au fait que celui-ci pourrait violer le droit à la vie privée prévue à l'article 8 de la *Charte canadienne des droits et libertés* (ci-après « *Charte* »)[62]. Dans une autre affaire, les services de police de Toronto et de Calgary ont cessé d'utiliser le logiciel de surveillance de médias sociaux de *Media Sonor*, car celui-ci était devenu inutile après que les services de police aient été bannis de Facebook et Twitter pour avoir enfreint les politiques en matière de protection à la vie privée[63] .

1.2. Réception des outils IA au Canada : précaution face à l'expérience américaine

En raison du peu d'informations circulant dans le domaine public concernant les outils d'IA utilisés actuellement par les corps policiers canadiens, il est difficile de trouver des études d'envergure sur *l'efficacité* ou *l'impartialité* des outils particuliers présentement en cours d'utilisation. Par conséquent, l'évaluation des risques liés à ces technologies se traduit plutôt par une attitude de précaution *générale*

---

[58] Céline CASTETS-RENARD, Émilie GUIRAUD et Jacinthe AVRIL-GAGNON, « Cadre juridique applicable à l'utilisation de la reconnaissance faciale par les forces de police dans l'espace public au Québec et au Canada », Observatoire international sur les impacts sociétaux de l'IA et du numérique, Chaire de recherche I.A. responsable à l'échelle mondiale, 2020, p. 12.

[59] Au Canada il y a un Commissariat à la protection de la vie privée du Canada au niveau fédéral et des commissariats provinciaux qui ont pour fonction d'appliquer les lois canadiennes et provinciales concernant la protection des renseignements personnels.

[60] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, Enquête conjointe sur Clearview AI, Inc. : https://www.priv.gc.ca/fr/mesures-et-decisions-prises-par-le-commissariat/enquetes/enquetes-visant-les-entreprises/2021/lprpde-2021-001/#toc7 [ci-après « Enquête sur Clearview AI »]

[61] Tristan PÉLOQUIN, « Clearview AI sommée de détruire ses photos de québécois », *La Presse*, 2021, en ligne : https://www.lapresse.ca/actualites/2021-12-14/commission-d-acces-a-l-information/clearview-ai-sommee-de-detruire-ses-photos-de-quebecois.php Récemment, Clearview AI contestait devant les tribunaux cette ordonnance en affirmant que leur technologie ne leur permettait pas d'identifier les photos des québécois afin de les détruire, Isabelle DUCAS, « Clearview AI dit qu'elle ne peut détruire les photos de Québécois », *La Presse*, 7 février 2022, en ligne : https://www.lapresse.ca/actualites/national/2022-02-07/logiciel-de-reconnaissance-faciale/clearview-ai-dit-qu-elle-ne-peut-detruire-les-photos-de-quebecois.php

[62] Jeff GRAY, "Toronto police end ShotSpotter project over legal concerns", *Globe and Mail*, 2019, en ligne : https://www.theglobeandmail.com/canada/toronto/article-toronto-police-end-shotspotter-project-over-legal-concerns/; ASSOCIATION CANADIENNE DES LIBERTÉS CIVILES, « Shotspotter ne vient pas à Toronto, et c'est une victoire », 2019, en ligne : https://ccla.org/fr/privacy/surveillance-technology/shotspotter-is-not-coming-to-toronto-and-thats-a-win/ ; Andrea JANUS, « Toronto police scrap plans to acquire controversial gunshot-detection system », *CBC news*, 2019, en ligne, https://www.cbc.ca/news/canada/toronto/toronto-police-scrap-plans-to-acquire-controversial-gunshot-detection-system-1.5019110

[63] Citizen Lab, p. 58.

informée par l'expérience américaine. Le Rapport « The Rise and Fall of AI and Algorithms in American Criminal Justice Lessons for Canada » publié en octobre 2020 par la *Commission du droit de l'Ontario* (CDO) illustre bien l'*approche canadienne*. On y exprime des inquiétudes face à l'arrivée de ces nouvelles technologies au Canada en raison de leur impact potentiel sur les droits de la personne puis, on lance un appel à la précaution en établissant dix leçons que le Canada devrait retenir de l'expérience américaine[64].

Dans la même lignée, le rapport du *Citizen Lab* donne une place importante aux craintes exprimées par les intellectuels et les militants des ONG au Canada. Ceux-ci font état, à la lumière de l'expérience américaine, d'inquiétudes raisonnables quant à l'impact potentiel de ces outils sur nos populations déjà marginalisées et sur-représentées dans les interventions de la police au Canada, comme les personnes racisées, les personnes atteintes de troubles mentaux, les personnes issues de la diversité sexuelle et les autochtones[65]. Ces craintes sont partagées par les rapporteurs du *Citizen Lab*. Celles-ci se fondent sur plusieurs études ou enquêtes journalistiques réalisées dans d'autres pays, notamment aux États-Unis et en Grande-Bretagne. Entre autres, les rapporteurs ont émis des craintes quant à l'utilisation par les services policiers de Calgary et de Toronto des technologies de RF de *NEC Corporation* à la lumière d'une étude réalisée en Grande-Bretagne sur d'autres de ses produits et qui établissait des problèmes d'inexactitude et de biais dans le traitement des données[66]. Dans un mémoire déposé en 2020 devant la *Commission de la sécurité publique de Montréal*, la LDL fait également état, à la lumière de la « tendance préoccupante [du recours par la police aux technologies d'IA] qui est croissante en Amérique du Nord depuis environ 2011 », de leurs inquiétudes quant à l'impact du recours aux technologies de RF ou des technologies de prédiction sur les populations qui font déjà l'objet de profilage racial[67]. La LDL fonde ses inquiétudes sur le Rapport Armony-Hassaoui-Mulone déposé auprès du SPVM en 2019 qui faisait état de la sur-interpellation des personnes racisées et autochtones à Montréal. Ce rapport anticipait, lui aussi, les effets néfastes de l'utilisation des technologies prédictives d'IA sur ces populations :

> « Le profilage criminel s'est raffiné, ces dernières années, par l'entremise du développement des technologies de l'information et l'avènement du big data qui ont permis de mettre sur pied des outils de prédiction du crime de plus en plus élaborés, que ce soit dans le domaine de l'analyse géospatiale (identifier les lieux probables des futurs crimes) ou de la récidive (identifier les individus à potentiel élevé de (re)commettre un crime). Si tous les services de police de moindre envergure ont déjà intégré ces techniques de prédiction, il y a tout lieu de croire que ces stratégies d'analyses de la criminalité vont prendre une place de plus en plus prépondérante à l'avenir. Or, au-delà de leur possible efficacité ou inefficacité en termes de baisse de la criminalité, l'emphase mise sur ces outils peut cependant avoir pour effet de renforcer les profilages existants. En effet, dès que le profilage criminel (la prédiction) s'appuie sur des éléments liés directement ou indirectement à l'appartenance « raciale » (la couleur de peau, certes, mais également l'habillement, la démarche, la gestuelle corporelle ou tout simplement le lieu de résidence), il va forcément avoir pour effet d'accentuer les disparités raciales existantes. Et, dans le même mouvement, il va

[64] COMMISSION DU DROIT DE L'ONTARIO, "The Rise and Fall of AI and Algorithms In American Criminal Justice: Lessons for Canada", 2020. [ci-après « CDO1 »]
[65] Citizen Lab, p. 26-28.
[66] Citizen Lab, p. 92 : "Further, a 2018 report by Big Brother Watch indicated that NeoFace Watch, a facial recognition product by NEC Corporation—the company from which the CPS and the TPS procured their facial recognition technologies—was found to produce inaccurate matches 91 to 98 percent of the time, in usage by the Metropolitan Police and South Wales Police in the United Kingdom. (…) Such findings raise questions about the reliability of a technique that can lead to arrests or criminal charges on the basis of misidentification." ; Pour l'étude, cf. BIG BROTHER WATCH, "Face Off: The lawless growth of facial recognition in UK policing", 2018.
[67] LIGUE DES DROITS ET LIBERTÉS, « Mémoire : Étude des technologies de reconnaissance faciale et des lecteurs automatiques de plaques d'immatriculation », 30 octobre 2020, en ligne : https://liguedesdroits.ca/memoire-reconnaissance-faciale-lapi-csp-montreal-2020/#_ftnref10 ; Elle fonde ses inquiétudes, notamment, sur l'étude suivante Will Douglas HEAVEN, « Predictive policing algorithms are racist. They need to be dismantled », *MIT Technology Review*, 17 juillet 2020, en ligne: https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/

> probablement augmenter le nombre d'interpellations portant sur des citoyens non criminels mais appartenant au groupe ciblé. »[68]

Dans le cadre de l'*Enquête sur Clearview AI, Inc*., les différents commissaires à la vie privée fédéraux et provinciaux, sans avoir fait une « évaluation technique de l'exactitude de la technologie de reconnaissance faciale », ont tout de même exprimé leurs « préoccupations liées à la technologie de reconnaissance faciale en général »[69]. Ces préoccupations concernent l' « efficacité », « l'exactitude » et les possibles erreurs d'identification des technologies de RF et se fondent sur une étude de la *National Institute of Standards and Technology* aux États-Unis[70]. Les commissaires y expriment également leur préoccupation particulière quant au taux élevé de faux positifs lors de « l'évaluation du visage des personnes de couleur, et en particulier celui des femmes de couleur, ce qui pourrait entraîner un traitement discriminatoire pour ces personnes. »[71]

Le Canada aurait avantage à créer une liste publique des différents outils d'IA utilisés par les services de police afin de permettre et d'encourager la recherche indépendante sur l'*efficacité*, la *fiabilité* et l'*impartialité* de ces outils. Pour l'instant, il n'existe aucune autre liste que le recensement effectué par les chercheurs du *Citizen Lab*. Il pourrait être intéressant de réaliser au Canada des études similaires à celles effectuées ailleurs dans le monde sur les outils qui sont présentement utilisés par nos corps policiers.

2. Cadre normatif

À ce jour, il n'existe aucune loi, directive, ou politique d'envergure de la part du gouvernement canadien ou des gouvernements provinciaux encadrant *spécifiquement* l'usage des outils IA aux fins de police prédictive ou de surveillance algorithmique[72]. Il n'en existe pas non plus pour encadrer *spécifiquement* le recours à l'utilisation de la technologie d'IA pour rendre une décision dans le cadre d'un procès en matière criminelle (Partie II sur la Justice prédictive et Partie III sur la preuve). En somme, le cadre légal demeure lacunaire et ne permet pas de répondre aux « garanties de fond » en matière de *transparence*, d'*imputabilité*, d'*efficacité*, d'*impartialité*, de *fiabilité*, de *certification* et de *labellisation*, qui sont généralement identifiées dans le droit des technologies comme étant nécessaires pour protéger l'étendue actuelle de nos droits et libertés face aux potentialités de ces nouvelles technologies. Nous proposons **(2.1.)** d'examiner les principes du droit qui fondent les premières tentatives d'encadrement de ces nouvelles technologies et **(2.2.)** de voir comment les lois actuelles sur l'exactitude des renseignements personnels contribuent minimalement à assurer une certaine *fiabilité* aux outils d'IA utilisés aux fins de police prédictive. **(2.3.)** Nous mentionnerons également ce qui constituerait les principaux obstacles au respect de la garantie de *transparence* dans le fonctionnement des outils d'IA.

2.1. Cadre normatif et principes du droit

[68] Victor ARMONY, et al., « Les interpellations policières à la lumière des identités racisées. », 2019, p. 19-20.
[69] Enquête sur Clearview AI, Inc., par. 91-97.
[70] Patrick GROTHER Mei NGAN et Kayee HANAOKA, « Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects », en ligne : Face Recognition Vendor Test, Part 3: Demographic Effects
[71] Enquête sur Clearview AI, Inc., par. 95.
[72] Citizen Lab, p. 9.

Densification normative. Même si les normes que nous présenterons n'ont pas le statut de lois ou de règlements, elles aspirent tout de même à encadrer certains comportements dans le recours aux technologies d'IA par la police ou par un décideur. Les normes, en tant qu'ordonnances visant à réguler, normaliser et prescrire des comportements, peuvent emprunter plusieurs *formes* (orale ou écrite, publiées sur différents formats), régir un nombre plus ou moins grand de *justiciables* (directives publiques ou internes), suivre différents *processus d'élaboration* (ententes diplomatiques, enquête publique, collaboration avec des acteurs privés), être énoncées avec des *forces* de contrainte variables (recommandations, guide, obligations, principes), provenir de diverses *sources* et de *type d'autorité* différent (entente diplomatique entre deux ministres de pays différents, organisme de normalisation, Conseil du Trésor, Commissariat à la protection de la vie privée du Canada, hauts-gradés dans les services policiers, société civile). On observe actuellement au Canada une *multiplication* et une *diversification* des sources de la norme (organisme de normalisation, compagnies privées, département de police, conseil multi-sectoriel, etc.), l'emprunt par la norme de différentes *formes* (directive, déclaration, standard, document d'orientation), une certaine, quoiqu'encore timide, *intensification* de la norme (des normes qui initialement se trouvaient dans la société civile, sous forme de principes ou de directives internes, peuvent avoir été reprises essentiellement ou en partie ou avoir été corrigées par des autorités administratives ou gouvernementales pour fonder leur propre directive), un *enrichissement* du contenu normatif (la *Directive* du gouvernement fédéral – que nous présenterons – peut être modifiée et peut évoluer; d'autres directives internes ont également été modifiées suite à des enquêtes). À la seule vue des activités normatives ayant eu cours durant les 4 dernières années, nous prédisons également une augmentation du *volume* normatif, une éventuelle *extension* des champs de la norme (du domaine administratif au domaine pénal) et nous voyons déjà une *augmentation des acteurs* concernés par ces normes (policiers, concepteurs d'outil d'IA, techniciens en laboratoire, décideurs)[73].

Nous tâcherons donc de présenter ces normes en portant une attention à leur « densité normative », c'est-à-dire leur aspiration et leur capacité à encadrer les comportements, leur degré de détails, leur force de contrainte, voire leur autorité. Il faut savoir que la « densification normative » est un phénomène à la fois *quantitatif* et *qualitatif* – nous tâcherons de décrire les normes en respectant ces deux dimensions. La densification normative est également décrite comme un « processus polarisé » : d'un côté, il y a une *expansion* des *champs* régulés par la norme, une augmentation des *sources* et du *volume* et, de l'autre, il y a une *concentration* de la norme qui s'exprime avec plus de *précision*, avec plus de *force*[74]. À partir de ces normes, de ces premières tentatives d'encadrement des technologies d'IA, il est également possible de dégager un socle de grands principes du droit qui pourraient probablement être repris lors de futurs efforts législatifs pour encadrer précisément le recours aux technologies d'IA par la police ou par nos décideurs.

Principes soutenus par le Canada à l'international. Les engagements du gouvernement fédéral canadien à l'étranger témoignent de l'importance qu'il accorde à la protection des droits humains lors du recours aux technologies d'IA. Nous pensons, entre autres, à la *Déclaration franco-canadienne sur l'Intelligence artificielle* (2018) où le Canada et la France se sont engagés à établir un groupe international d'étude sur ces technologies et à « promouvoir une vision de l'intelligence artificielle centrée sur l'humain et axée sur le respect des droits de la personne, l'inclusion, la diversité, l'innovation et la croissance économique. »[75].

[73] Catherine THIBIERGE (dir.), *La densification normative. Découverte d'un processus*, Paris, Éditions mare & martin, 2013, 1123-1124.
[74] Catherine THIBIERGE (dir.), *La densification normative. Découverte d'un processus*, Paris, Éditions mare & martin, 2013, p. 1108.
[75] *La déclaration franco-canadienne sur l'intelligence artificielle*, 2018, en ligne : https://www.international.gc.ca/world-monde/international_relations-relations_internationales/europe/2018-06-07-france_ai-ia_france.aspx?lang=fra

Cette *Déclaration* se veut un rappel de la *Déclaration des ministres de l'Innovation du G7* adoptée à Montréal le 28 mars 2018 où les représentants du G7 se sont engagés à défendre un développement de la technologie d'IA qui serait « axé sur l'humain » tout en assurant la croissance économique et l'innovation[76].

Nous pensons également à la *Recommandation du Conseil sur l'intelligence artificielle* de l'OCDE (2019) à laquelle adhère le Canada. Il s'agit de la « première norme intergouvernementale » sur l'intelligence artificielle. Celle-ci est organisée autour de « cinq principes complémentaires fondés sur des valeurs » visant une « approche responsable en appui d'une IA digne de confiance », ces principes directeurs sont la « croissance inclusive, [le] développement durable et [le] bien-être », les « valeurs centrées sur l'humain et [l'] équité », la « transparence et [l'] explicabilité », la « robustesse, [la] sûreté et [la] sécurité » et la « responsabilité ». Le deuxième principe - « [les] valeurs centrées sur l'humain et [l']équité » - est particulièrement pertinent pour encadrer l'implémentation d'une technologie IA en droit pénal. Le principe est exprimé en ces termes : « Les acteurs de l'IA devraient respecter l'Etat de droit, les droits de l'homme et les valeurs démocratiques tout au long du cycle de vie des systèmes d'IA. Ces droits et valeurs comprennent la liberté, la dignité et l'autonomie, la protection de la vie privée et des données, la non-discrimination et l'égalité, la diversité, l'équité, la justice sociale, ainsi que les droits des travailleurs reconnus à l'échelle internationale. »[77] Ce principe assure que les nouvelles technologies devront se conformer à l'état du droit dans lequel ils sont déployés, et non l'inverse.

Conformément à la *Déclaration franco-canadienne* qui annonçait la volonté de créer un regroupement international d'experts en IA, la France et le Canada ont mis sur pied le *Groupe international d'experts en intelligence artificielle* qui vise à « soutenir et guider une adoption responsable de l'IA centrée sur l'humain et axée sur le respect des droits de la personne, l'inclusion, la diversité, l'innovation et la croissance économique. »[78] Ce groupe a finalement été renommé le *Partenariat Mondial sur l'Intelligence Artificielle* (PMIA). Le Canada est membre du PMIA, dont l'un des centres d'expertise est à Montréal (*Centre international d'expertise de Montréal pour l'avancement de l'intelligence artificielle*). Le Président sortant du PMIA est le Ministre canadien de l'Innovation, des Sciences et de l'Industrie François-Philippe Champagne (2020-2021). Ce *Partenariat* est reconnu par l'OCDE, il est organisé autour des *Principes de l'OCDE sur l'IA* contenus dans la *Recommandation du Conseil sur l'intelligence artificielle* de l'OCDE; l'OCDE est d'ailleurs un observateur permanent du PMIA et il héberge le Secrétariat du PMIA. Le mandat préliminaire du PMIA confirme que celui-ci vise finalement à créer des groupes de travail et de recherche afin de développer l'IA dans le respect des droits humains, de l'inclusion, de la diversité, de l'innovation et de la croissance économique[79]. À Montréal, le centre d'expertise héberge deux groupes de travail, l'un sur l'IA responsable et, l'autre, sur la gouvernance des données.

---

[76] G7 INNOVATION MINISTERS, « Annex B: G7 Innovation Ministers' Statement on Artificial Intelligence, Montreal, Canada », 28 mars 2018, en ligne : http://www.g8.utoronto.ca/employment/2018-labour-annex-b-en.html

[77] OCDE, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449, 2019, en ligne : OECD Legal Instruments

[78] Liens Web : Mandat pour le Groupe international d'experts en intelligence artificielle | Premier ministre du Canada (pm.gc.ca) ; France and Canada create new expert International Panel on Artificial Intelligence | Gouvernement.fr ; https://www.canada.ca/fr/innovation-sciences-developpement-economique/nouvelles/2019/05/le-canada-et-la-france-uvrent-aux-cotes-de-la-communaute-internationale-a-la-promotion-dune-utilisation-responsable-de-lintelligence-artificielle.html ; https://www.montrealinternational.com/fr/actualites/le-partenariat-mondial-sur-lintelligence-artificielle-officiellement-lance/ ; La *Déclaration du Groupe international d'experts en intelligence artificielle* , 2019, en ligne : https://www.canada.ca/fr/innovation-sciences-developpement-economique/nouvelles/2019/05/declaration-du-groupe-international-dexperts-en-intelligence-artificielle.html

[79] Voir le « mandat préliminaire » du PMIA, Site Web du PMIA : https://www.gpai.ai/fr/a-propos/pmia-mandat.pdf ; Site Web du PMIA : https://gpai.ai/fr/a-propos/

Le standard national CAN/CIOSC 101:2019. Afin d'« aider toutes les entités ouvertes, fermées, sans but lucratif et publiques à s'harmoniser avec les principes fondés sur les valeurs de l'OCDE », le *Conseil stratégique des dirigeants principaux de l'information* (CSDPI) a mis en place « la première norme du monde qui établit des protections éthiques minimales dans la conception et l'utilisation de systèmes de décision automatisés. »[80] Le CSDPI est « une société nationale à but non lucratif créée en juillet 2017 »[81] rassemblant les *Dirigeants principaux de l'information* de plusieurs secteurs, notamment des entreprises, des gouvernements provinciaux et fédéraux, des municipalités et des organismes sans but lucratif. Elle est chargée de construire et d'influencer l'écosystème technologique du Canada et de supporter le pays dans cette nouvelle économie fondée sur les données en développant des standards nationaux de conformité pour encadrer les nouvelles technologies[82]. En l'absence de lois ou de règlements adoptés par notre Parlement, ces standards nationaux se veulent être une règlementation de première ligne des outils d'IA comme en témoigne cette citation d'Alex Benay, Co-fondateur et ex-coprésident : « Nous devons axer nos efforts comme jamais sur l'élaboration de normes technologiques de la prochaine génération afin de combler le vide laissé par les lois et les règlements déjà existants qui n'ont pas suivi le rythme des changements. »[83] Le CSDPI a été accrédité par le *Conseil canadien des normes*, le principal organisme d'accréditation au Canada; ce dernier est une société d'État fédérale créée en vertu de la *Loi sur le Conseil canadien des normes* « en vue d'améliorer la compétitivité du Canada et le bien-être de sa population »[84]. Le standard national CAN/CIOSC 101:2019 élaboré par le CSDPI en 2019 s'applique à tous les organismes privés, publics, sans but lucratif et gouvernementaux qui voudraient recourir à l'IA par apprentissage-machine afin de faire fonctionner un système automatisé de prise de décisions. Fidèle à la mission de l'OCDE, le standard national stipule que « l'IA devrait profiter aux personnes et à la planète en favorisant une croissance inclusive et un développement durable et bien-être »[85]. Il met en place des normes minimales en matière de *transparence*, d'*imputabilité*, d'*impartialité* et d'*efficacité* dans le but de « prot[éger] des valeurs humaines et d'incorpor[er] de l'éthique dans la conception et l'utilisation de systèmes de décision automatisés. »[86]

Conclusion préliminaire. Nous observons à travers ces premières initiatives normatives que les principes qui en découlent et risquent de se répercuter dans nos prochains efforts législatifs proviennent avant toute chose de l'engagement diplomatique du Canada à l'égard de l'OCDE et de sa participation au G7 – deux organismes essentiellement économiques. Les « principes » défendus par le Canada dans son engagement international comportent à la fois une dimension éthique, mais aussi une importante dimension économique. Comme on peut le remarquer à la fois dans la *Déclaration Franco-Canadienne*, dans la *Recommandation du Conseil sur l'intelligence artificielle* de l'OCDE, dans le mandat du PMIA et dans le *Standard national CAN/CIOSC 101:2019*, les « principes » promus dans ces initiatives réfèrent tous au potentiel de « croissance économique » qu'offrent les outils d'IA. Ces initiatives axées sur la croissance ne semblent pas à même de s'appliquer à un domaine aussi particulier et aussi sensible que le droit pénal ou le droit policier.

---

[80] Sénateur Colin DEACON, « Focusing on ethical AU will unlock social and economic opportunity », 13 Avril 2021, en ligne : https://sencanada.ca/en/sencaplus/opinion/focusing-on-ethical-ai-will-unlock-social-and-economic-opportunity-senator-colin-deacon/ On dit de cette norme qu'elle est également « inspirée par la *Directive sur la prise de décision automatisée* du gouvernement du Canada ».

[81] Site Web du CSDPI : https://ciostrategycouncil.com/normes/?lang=fr

[82] Site Web du CSDPI : https://ciostrategycouncil.com/about/

[83] Site Web du CSDPI : https://ciostrategycouncil.com/histoire/?lang=fr

[84] Site Web du CCN : https://www.scc.ca/fr/notre-organisme/ce-que-nous-faisons

[85] Site Web du CSDPI : https://ciostrategycouncil.com/normes/conception-ethique/?lang=fr

[86] Site Web du CSDPI : https://ciostrategycouncil.com/normes/conception-ethique/?lang=fr

Nous verrions évidemment un problème au fait de considérer la « croissance économique » comme une valeur au même titre que les principes relevant des « droits humains ». Une fois appliqué en droit pénal, un tel ordonnancement des valeurs pourrait avoir des impacts disproportionnés sur les droits des accusés et des suspects. Nous ne pourrions tolérer, par exemple, que le droit à la propriété intellectuelle ou au brevet vienne limiter ou concurrencer un droit aussi fondamental que celui à une défense pleine et entière.

Principes se dégageant de la Directive sur la prise de décisions automatisées. Animé par ces mêmes principes, le Gouvernement du Canada, par l'entremise du *Président du Conseil du Trésor*, a émis sa toute première *Directive sur la prise de décisions automatisée* (en vigueur depuis le 1er avril 2019). Elle s'adresse, d'une manière générale, et sans autres précisions, « aux services du secteur public » fédéral qui souhaitent intégrer des technologies d'IA dans leur prise de décisions. Dans l'éventualité où la *Directive* serait élargie pour s'appliquer au contexte particulier du droit criminel et policier, il vaut la peine de s'y intéresser plus substantiellement[87]. Cette directive vise à garantir la *qualité* et la *fiabilité* des résultats (notamment en assurant la qualité des données, art. 6.3.3), l'*impartialité* (par le respect de la primauté du droit, art. 6.3.8, et la nécessité de contrer les biais, art. 6.3.1) et l'*efficacité* de la technologie (Section 6.5), ainsi qu'à assurer une plus grande *transparence* de ces nouvelles technologies (notamment par le partage du code source, malgré certaines exceptions, art. 6.26-6.27, et par certaines exigences en matière d'explication de la décision, appendice C). Elle stipule également que le service public devra « offrir aux clients toute possibilité à leur disposition en matière de recours applicable afin de contester la prise de décision administrative automatisée » (art. 6.4.1). La *Directive* exige encore que la technologie d'IA destinée à prendre des décisions de manière automatisée soit testée à l'aide d'un outil d'évaluation des risques *avant* d'être mise en fonction. Afin de respecter cette obligation, le *Conseil du Trésor* a rendu disponible un *Algorithmic Impact Assessment tool* (AIA).

La *Directive* a été critiquée pour l'imprécision de son objet d'application[88]. Elle semble difficilement applicable aux pratiques de police prédictive et ne semble pas avoir été rédigée pour s'appliquer dans ce contexte. En effet, elle a été conçue pour s'appliquer aux « décisions administratives » et ne fait aucunement référence aux procédures policières ou criminelles. De plus, elle concerne uniquement les institutions fédérales si bien qu'elle ne lie pas les services de police provinciaux ou municipaux où la plupart des technologies d'IA sont utilisées[89]. Si elle s'applique aux institutions fédérales publiques, peut-on penser que la GRC serait liée par ces exigences ? L'*Agence fédérale du revenu* serait-elle également liée par ces exigences lorsqu'elle utilise un outil d'IA aux fins d'enquête sur la fraude fiscale ? Selon la Clause 5.4 de la directive, il semblerait que non. Pour la chercheuse canadienne Céline Castets-Renard, il est clair que la *Directive* actuelle n'est pas applicable en droit criminel : "It applies to any ADS developed or procured after 1 April 2020 and to any system, tool, or statistical model used to recommend or make an administrative decision about a client (the recipient of a service). Consequently, this provision does not apply in the

[87] Nous vous renvoyons, pour plus de détails concernant la *Directive* vers l'étude faite par S. DU PERRON et K. BENYEKHLEF, « Les algorithmes et l'État de droit », Document de travail No 27 Version 1.0 – Juin 2021 à la Section 2.2.1.1, en ligne : https://cyberjustice.openum.ca/simon-du-perron-et-karim-benyekhlef_les-algorithmes-et-letat-de-droit/. *Cf.* également l'étude critique faite par T. SCASSA, « Administrative Law and the Governance of Automated Decision-Making: A Critical Look at Canada's Directive on Automated Decision-Making », October 30, 2020, Forthcoming, (2021) 54:1 *University of British Columbia Law Review*, Available at SSRN: https://ssrn.com/abstract=3722192

[88] CDO1, p. 39.

[89] Citizen Lab, p. 142.

criminal justice system or criminal proceedings."[90] Si elle devait être appliquée aux activités policières, la chercheuse avance que les décisions policières devraient évidemment être considérées comme ayant un très grand impact sur la personne judiciarisée.[91]

De toute façon, nous sommes d'avis que cette *Directive* ne serait pas suffisante pour atteindre les garanties élevées qu'exige une procédure en matière criminelle. Premièrement, la *Directive* a été édictée précisément pour le recours aux outils d'IA par les organismes publics *administratifs* - le droit administratif et le droit criminel ne font pas appel aux mêmes enjeux et n'ont pas le même impact sur la personne judiciarisée. Il serait alors nécessaire d'adopter des règles plus précises et plus rigoureuses afin d'encadrer l'utilisation des outils d'IA en matière criminelle. Deuxièmement, le questionnaire (« checklist ») de l'*AIA* qui accompagne la *Directive* ne permettrait qu'une évaluation « superficielle » des risques liés à l'utilisation de la technologie et s'éloignerait substantiellement des garanties exigées par d'autres outils de conformité au Canada[92], comme les *privacy impact assessments* requis par la *Loi* fédérale *sur la protection des renseignements personnels*, L.R.C. (1985), ch. P-21 : « What a PIA is not: - a superficial legal checklist - a one-time exercise »[93]. Troisièmement, la *Directive* a également fait l'objet de critiques dans un rapport de l'OCDE puisqu'elle serait peu respectée, en pratique, en raison de son caractère non obligatoire[94]. Afin d'assurer le respect de ces garanties, certains chercheurs avancent que l'application d'un cadre normatif visant à assurer la *transparence*, la *fiabilité* et l'*exactitude* devrait être chapeautée par un organisme de surveillance afin d'assurer le respect de ces garanties et l'*imputabilit*é des décideurs ayant recours à l'IA. Ceux-ci proposent d'élargir les pouvoirs de certains organismes déjà existants, comme le *Commissariat à la protection de la vie privée du Canada* ou le *Centre canadien pour la cybersécurité*, afin que ceux-ci soient désormais chargés de veiller à ce que l'utilisation des technologies d'IA par nos services policiers respecte ces exigences en matière de *fiabilité*, d'*impartialité* et de *transparence*[95].

Principes émanant du Document d'orientation préliminaire sur la reconnaissance faciale. Le *Commissariat à la protection de la vie privée du Canada* a publié en 2021, pour faire suite à son *Enquête sur AI Clearview Inc.*, un *Document d'orientation préliminaire* à l'intention des policiers fédéraux, provinciaux et municipaux visant à encadrer leur usage des technologies de RF[96]. Ce cadre normatif vise à renforcer les protections actuelles en matière de renseignements personnels qui peuvent être mises à mal en raison des potentialités techniques des outils de RF. Des obligations préalables à l'utilisation de la technologie de RF y sont prévues. Ces obligations se rapportent aux principes directeurs suivants : « la nécessité et la proportionnalité » dans l'usage de la technologie, l' « exactitude » des résultats, la

[90] Céline CASTETS-RENARD, "Human Rights and Algorithmic Impact Assessment for Predictive Policing", p.106 dans Micklitz, H., Pollicino, O., Reichman, A., Simoncini, A., Sartor, G., & De Gregorio, G. (Eds.), *Constitutional Challenges in the Algorithmic Society*, Cambridge, Cambridge University Press, 2021.
[91] *Id*., p. 107-110.
[92] Citizen Lab, p. 143.
[93] COMMISSARIAT À LA VIE PRIVÉE, en ligne : Expectations: OPC's Guide to the Privacy Impact Assessment Process - Office of the Privacy Commissioner of Canada;
[94] OCDE, "State of implementation of the OECD AI principles: insights from national AI policies", p. 45.
[95] COMMISSION DE DROIT DE L'ONTARIO, Jill R. PRESSER et Kate ROBERTSON (aut.), "AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts", (Toronto: June 2021), p. 24 [ci-après « CDO2 »]
[96] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, *Document d'orientation préliminaire sur la protection de la vie privée à l'intention des services de police relativement au recours à la reconnaissance faciale,* 2021, https://www.priv.gc.ca/fr/a-propos-du-commissariat/ce-que-nous-faisons/consultations/gd_frt_202106/. Cette directive doit s'harmoniser et se rapporter à d'autres directives connexes, notamment la *Directive sur l'évaluation des facteurs relatifs à la vie privée* du SECRÉTARIAT DU CONSEIL DU TRÉSOR, en ligne : https://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=18308

« minimisation des données », la « responsabilité » des décideurs et l' « ouverture, [la] transparence et [l']accès aux renseignements personnels ».

Principes émanant de la société civile. Certaines déclarations de principe provenant d'organismes non-gouvernementaux sont également à même de guider les pratiques policières ou d'inspirer de futurs efforts législatifs. Nous pensons tout d'abord à la *Déclaration de Montréal IA responsable* (2018)[97] qui rappelle l'importance de conjuguer le développement technologique avec le droit à l'équité, le respect de la vie privée et la nécessité d'assurer la participation démocratique. Une autre déclaration de principe au Canada – *The Toronto Declaration* (2018) par *Access Now*[98] – fait état de la nécessité d'assurer l'imputabilité des décideurs lorsqu'ils ont recours aux outils d'IA. Elle exige de prendre des mesures pro-actives pour minimiser l'impact sur le droit à l'égalité et de délimiter clairement l'utilisation de ces technologies à ce qui est nécessaire. Ces déclarations, issues de la collaboration entre différents acteurs de la société civile, visent à promouvoir un usage éthique de ces technologies en se fondant sur les principes d'équité, d'imputabilité, de transparence et d'éthiques (connu sous l'expression « FATE principles »)[99]. Les *Sedona Principles E-discovery* sont également à même d'offrir un cadre normatif principiel aux policiers canadiens dans le cadre de la collecte de preuve. Nous y reviendrons dans la partie III sur la preuve.

Initiatives et directives internes des services de police. Encore à ce jour, l'activité normative la plus foisonnante se trouve au sein des départements de police. En l'absence de cadre normatif contraignant, les corps policiers font preuve d'auto-régulation dans l'utilisation des nouvelles technologies IA. Nous nous désolons de voir que leur encadrement soit laissé à la discrétion des services de police. Contrairement à ce qu'affirmait le SPVM lors d'une conférence de presse, le choix de recourir ou non à ces nouvelles technologies et la manière de les utiliser ne se résument pas à de simples « technicalités d'enquêtes policières »[100], mais font appel à des enjeux collectifs plus larges. En effet, le recours à ces technologies requiert la tenue d'un débat public sur la manière dont nous voulons assurer l'harmonie et la paix sociale dans l'espace public et la manière dont nous voulons ordonnancer et mettre en équilibre l'impératif lié à la sécurité et nos droits fondamentaux. Par le passé, quelques-unes de ces directives internes ont été jugées insuffisantes par les commissaires à la vie privée.

Par exemple, dans le cadre de l'*Enquête sur AI Clearview, Inc.,* concernant l'utilisation de la RF, on apprenait que la GRC avait préalablement ordonné, dans ses consignes de sécurité, de traiter les associations effectuées par la technologie comme des pistes et non comme des résultats confirmés. Les différentes commissaires à la vie privée fédéraux et provinciaux responsables de cette enquête ont jugé qu'une telle consigne n'était pas suffisante pour encadrer la technologie de RF et que des mesures supplémentaires pour assurer l'exactitude et contrecarrer la possibilité de « faux positifs » ou de biais discriminatoires étaient nécessaires[101]. C'est uniquement à la suite de cette enquête, en mars 2021, que la GRC a instauré le

[97] Site Web de la Déclaration de Montréal : https://www.declarationmontreal-iaresponsable.com/la-declaration
[98] Site Web de la Déclaration de Toronto : https://www.torontodeclaration.org/#:~:text=The%20Toronto%20Declaration,-Protecting%20the%20right&text=It%20calls%20on%20governments%20and,to%20equality%20and%20non%2Ddiscrimination
[99] Citizen Lab, p. 9.
[100] LIGUE DES DROITS ET LIBERTÉS, « Étude des technologies de reconnaissance faciale et des lecteurs automatiques de plaques d'immatriculation », 30 Octobre 2020, en ligne : https://liguedesdroits.ca/memoire-reconnaissance-faciale-lapi-csp-montreal-2020/.
[101] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, « Rapport spécial au Parlement sur l'enquête réalisée par le Commissariat à la protection de la vie privée du Canada sur l'utilisation par la GRC de la technologie de Clearview AI », 2021, par. 81. [ci-après : « Rapport spécial sur la Technologie de RF »]

Programme national d'intégration des technologies visant à examiner systématiquement la conformité des nouveaux outils technologiques utilisés dans le cours d'une enquête à la *Loi sur la protection des renseignements personnels* et à la *Charte canadienne des droits et libertés*[102].

De son côté, la CPS a élaboré sa propre politique écrite en matière de collecte de renseignement sur les réseaux sociaux. Les policiers sont autorisés à collecter ces informations, mais doivent se limiter aux fins précises de l'enquête. Elle permet également de recueillir toute information qui pourrait constituer une « menace ». Les rapporteurs du *Citizen Lab* ont critiqué cette directive en raison de son imprécision[103].

En l'absence de normes étatiques contraignantes, la VPD a elle aussi, de sa propre initiative, essayé de limiter l'impact de sa technologie d'IA sur les populations déjà sur-représentées dans les interventions policières. Elle s'est assurée que les données qui alimentent son système proviennent exclusivement des cas d'entrée par effraction signalés par un citoyen ordinaire afin de limiter l'impact des biais policiers qui sont présents dans les données historiques. Elle a également exclu de sa cartographie de surveillance certaines zones plus « sensibles », comme le quartier *Downtown Eastside* où la population est plus pauvre, et elle a finalement donné aux patrouilleurs la directive de ne pas utiliser l'application pour justifier un contrôle de routine d'identification volontaire (*street check*)[104]. Les rapporteurs du Citizen Lab ont reconnu qu'il s'agissait de normes techniques intéressantes pour limiter certains effets néfastes de la technologie, mais ils ont également émis certaines réserves sur leur capacité à contrer d'autres formes de biais discriminatoires et d'enrayer les risques de biais en faveur de la prédiction offerte par la technologie (*automation bias*)[105].

2.2. L'exactitude des renseignements personnels comme garantie minimale de la *fiabilité* des outils d'IA

L'adoption de règles particulières pour encadrer les technologies d'IA en droit criminel et policier, notamment en ce qui a trait à la RF, s'avère nécessaire compte tenu de l'information *sensible* en cause, par exemple les données biométriques qui sont inaltérables, et de la nécessité d'étendre nos protections actuelles au même rythme qu'augmente le niveau de performance des outils technologiques. La capacité de traitement des outils d'IA et la nature des données traitées (données historiques, multi-sources et décontextualisées) commandent la révision du cadre actuel. S'il n'existe pas de règles encadrant précisément le degré de *fiabilité* nécessaire des outils d'IA, nous jugeons que les garanties actuelles en matière d'exactitude des renseignements personnels constituent indirectement une garantie minimale pour assurer leur fiabilité.

Ces nouveaux outils requièrent néanmoins un rehaussement du standard de fiabilité et d'exactitude des banques de données policières ainsi qu'une réactualisation constante de celles-ci en raison du traitement autonome, quasi-instantané et continu effectué par la technologie d'IA. Les banques de données

---

[102] GENDARMERIE ROYALE DU CANADA, "Response to the Report by the Office of the Privacy Commissioner into the RCMP's use of Clearview AI", 10 juin 2021, Ottawa, Ontario, en ligne : https://www.rcmp-grc.gc.ca/en/node/91915. *Cf.* également le Rapport spécial sur la Technologie de RF.

[103] Citizen Lab, p. 58 : "Under the policy, officers may collect publicly available data, including data processed by third-party social network aggregators and software. Officers are restricted, however, to collecting only information that is linked to a specific investigative purpose, including "threat-related information." The policy does not define what "threat-related information" means nor does it restrict the CPS from using products like Media Sonar in the future, should they become useful once more for investigations. »

[104] Citizen Lab, p. 44.

[105] Citizen Lab, pp. 109-110 et 125-126.

couramment utilisées par les services policiers ne semblent pas respecter un standard adéquat pour alimenter des outils d'IA. **(1)** Par exemple, la CIPC, qui est la principale banque de données en matière d'informations sur les casiers judiciaires actuellement utilisée par les services de police et les départements gouvernementaux canadiens est reconnue comme contenant des informations qui ne sont pas à jour et qui sont inexactes[106]. **(2)** La banque de données des dossiers inconsultables de la GRC a également fait l'objet par le passé de sévères critiques quant à leur méthode de conservation. En 2008, la *Commissaire à la vie privée du Canada* a effectué une enquête afin de déterminer si l'exactitude des données contenues dans les fichiers inconsultables de la GRC, avait été évaluée et consignée[107]. En vertu de l'article 18 de la *Loi sur les renseignements personnels*, la GRC peut déclarer certains de ses dossiers comprenant des renseignements personnels collectés dans le cadre d'enquête criminelle comme étant « inconsultable au public »; un tel dossier « pourrait éventuellement renfermer des renseignements sur des personnes [innocentes] si celles-ci se sont trouvées au mauvais endroit, au mauvais moment, en compagnie des mauvaises personnes »[108]. Pour ces raisons, les organismes doivent s'assurer que le contenu de ces fichiers soit limité à ce qui est « légitime » et il leur incombe de *classer* et d'*organiser* l'information dans des dossiers qui sont « repérables ». On exige également que les dossiers soient classés sous des numéros uniques, qu'ils aient été vérifiés et qu'ils aient fait l'objet d'une procédure d'examen périodique. L'enquête du *Commissariat* a révélé que la quasi-totalité « des dossiers n'ont pas fait l'objet de contrôles réguliers destinés à s'assurer qu'il convient toujours de les classer parmi les fichiers inconsultables, conformément à la politique de la GRC »[109]. Il serait évidemment problématique qu'une telle banque de données ou des informations obtenues à partir d'une telle banque de données servent à alimenter un outil d'IA[110]. **(3)** Finalement, nous voyons également un problème avec le fait d'alimenter les outils d'IA à partir des rapports d'incidents remplis par les policiers, en raison des potentiels biais policiers et du manque de standardisation révélée au sein de cette pratique : "The analyst above highlights the lack of standardization and training, as well as selective reporting, which raises concerns regarding data quality and integrity."[111] L'exactitude des données alimentant l'outil d'IA constitue une garantie *minimale* à toute utilisation raisonnable de cette technologie. Des données inexactes ou incomplètes peuvent générer des résultats faussés et conduire à une intervention disproportionnée ou sans motif de la part d'un policier. Par exemple, **(i)** des données biaisées ou inexactes peuvent mener à une détention arbitraire, car il s'agirait de motifs déraisonnables de soupçonner; **(ii)** ou

---

[106] Citizen Lab, p. 85 citant Alyshah HASHAM, "Criminal-record database spotty and out of date, lawyers lament", The Toronto Star, 9 décembre 2016, en ligne : Criminal-record database spotty and out of date, lawyers lament | The Star ; Brigitte BUREAU, "RCMP database remains out of date, police and prosecutors say", *CBC News*, 10 mars 2015, en ligne : RCMP database remains out of date, police and prosecutors say | CBC News. V. également plus récemment, Nicole BROCKBANK, "How a criminal charge laid in Calgary was linked to a Toronto woman who's never been there", *CBC News*, 21 janvier 2021, en ligne : https://www.cbc.ca/news/canada/toronto/false-identity-rcmp-database-1.5881006. Voir aussi, Renata D'ALIESIO et Kathryn blaze CARLSON, "Substantial gap discovered in RCMP database of anonymous dead", *Globe and mail*, 16 Mars 2015, en ligne : https://www.theglobeandmail.com/news/national/substantial-gap-discovered-in-rcmp-database-of-anonymous-dead/article23467796/

[107] COMMISSAIRE À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, *Examen des fichiers inconsultables de la GRC*, 2008, p. 13, en ligne : https://www.priv.gc.ca/fr/mesures-et-decisions-prises-par-le-commissariat/verifications/rcmp_080213/.

[108] *Id*., p. 2.

[109] *Id*., p. 3.

[110] Pour un exemple d'un renseignement non à jour trouvé, *Id*., p. 25 : « Un homme a vu quelqu'un entrer dans une maison de chambre près de chez lui. Croyant qu'il pouvait s'agir d'une affaire de drogue, il a contacté la police. L'enquête a révélé que l'homme en question venait de reconduire sa fille à l'école (à l'autre bout de la rue) et était sorti de son véhicule pour fumer une cigarette. Le dossier a été fermé il y a environ sept ans. »

[111] Carrie B. SANDERS, Crystal WESTON et Nicole SCHOTT, "Police innovations, 'secret squirrels' and accountability: empirically studying intelligence-led policing in Canada", *The British Journal of Criminology*, Volume 55, Issue 4, July 2015, p. 720.

conduire à des techniques policières discriminatoires, comme le profilage racial, enfreignant ainsi le droit à l'égalité et à la non-discrimination[112].

Une obligation pro-active de lutter contre les risques d'erreur ? Du droit actuel, semble se dégager une forme d'obligation pro-active de la part des policiers de s'assurer de l'exactitude des données lorsqu'elles sont utilisées par une technologie d'IA. Tout d'abord, l'art. 6(2) de la *Loi sur la protection des renseignements personnels* oblige tout organisme public fédéral, comme la GRC, « de veiller, dans la mesure du possible, à ce que les renseignements personnels qu'elle utilise à des fins administratives soient à jour, exacts et complets. » Lorsqu'un policier a recours à une technologie de RF, le *Commissariat à la vie privée du Canada* a jugé qu'une consigne interne de précaution générale adressée aux enquêteurs visant à leur demander de considérer les résultats générés par la technologie de RF « comme des pistes, et non comme des correspondances d'identité confirmées », pouvait ne pas être suffisante pour respecter cette obligation; c'est donc dire que des mesures *supplémentaires* pour assurer l'exactitude des données et contrecarrer la possibilité de faux positifs ou de biais discriminatoires seraient exigées lors du recours à une technologie d'IA[113]. Dans le même ordre idée, une loi similaire au Québec prévoit qu'un service de police québécois doit « veiller à ce que les renseignements personnels qu'il conserve soient à jour, exacts et complets pour servir aux fins pour lesquelles ils sont recueillis ou utilisés »[114]: « on peut là aussi se demander si cette obligation d'exactitude pourrait s'appliquer au procédé et aux données biométriques, ce qui permettrait de fonder une obligation de nature à lutter contre le risque d'erreur. »[115] Les organismes publics des autres provinces ont également l'obligation de prendre des mesures raisonnables afin de s'assurer que les renseignements personnels qu'ils collectent et utilisent sont exacts et mis à jour[116], bien que certains corps policiers, comme ceux de l'Ontario, soient exemptés de cette obligation particulière[117].

2.3. Obstacles à la garantie de transparence des outils algorithmiques

L'opacité et la complexité inhérentes au fonctionnement des outils d'IA menacent l'équité du procès en matière criminelle garantie par l'article 7 de la *Charte*. La garantie de *transparence* des outils d'IA est essentielle afin d'assurer à l'accusé son droit à une défense pleine et entière. Pour exercer ce droit, l'accusé doit avoir accès à toute l'information nécessaire pour faire valoir sa défense et répondre à l'infraction dont il est accusé[118]. Le fonctionnement opaque des outils d'IA le prive de ce droit. Incidemment, l'exercice de ses autres droits constitutionnels est également menacé par cette opacité; la *présomption d'innocence* (art. 11d) de la *Charte*), le *droit à la non-discrimination* (art. 15 de la *Charte*), la *protection contre la détention*

---

[112] Citizen Lab, p. 18-25 et 85; *Cf.* également Citizen Lab, à la Section 2.2 ("Bias and Inaccuracies in Police Data").
[113] Rapport spécial sur la Technologie de RF, par. 80-85.
[114] *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, chapitre A-2.1, art. 72. **(Québec)**
[115] Céline CASTETS-RENARD, Émilie GUIRAUD et Jacinthe AVRIL-GAGNON, « Cadre juridique applicable à l'utilisation de la reconnaissance faciale par les forces de police dans l'espace public au Québec et au Canada », Observatoire international sur les impacts sociétaux de l'IA et du numérique, Chaire de recherche I.A. responsable à l'échelle mondiale, 2020, p. 41.
[116] *Freedom of Information and Protection of Privacy Act*, RSA 2000, c F-25, s 35 **(Alberta)**; *Freedom of Information and Protection of Privacy Act*, RSBC 1996, c 165, s 28 **(C.-B.)**; *Local Authority Freedom of Information and Protection of Privacy Act*, SS 1990-91, c L-27.1, s 26 **(Saskatchewan)**.
[117] Citizen Lab, p. 86, cite à titre d'exemple *Freedom of Information and Protection of Privacy Act*, RSO 1990, c F.31, art. 40(3) **(Ontario)** et *Municipal Freedom of Information and Protection of Privacy Act*, RSO 1990, c M.56, art. 30(3) **(Ontario)**.
[118] *R.* c. *Seaboyer*, [1991] 2 SCR 577; Anastasia KONINA, "The Privatization of Law Enforcement: Promoting Human Rights through Procurement Contracts", *McGill GLSA Research Series*, *1*(1), 1-36. p. 16-18:« The opacity of algorithms, also referred to as the black box problem, strongly suggests that the law enforcements' use of technology is incompatible with the *Charter* right to make full answer and defence. » Nous reviendrons plus en détails sur ce point dans la Partie III sur le droit de la preuve.

ou l'arrestation arbitraire (art. 9 de la *Charte*), le *droit à la réparation* (art. 24(1) de la *Charte*) sont parmi les droits constitutionnels touchés par l'opacité technique des outils d'IA.

La pleine transparence est également essentielle pour assurer l'*imputabilité* et la responsabilité des décideurs (policier et juge) qui doivent être conscients du fonctionnement de l'outil d'IA et de ce sur quoi reposent ses recommandations[119]. La structure de l'outil d'IA est elle-même normative[120] : elle est le fruit de décisions humaines, d'une réflexion consciente ou non, de choix politiques ou éthiques dans l'ordonnancement des paramètres et qui, à travers sa recommandation, vient influencer et contraindre le décideur. En ce sens, son architecture doit aussi pouvoir faire l'objet d'un débat contradictoire, sans quoi la décision de justice se retrouverait usurpée par les choix du concepteur de l'algorithme. Comme l'expliquent les auteurs Antoine Garapon et Jean Lassègue, « [s]i elle ne veut pas passer pour une justice divinatoire, aussi mystérieuse et intimidante que les oracles antiques, la justice prédictive doit rendre publics ses algorithmes et ne pas se réfugier derrière le secret de fabrication (ce qui implique de changer le droit d'auteur) (…) [car] sans espace de contradiction, il n'y a plus de droit possible. »[121] Pour ces raisons, nos prochains efforts législatifs devraient reconnaitre une certaine garantie de *transparence* à la personne qui se retrouve en interaction avec l'institution pénale en raison de l'utilisation d'une technologie d'IA : un droit d'accès au code source et un droit à une explication intelligible quant à son fonctionnement.

Actuellement, les principaux obstacles au respect de cette garantie de *transparence* au Canada sont **(i)** le secret commercial qui empêche de rendre le code source accessible à tous, **(ii)** l'absence de formation et de connaissance par les parties au procès de son fonctionnement (« *technical illiteracy* ») et **(iii)** le fonctionnement même de l'apprentissage-machine des outils d'IA qui, avec le temps, fait évoluer la structure du code au point où elle finit par échapper à celui qui l'utilise, et parfois même au codeur[122].

Il n'existe pas pour l'instant de loi au Canada exigeant des vendeurs d'un outil IA de dévoiler le code source de l'outil qu'il fournit aux services de police. Le secret commercial, la propriété intellectuelle et les normes en matière de brevet semblent être les principales limites à cette garantie de *transparence*. Par exemple, l'*Accord Canada–États-Unis–Mexique* (C. 19, art. 19.16) empêcherait d'imposer au fournisseur de la technologie une exigence préalable de dévoilement du code source avant son importation. C'est donc uniquement *a posteriori* – c'est-à-dire après l'atteinte aux droits de l'accusé, dans le cadre d'une enquête spécifique – que le code source pourrait être consulté et ce, à certaines conditions[123]. Comme le révèle l'arrêt *Aithaqi* c. *R*., la propriété intellectuelle peut constituer un obstacle au droit de l'accusé à une défense pleine et entière[124]. Dans cet arrêt, le laboratoire d'expertise médico-légale était réticent au fait de dévoiler l'étude complète de validation interne de l'outil d'IA *STRmix*™[125] qu'il avait réalisée, afin de protéger sa propriété intellectuelle. Par conséquent, l'accusé devait requérir une ordonnance de communication de la preuve auprès du juge et supportait le fardeau de prouver la « vraisemblable pertinence » de l'accès à cette étude

[119] Citizen Lab, p. 129-132.
[120] Lawrence LESSIG, *Code and Other Laws of Cyberspace*, New York, Basic Books, 1999; Joel R. REIDENBERG, *Lex Informatica: The Formulation of Information Policy Rules through Technology* , 76 *Tex. L. Rev*. 553, 1997; L. WINNER, *La baleine et le réacteur. À la recherche des limites de la haute technologie*, 1986, trad. par M. Puech, Éditions Descartes & Cie, Paris, 2002.
[121] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 242.
[122] Angèle CHRISTIN, "Predictive algorithms and criminal sentencing", p. 283 dans N. Guilhot et D. Bessner (dir.), *The Decisionist Imagination*, Berghahn Books, 2018.
[123] Citizen Lab, p. 131-132.
[124] 2020 QCCS 870, par. 37; CDO2, p. 26 et 35.
[125] Cet outil sera présenté dans la partie III sur le droit de la preuve.

complète de validation interne. Même s'il ne s'agit pas d'un fardeau onéreux, il s'agit à notre avis d'un fardeau *inusité* à faire porter sur l'accusé alors que celui-ci se retrouve devant une technologie nouvelle, complexe et dont le fonctionnement demeure, pour le commun des mortels, incompréhensible. Face au fonctionnement opaque de ces nouvelles technologies, l'accusé peut tout à fait ne pas savoir par où commencer sa défense, ce qui l'amènerait tout naturellement à exiger des documents-sources telle l'étude complète de validation interne. Or, une telle investigation pourrait représenter aux yeux du juge une « recherche à l'aveuglette » d'éléments de preuve ce qui est interdit par l'arrêt *Gubbins*[126]. Nous verrons dans la partie III comment cette garantie de *transparence* pourrait être renforcée par une réforme du droit de la preuve.

3. Principes généraux du droit : Droits constitutionnels et garanties en matière criminelle à l'épreuve des possibilités techniques de l'IA

3.1. Droit à la vie privée

En matière de droit à la vie privée, le fédéral et les provinces ont compétence à l'égard des organisations qui sont sous leur juridiction. Il existe donc des lois différentes en matière de protection des renseignements personnels pour les organisations publiques et privées dépendamment si elles sont fédérales ou provinciales. Or, les principes organisant le droit à la vie privée sont sensiblement les mêmes au niveau fédéral et dans les provinces. La particularité des lois en matière de renseignement à la vie privée s'appliquant au secteur privé est que c'est la *Loi* fédérale *sur la protection des renseignements personnels et les documents électroniques*, L.C. 2000, ch. 5 qui s'applique à toutes les provinces qui n'ont pas édicté une loi « essentiellement similaire » à celle-ci. À ce jour, seuls le Québec, la Colombie-Britannique et l'Alberta ont adopté une loi pour régir le secteur privé qui a été jugée essentiellement similaire à la loi fédérale.

Malgré un libellé qui semble se restreindre à interdire les fouilles et les perquisitions abusives par l'État, l'article 8 de la *Charte* confère une protection *morale* plus large aux personnes « contre les intrusions injustifiées de l'État dans leur vie privée »[127]. En ce sens, la protection de l'article 8 englobe « trois sphères de revendications du droit à la vie privée », trois expressions particulières de ce droit : **(i)** la vie privée relative au corps (sphère *personnelle*), **(ii)** la vie privée relative aux lieux (sphère *spatiale*) et **(iii)** la vie privée relative aux renseignements personnels (sphère *informationnelle*)[128]. Cette manière de catégoriser ces sphères non étanches de revendication permet d'illustrer toute la latitude, insoupçonnée, de la protection du droit à la vie privée : un droit riche, complexe et fragmenté. Complexe, car il bénéficie à la fois d'une protection *constitutionnelle* et de protections *législatives* particulières. Fragmenté, puisqu'en raison du partage constitutionnel des compétences entre les provinces et le gouvernement fédéral[129], le droit à la vie privée est protégé à la fois par nos lois *provinciales* et *fédérales* dans leurs champs de compétence respectifs et, il est une fois de plus fragmenté, puisque les exigences diffèrent dépendamment si l'on s'adresse à une institution *publique* ou *privée*. Le fonctionnement des outils d'IA étant garanti par la collecte, la conservation et l'utilisation massive de renseignements personnels, nous nous intéresserons particulièrement à la *sphère informationnelle* du droit à la vie privée. Nous présenterons, tout d'abord,

[126] 2018 CSC 44, par. 29.
[127] *Hunter* c. *Southam*, [1984] 2 R.C.S. 145, p. 160.
[128] Karim BENYEKHLEF et Pierre-Luc DÉZIEL, *Le droit à la vie privée en droit québécois et canadien*, Montréal, Éditions Yvon Blais, 2018 [ci-après : « Benyekhlef et Déziel »] se référant à la classification proposée dans *R*. c. *Dyment*, [1988] 2 RCS 417.
[129] *Loi constitutionnelle de 1867*, 30 & 31 Victoria, ch. 3 (R.-U.), art. 91 et 92.

(3.1.1) la protection *législative* de la vie privée informationnelle s'imposant aux institutions *publiques*, autant *fédérales que provinciales*, comme les services de police, et les obligations en matière de protection des données qui incombent aux entreprises *privées* lors de leur interaction avec la police. Ensuite, **(3.1.2.)** nous nous intéresserons à la protection *constitutionnelle* de la vie privée. Pour les fins du présent rapport, nous nous intéresserons surtout à l'obligation d'obtenir un mandat dans le cadre d'une enquête policière afin de collecter des renseignements personnels sur les réseaux sociaux à l'aide d'un outil de surveillance algorithmique.

3.1.1. La protection législative de la vie privée informationnelle

En principe, le droit à la vie privée regroupe différentes protections portant sur les « renseignements personnels » à l'étape de leur collecte, leur utilisation et leur partage par les différents organismes. Si le libellé des lois varie d'une province à l'autre, les principes qui soutiennent l'organisation de ces lois sont similaires. Les « renseignements personnels » comprennent tout renseignement « identificatoire », ce qui inclut toutes les données ou informations « qui porte[nt] sur une personne identifiée ou qui permet[tent] de l'identifier. »[130] En respectant cette logique, une donnée qui, à elle seule, ne permettrait pas d'identifier un individu mais qui, une fois traitée par l'outil d'IA et couplée à d'autres renseignements collectés par l'algorithme, permettrait de l'identifier pourrait être considérée comme un « renseignement personnel » devant être couvert par les garanties de la loi[131].

Panorama général des protections*. Organisme public* – La personne peut généralement s'attendre à ce **(i)** que la collecte de ses renseignements ait un lien direct avec les activités de l'organisme public et ne concerne que ce qui est nécessaire aux fins de celles-ci[132], i.e. « essentielle » et non simplement commode[133]; **(ii)** qu'ils soient utilisés aux fins pour lesquelles ils ont été collectés ou pour des fins *compatibles* à celles-ci[134]; **(iii)** que la collecte soit effectuée directement auprès de la personne concernée et que celle-ci soit avisée de

[130] Benyekhlef et Déziel, p. 262; **Organisme public fédéral** : *Loi sur la protection des renseignements personnels*, L.C. (1985), ch. P-21, art. 3 (LPRP) qui s'applique à la GRC; **Organisme public provincial** : *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, RLRQ, c. A-2.1, art. 54 et la *Loi concernant le cadre juridique des technologies de l'information*, RLRQ, c. C-1.1 **(Québec)**; *Freedom of Information and Protection of Privacy Act* (FOIP), R.S.A. ch. F-25, art. 1 **(Alberta)**; *Freedom of Information and Protection of Privacy Act* (FIPPA), R. S.B.C. 1993, ch. 165, s. 1 **(Colombie-Britannique)**; *Freedom of Information and Protection of Privacy Act* (FIPPA), R. S.P.E.I. 1988, ch. F-15.01, art. 1.1. et 2 **(Île du Prince-Édouard)**; *Loi sur l'accès à l'information et la protection de la vie privée*, C.P.L.M., ch. F175, art. 1 **(Manitoba)**; *Loi sur le droit à l'information et la protection de la vie privée*, L.N.-B. 2009, ch. R-10.6, art. 1 **(Nouveau-Brunswick)**; *Freedom of Information and Protection of Privacy Act*, S.N.S. 1993, ch. 5, art. 4(i) **(Nouvelle-Écosse)**; *Loi sur l'accès à l'information et la protection de la vie privée*, L.R.O. 1990, ch. F.31**(Ontario)**; *Access to Information and Protection of Privacy Act*, S.N.L. 2015, ch. A-1.2, art.2 **(Terre-Neuve et Labrador)**; *Access to Information and Protection of Privacy Act*, S.S. 1990-91, ch. F-22.01, art. 24(1)a) **(Saskatchewan)**

[131] Benyekhlef et Déziel, p. 266, et à la p. 268 où on nous explique que le seuil au-delà duquel une information indirecte peut être considérée comme identificatoire n'est pas encore défini clairement en droit canadien, cf. *Gordon* c. *Canada (Santé)*, 2008 CF 258.

[132] **Organisme public fédéral** : *Loi sur la protection des renseignements personnels*, L.C. (1985), ch. P-21, art. 4 (interprété dans l'affaire Clearview); **Organisme public provincial** : *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, art. 64. **(Québec)**. Le projet de loi 64 adopté par le gouvernement du Québec prévoit désormais que la collecte devra être précédée d'une évaluation des facteurs relatifs à la vie privée et s'effectuer dans le respect d'une entente établie avec la CAI.

[133] *M.L.* c. *Gatineau (Ville de)*, 2010 QCCA168.

[134] **Organisme public fédéral** : *Loi sur la protection des renseignements personnels*, L.C. (1985), ch. P-21, art. 7. **Organisme public provincial** : *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, art. 65.1. **(Québec)**

la collecte et des fins de la collecte[135], sauf exceptions[136]; **(iv)** que la divulgation à d'autres organismes n'est permise que sous certaines conditions propres à chaque province, par exemple lorsqu'elle est « nécessaire » à l'application d'une loi[137]. ***Organismes privés* –** Actuellement, un corpus de lois fédérales et provinciales s'appliquant aux entreprises privées garantit que **(i)** le *consentement* de la personne devra être obtenu avant la collecte, l'utilisation ou la communication des renseignements personnels[138], sauf exception, comme lorsque ces renseignements sont déjà accessibles au public[139]. Autrement, des renseignements personnels détenus par un organisme privé peuvent être divulgués aux policiers lorsqu'ils ont l'autorité légale de l'obtenir, comme un mandat obtenu auprès d'un juge[140]. La loi fédérale prévoit également **(ii)** que « l'organisation ne peut recueillir, utiliser ou communiquer des renseignements personnels qu'à des fins

---

[135] **Organisme public fédéral** : *Loi sur la protection des renseignements personnels*, L.C. (1985), ch. P-21, Art. 5(1)(2); Benyekhlef et Déziel, p. 318

[136] **Organisme public fédéral** : *Loi sur la protection des renseignements personnels*, L.C. (1985), ch. P-21, Art. 8(2); **Organisme public provincial** : *Freedom of Information and Protection of Privacy Act* (FOIP), R.S.A. ch. F-25, art. 34 **(Alberta)**; *Freedom of Information and Protection of Privacy Act* (FIPPA), R.S.B.C. 1993, ch. 165, art. 26 **(Colombie-Britannique)**; Au Québec, cette règle « ne s'applique pas à une enquête de nature judiciaire, ni à une enquête ou à un constat faits par un organisme qui, en vertu de la loi, est chargé de prévenir, détecter ou réprimer le crime ou les infractions aux lois. », *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels,* art. 65 al.5. **(Québec)**

[137] Par exemple, la *Loi sur l'accès à l'information et la protection de la vie privée*, C.P.L.M., ch. F175, art. 44(r) prévoit qu'un organisme public peut communiquer des renseignements personnels sans consentement aux fins de la prévention du crime et de l'application de la loi **(Manitoba)**; *Freedom of Information and Protection of Privacy Act*, R.S.B.C. 1993, ch. 165, art. 33.2(i) : "A public body *may* disclose personal information referred to in section 33 inside Canada as follows: (i) to a public body or a law enforcement agency in Canada to assist in a *specific* investigation" et à l'art. 33.1 (1)(t) (interprétés dans l'affaire Denham): "to comply with a subpoena, a warrant or an order issued or made by a court, person or body in Canada with jurisdiction to compel the production of information." **(C.-B.)**; *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels,* Art. 67 : « Un organisme public *peut*, sans le consentement de la personne concernée, communiquer un renseignement personnel à toute personne ou organisme si cette communication est *nécessaire* à l'application d'une loi au Québec » et, le nouveau projet de loi 64 adopté par le gouvernement, prévoit, notamment, que cette communication doit être « 1° (…) prévue expressément par la loi » et si elle « 2° n'est pas prévue expressément par la loi, est *ponctuelle* et, s'il y a aussi communication de renseignements personnels concernant toute autre personne, lorsque ces renseignements n'en concernent qu'un nombre *restreint* ». **(Québec)**

[138] **Loi fédérale sur les organismes privés** : *Loi sur la protection des renseignements personnels et les documents électroniques* du Canada, L.C. 2000, ch. 5, art. 5(1) et principe #3 de l'Annexe A (LPRPDE) qui s'applique secteur privé si la province n'a pas adopté une loi « essentiellement similaire » régissant le secteur privé. **Lois provinciales particulières encadrant les organismes privés et remplaçant la loi fédérale dans ces provinces** : *Loi sur la protection des renseignements personnels dans le secteur privé*, RLRQ, c. P-39.1, art. 14 (LPRPSP) et la *Loi concernant le cadre juridique des technologies de l'information*, RLRQ, c. C-1.1 (**Québec**); *Personal Information Protection Act*, S.B.C. 2004, ch. 63 (Colombie-Britannique); *Personal Information Protection Act,* S. A. 2003, ch. P-6.5 **(Alberta)**. **N'ont pas été considérées comme essentiellement similaire à la loi fédérale** : *Privacy Act*, R.S.N.L. 1900, ch. P-22 **(Terre-Neuve-et-Labrador)**; *Loi sur la protection de la vie privée*, C.P.L.M., ch. P125 **(Manitoba)**; *Privacy Act*, R. S.S. 1978, ch. P-24 **(Saskatchewan)**

[139] **Loi fédérale sur les organismes privés** : *Loi sur la protection des renseignements personnels et les documents électroniques* du Canada, L.C. 2000, ch. 5, art. 7(1)d) (interprété dans l'affaire Clearview). **Lois provinciales particulières encadrant les organismes privés** : *Personal Information Protection Act*, S.B.C. 2004, ch. 63, art. 12(1)*e*), 15(1)*e*) et 18(1)*e*) **(C-B);** *Personal Information Protection Act,* S. A. 2003, ch. P-6.5, art. 14*e*), 17*e*) et 20*j*) **(Alberta)**

[140] **Loi fédérale sur les organismes privés** : *Loi sur la protection des renseignements personnels et les documents électroniques*, L.C. 2000, ch. 5, art. 7(3)(c.1) (interprété dans l'arrêt Spencer). **Lois provinciales particulières encadrant les organismes privés** : *Loi sur la protection des renseignements personnels dans le secteur privé*, RLRQ, c. P-39.1, Art. 18(3) : « Une personne qui exploite une entreprise peut, sans le consentement de la personne concernée, communiquer un renseignement personnel contenu dans un dossier qu'elle détient sur autrui: 3° à un organisme chargé en vertu de la loi de prévenir, détecter ou réprimer le crime ou les infractions aux lois, qui *le requiert dans l'exercice de ses fonctions*, si le renseignement est *nécessaire* pour la poursuite d'une infraction à une loi applicable au Québec ». **(Québec)** Ici, nous devons lire cet article à la lumière de l'interprétation qu'a fait la Cour suprême dans *R.* c. *Spencer,* (2014) 2 R.C.S. 212 de l'article 7(3)(c.1) de la loi fédérale. Les entreprises ont une obligation fiduciaire envers leurs clients et doivent protéger leurs renseignements personnels. Par conséquent, ils ne peuvent divulguer volontairement ces informations en l'absence d'un mandat légal obtenu par la police auprès d'un juge. En somme, l'article 18(3) doit être lu de manière à permettre à l'entreprise de communiquer des renseignements sans le consentement de la personne concernée lorsque la police québécoise le requiert dans l'exercice – *légal* – de ses fonctions, donc dans le respect de la primauté du droit, du droit à la vie privée prévu à la *Charte* et d'un mandat en bonne et due forme obtenu préalablement auprès d'un juge.

qu'une personne raisonnable estimerait acceptables dans les circonstances »[141] et **(iii)** que ces informations devront être protégées[142]. ***Protections particulières des données biométriques*** – Les données biométriques, comme celles comprises dans la représentation d'un visage, sont généralement caractérisées comme des « renseignements sensibles » et constituent des « renseignements personnels »[143]. La seule loi traitant au Canada directement des données biométriques est la *Loi concernant le cadre juridique des technologies de l'information* (« LCCJTI ») qui s'applique dans la province de Québec. L'organisme privé ou public qui constitue une base de données comportant un système biométrique doit non seulement **(i)** « obtenir le consentement exprès de la personne concernée à ce que la vérification ou la confirmation de leur identité soit fait à l'aide d'un procédé de reconnaissance faciale » (art. 44 LCCJTI) mais elle doit également **(ii)** « déclarer la création ou l'existence du système biométrique à la CAI » (art. 45 LCCJTI). La CAI peut interdire la mise en service de cette banque de données, exiger sa destruction ou ordonner des changements. On précise également que « tout autre renseignement concernant une personne qui pourrait être découvert à partir des caractéristiques ou mesures biométriques ne peut servir à fonder une décision à son égard. »[144] La CAI a d'ailleurs publié en juillet 2020 des principes et obligations contraignant les organisations publiques et les entreprises quand ils ont recours à des banques de données biométriques. Ces principes s'organisent autour de trois thèmes : « analyse préliminaire et proportionnalité de la récolte », « déclaration à la CAI » et « consentement exprès »[145]. ***Protections particulières face aux technologies d'IA*** – Au Québec, le projet de loi 64 adopté par le gouvernement de la province en septembre 2021 a créé deux nouvelles protections face à certaines technologies d'IA utilisées par nos services publics. ,Ces protections constituent une première au Canada. **(i)** Premièrement, le citoyen concerné doit en être avisé si l'organisme public a recours à une technologie de profilage et doit être informé des moyens qu'il a pour la désactiver[146]. **(ii)** Deuxièmement, « un organisme public qui utilise des renseignements personnels afin que soit rendue une décision fondée exclusivement sur un traitement automatisé de ceux-ci doit, au moment de la décision ou avant, en informer la personne concernée. » Si la personne le demande, d'autres informations sur le fonctionnement de l'outil d'IA devront lui être divulguées.[147]. À ce jour, on ignore comment ces protections s'appliqueront afin de baliser le recours par les policiers québécois d'outils d'IA de prédiction, de surveillance ou de RF.

Collecte d'informations sur Internet. ***Par l'organisme public directement*** – Dans une première enquête spéciale datant de 2013 – l'Affaire *Blackstock*, le *Commissariat à la vie privée* avait déjà statué que les informations accessibles sur une page Facebook personnelle constituaient des renseignements personnels protégés par la *Loi sur la protection des renseignements personnels :* « Under the Act, restrictions on the collection of personal information apply, whether the personal information is available publicly or not. »[148]

---

[141] **Loi fédérale sur les organismes privés** : *Loi sur la protection des renseignements personnels et les documents électroniques*, L.C. 2000, ch. 5, art. 5(3).
[142] **Loi fédérale sur les organismes privés** : *Id*., art. 5(1) et les 10 principes de l'Annexe A.
[143] Enquête sur Clearview AI, Inc.; Rapport spécial sur la technologie de RF.
[144] Rapport spécial sur la Technologie de RF donne pour exemple les renseignements contextuels découverts sur internet à partir de l'hyperlien rattaché à l'image, utilisée par l'outil de RF, qui mène vers l'adresse Internet où elle a été extraite.
[145] COMMISSION D'ACCÈS À L'INFORMATION DU QUÉBEC, « Biométrie : principes à respecter et obligations légales des organisations. Guide d'accompagnement pour les organismes publics et les entreprises », 2020, en ligne : https://www.cai.gouv.qc.ca/biometrie/pour-davantage-dinformation/
[146] *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, art. 65.0.1
[147] *Id*., art. 65.2.
[148] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, "Aboriginal Affairs and Northern Development Canada wrongly collects information from First Nations activist's personal Facebook page", en ligne : Aboriginal Affairs and Northern Development Canada wrongly collects information from First Nations activist's personal Facebook page - Office of the Privacy Commissioner of Canada.

Ainsi, les organismes publics mis en cause dans cette enquête, soit le *Ministère de la justice du Canada* et *Affaires autochtones et du Nord Canada*, ne pouvaient collecter les renseignements accessibles sur la page Facebook personnelle d'une activiste autochtone car ces renseignements n'étaient pas directement et précisément liés à leur programme ou à leurs activités : « not obviously relevant to policy development by AANDC, as the department contended, or to the human rights lawsuit with which the Department of Justice was particularly concerned. »[149]. Le *Commissariat* y a vu également une atteinte à « l'esprit de la loi, si ce n'est à la lettre même de la loi » (notre traduction), en raison d'une contravention au principe de transparence qui la sous-tend; nous en déduisons alors que la collecte a été faite sans qu'elle ait été avisée et sans qu'elle ait été effectuée directement auprès de la personne concernée[150]. ***Par un organisme public auprès d'un organisme privé*** – Conformément à cette décision, le *Commissaire* a réitéré lors de *l'Enquête conjointe sur Clearview AI, Inc*. que les renseignements personnels accessibles sur Internet ne constituaient pas des renseignements de nature publique; il faut savoir que seuls des renseignements « auxquels le public a accès » peuvent être collectés sans obtenir de consentement[151]. Ainsi, il a été décidé qu'un organisme privé, comme *Clearview AI. Inc*, ne pouvait collecter en toute légalité ces images afin d'alimenter leur outil de RF sans le consentement des personnes concernées. *A fortiori*, un organisme public, comme la GRC, ne pourrait pas non plus chercher à obtenir ces informations auprès d'un tiers privé lorsque celui-ci les aurait collectées illégalement. En effet, une telle collecte enfreindrait alors l'article 4 de la *Loi* fédérale *sur la protection des renseignements personnels,* L.R.C. (1985), ch. P-21, qui concerne les organismes publics fédéraux, et qui les oblige à ne collecter que les renseignements qui ont un « lien direct avec ses programmes ou ses activités ». Le *Commissariat* a interprété cette disposition de manière à s'assurer que la collecte effectuée par l'organisme public n'entretienne de lien qu'avec des « programmes ou des activités » qui sont « légaux », afin de s'assurer que l'organisme public respecte activement le principe constitutionnel non-écrit de la PRIMAUTÉ DU DROIT: « Conclure autrement reviendrait à permettre aux institutions fédérales de réaliser leur mandat tout en récompensant les organisations dont les pratiques de collecte de renseignements personnels sont illégales et notamment non conformes aux lois canadiennes sur la protection des renseignements personnels. »[152] Ce principe constitutionnel est consacré expressément dans le préambule de la *Loi constitutionnelle de 1982* et a été reconnu par la Cour suprême comme faisant partie implicitement du préambule de la *Loi constitutionnelle de 1867*[153].Un organisme public fédéral, comme la GRC, aurait donc l'obligation de s'assurer de la légalité des pratiques en matière de collecte de renseignements personnels du tiers privé avec qui il veut faire affaire. Cette obligation constituerait donc une limite au pouvoir général d'enquête de la GRC[154]. Étant donné que l'organisme public doit veiller proactivement au respect de la primauté du droit, l'organisme public aurait également l'obligation d'évaluer les risques de la technologie de RF et de vérifier sa conformité avec les principes de *common law* et les droits et libertés prévus à la *Charte canadienne*[155].

Partage de données entre organisations : la problématique du trans-fonctionnalisme. Le contexte actuel est caractérisé par la convergence des fonctions institutionnelles – nous pensons, entre autres, aux

[149] *Id.*
[150] *Id*.
[151] Enquête conjointe sur Clearview AI, Inc., par. 44-47.
[152] Enquête conjointe sur Clearview AI, Inc, par. 22 et 26-27.
[153] *Colombie-Britannique* c. *Imperial Tobacco Canada Ltée*, 2005 CSC 49, par. 57; *Renvoi relatif aux droits linguistiques au Manitoba*, [1985] 1 R.C.S. 721.
[154] Rapport spécial sur la Technologie de RF, par. 26, citant l'art. 18 de la *Loi sur la Gendarmerie royale du Canada* et l'al. 14(1)*a*) du *Règlement de la Gendarmerie royale du Canada.*
[155] Rapport spécial sur la Technologie de RF, par. 42.

services policiers appelés à jouer un rôle « communautaire » et « social » et aux services sociaux, comme le système de santé, – et, conséquemment, par des partenariats de plus en plus fréquents entre plusieurs organismes publics et entre les organismes publics et des tiers privés. Cela pose le problème du partage de données entre des organisations qui ont des fonctions divergentes. Les pratiques policières actuelles, ou raisonnablement prévisibles, nous amènent à relever les enjeux liés à l'utilisation d'un outil d'IA dont le fonctionnement serait assuré grâce à une banque de données privée ou à partir de renseignements détenus initialement par d'autres organismes publics (système de santé, services sociaux, système de protection de l'enfance, etc.).

Partage Public-Public – En général, un organisme public peut partager des renseignements personnels avec un autre organisme public, sans le consentement de la personne, si le partage est fait aux fins de la collecte initiale, s'il est fait pour un usage qui est *compatible* avec les fins de la collecte initiale ou lorsque le partage est autorisé par une autre loi[156]. Autrement dit, l'utilisation des renseignements personnels doit se limiter aux fins pour lesquelles ils ont été collectés[157]. Chaque loi provinciale prévoit néanmoins des situations particulières, très précises, où un partage de renseignements est autorisé[158]. Aux yeux des rapporteurs du *Citizen Lab*, la conception d'outils d'IA aux fins de police prédictive qui sont axés sur une approche intersectorielle (modèle Hub), comme ceux possiblement en développement par le SPPAL de la police de Saskatchewan, représente un risque accru de partage de renseignements personnels, parfois très sensibles, qui, d'ailleurs, n'avaient pas été collectés initialement pour servir à ces fins. Une telle approche pourrait également avoir un effet contreproductif puisqu'elle serait susceptible de miner la confiance du public envers les services sociaux publics et de dissuader ceux qui en ont besoin d'y recourir[159]. En ce sens, d'autres chercheurs au Canada craignent que les renseignements personnels détenus par les organismes publics (par exemple, « passeports, visa, permis de travail, d'étude ou de conduire ») finissent par être recyclés pour servir aux enquêtes policières[160]. À ce sujet, la Commissaire Denham, dans le cadre d'une enquête spéciale menée par la *Commissariat à l'information et à la protection de la vie privée de la Colombie-Britannique* en 2012, s'est inquiétée du recyclage informationnel aux fins d'enquête policière qu'avaient fait subir les policiers à la banque de données des permis de conduire de la province. Sans que la police n'ait eu à obtenir de mandat et sans même avoir de piste précise sur l'identité des vandales recherchés, l'assureur automobile de la province aurait volontairement permis aux policiers d'utiliser l'*entièreté* de la banque de données de permis de conduire de la province ainsi que la technologie de RF détenue par l'assureur public dans le but de permettre à la police de découvrir l'identité des vandales. Évidemment, la Commissaire a reconnu que cette méthode contrevenait à la *Freedom of information and protection of privacy act*[161]. En effet, il s'agit d'une utilisation des renseignements personnels des détenteurs de permis de conduire qui est différente et

---

[156] Benyekhlef et Déziel, p. 320-322. **Organisme public fédéral** : *Loi sur la protection des renseignements personnels*, L.C. (1985), ch. P-21, art. 8(2)

[157] **Organisme public fédéral** : *Loi sur la protection des renseignements personnels*, L.C. (1985), ch. P-21, art. 7. **Organisme public provincial** : *Loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels*, art. 65.1 et 65.3. **(Québec)**

[158] Benyekhlef et Déziel, tableau 4.6, p. 324-325.

[159] Citizen Lab, p. 81-83.

[160] Céline CASTETS-RENARD, Émilie GUIRAUD et Jacinthe AVRIL-GAGNON, « Cadre juridique applicable à l'utilisation de la reconnaissance faciale par les forces de police dans l'espace public au Québec et au Canada », Observatoire international sur les impacts sociétaux de l'IA et du numérique, Chaire de recherche I.A. responsable à l'échelle mondiale, 2020, p. 41. Par exemple au Québec, cela pourrait contrevenir à l'art. 65.1 de la *loi sur l'accès aux documents des organismes publics et sur la protection des renseignements personnels.*

[161] INFORMATION AND PRIVACY COMMISSIONER, *Investigation Report F12-01. Investigation into the use of facial recognition technology by the insurance corporation of British-Columbia*, [2012] B.C.I.P.C.D. No. 5.

incompatible avec les fins pour lesquelles l'assureur public les avait collectés initialement[162]. Étant donné que cette technologie de RF offrait aux services policiers un accès à *l'entièreté* de la banque de données de l'assureur public de la province, sa finalité se retrouvait *détournée* par l'usage qu'en ont fait les policiers. Ceux-ci ne pouvaient donc y avoir recours sans « l'autorité légale nécessaire », comme un mandat obtenu auprès d'un juge[163]. Généralement, les lois qui traitent de la divulgation d'informations entre organismes publics ne permettent la collaboration *volontaire* de l'organisme public avec les policiers que lorsque ces derniers formulent une demande d'information *unique* sur un usager préalablement identifié dans le cadre d'une enquête *précise*[164]. Ici, les policiers se sont servis de l'entièreté de la banque de données sans même avoir de piste quant à l'identité des vandales et comptaient sur cette banque de données pour leur en fournir une; cela équivalait alors à mettre sous enquête l'entièreté des détenteurs de permis de conduire de la province uniquement en raison du fait qu'ils détenaient un permis de conduire.

***Partage Privé-Public* –** Dans l'optique où un organisme public, comme un service de police, établirait un partenariat avec un organisme privé dans le but de développer un outil d'IA aux fins de police prédictive, le partage d'informations serait limité en vertu des lois particulières régissant les organismes privés : « Commercial privacy legislation in Canada (…) also does not authorize disclosure without consent to law enforcement unless law enforcement has "lawful authority" to access the information. »[165] Conformément à ce principe, il a été décidé, afin d'assurer un degré d'anonymat aux Canadiens lors de leurs activités sur Internet, que les internautes avaient une attente raisonnable en matière de vie privée à l'égard de leur adresse IP une fois associée à leur *nom*, leur *numéro de téléphone* ou à leur *adresse*. Ainsi, une entreprise privée ne pourrait pas divulguer *volontairement* à la police ces renseignements personnels sans que cette dernière n'obtienne un mandat judiciaire, au risque de commettre une infraction à la protection de la vie informationnelle protégée par l'article 8 de la *Charte*. Ce sont les conclusions de la Cour suprême dans l'arrêt *Spencer*[166]. L'article 7(3)c.1) de la *Loi* fédérale *sur la protection des renseignements personnels et les documents électroniques*, L.C. 2000, ch. 5, doit être interprétée comme exigeant des organismes privés, comme les fournisseurs de service Internet, l'obligation de protéger la vie privée de leurs clients[167].

Conclusion. En dépit du flou du cadre juridique actuel concernant les outils d'IA, les policiers ont une obligation de fond de veiller au respect de la primauté du droit et de la protection de la vie privée , qui repose sur l'exigence constitutionnelle de fonder toute intrusion dans la vie privée sur une autorisation et un contrôle judiciaire : « Allowing law enforcement agencies to access data they could not constitutionally obtain, through a private company that obtained the data legally, could represent an unconstitutional expansion of the state's ability to monitor and track individuals without justification or judicial oversight. »[168] Malgré la volonté de développer des stratégies de prévention plus efficaces, axées sur l'intersectorialité, le trans-fonctionnalisme et des partenariats entre institutions, il faut comprendre qu'*aux yeux de la personne concernée,* ces institutions et ces compagnies relèvent de fonctions différentes et sont,

---

[162] INFORMATION AND PRIVACY COMMISSIONER, *Investigation Report F12-01,* par. 111.
[163] Citizenlab, p. 81; Céline CASTETS-RENARD, Émilie GUIRAUD et Jacinthe AVRIL-GAGNON, « Cadre juridique applicable à l'utilisation de la reconnaissance faciale par les forces de police dans l'espace public au Québec et au Canada », Observatoire international sur les impacts sociétaux de l'IA et du numérique, Chaire de recherche I.A. responsable à l'échelle mondiale, 2020, p. 84. Comme le prévoit l'art. 33.1 (1)(t) *Freedom of Information and Protection of Privacy Act*, R.S.B.C. 1993, ch. 165 **(C.-B.)**
[164] 33.2(i) *Freedom of Information and Protection of Privacy Act*, R.S.B.C. 1993, ch. 165 **(C.-B.)**
[165] Citizen Lab, p. 84.
[166] *R.* c. *Spencer*, (2014) 2 R.C.S. 212
[167] Benyekhlef et Déziel, p. 188.
[168] Citizen Lab, p. 84.

dans son imaginaire, hermétiquement confinées. Ainsi, le consentement accordé à l'un de ces organismes peut ne pas correspondre aux fins nouvelles pour lesquelles on souhaite redestiner ces informations.

3.1.2. La protection constitutionnelle de la vie privée informationnelle

Après avoir étudié les protections législatives du droit à la vie privée, passons à la protection constitutionnelle de la vie privée, qui résulte de l'article 8 de la *Charte* : « chacun a droit à la protection contre les fouilles, les perquisitions ou les saisies abusives ». Ainsi, lorsque la personne peut raisonnablement s'attendre à ce que sa vie privée soit respectée par l'État, les policiers ont l'obligation d'obtenir un mandat délivré par un juge. Ce dernier est délivré sur la base de motifs raisonnables de croire qu'une infraction a été ou sera commise et que des renseignements liés à ce crime pourront être obtenus grâce à cette intrusion dans la vie privée (art. 487; 487.01; 184.2(3)[169]. En droit canadien, une fouille menée sans mandat est présumée abusive et il revient alors à la couronne de repousser ce caractère abusif selon la prépondérance des probabilités[170].

Respect du principe de proportionnalité. L'obtention d'un mandat permet de limiter l'intrusion à ce qui est *nécessaire* et *raisonnable* aux fins de la collecte. Cette exigence concerne directement l'exercice d'harmonisation entre certains droits concurrents : "The task of any section 8 analysis is to balance competing values : individual interests and rights against collective preferences and desire for security"[171]. Elle permet de respecter le critère de proportionnalité qui doit gouverner tout empiètement de l'État sur les droits et libertés individuels (art. 1er de la *Charte*) : « Le fait de demander un mandat et une autorisation du tribunal pourrait contribuer à faire en sorte qu'une utilisation proposée de la technologie [par exemple, de reconnaissance faciale] respecte le critère de proportionnalité. »[172] De même, dans l'enquête de la Commissaire Denham évoquée plus haut, l'exigence d'obtenir un mandat judiciaire pour utiliser la technologie de RF et la banque de données détenue par l'assureur public de la province permet d'assurer que : « any *change in use* [dans la finalité initiale de la collecte] of this magnitude is proportional to the public good served by the infringement on privacy rights of citizens. »[173].

Expectative de vie privée : Par-delà la frontière traditionnelle du public/privé. Au cours des dernières décennies, la frontière traditionnelle qui séparait l'espace « privé » de l'espace « public » s'est dissipée, notamment en raison de l'accroissement de la performativité des outils de surveillance. Pour cette raison, les décideurs ont eu besoin d'une nouvelle mesure afin d'appliquer les protections en matière de vie privée. La Cour suprême a alors reconnu que la quotidienneté contemporaine exigeait un certain degré d'anonymat et ce, même lorsqu'on agit « en public », « à la vue de tous » comme lorsqu'on interagit sur Internet[174]. Si l'expectative raisonnable de vie privée est une protection *variable*, elle devrait vraisemblablement

[169] Rapport spécial sur la Technologie de RF, par. 42-48; *Hunter* c. *Southam Inc.*, [1984] 2 R.C.S. 145, par. 27-29 concernant le mandat comme pré-condition à la validité de la fouille ou de la perquisition. *Cf.* également, *R.* c. *Collins*, [1987] 1 SCR 265, par. 34.
[170] *Hunter* c. *Southam Inc.*, [1984] 2 R.C.S. 145; *R. c. Collins*, [1987] 1 R.C.S. 265.
[171] Lee-Ann CONROD, "Smart Devices in Criminal Investigations: How Section 8 of the Canadian Charter of Rights and Freedoms Can Better Protect Privacy in the Search of Technology and Seizure of Information" (2019) 24 *Appeal: Rev Current L & L Reform* 115, p. 122; Citizen Lab, p. 78.
[172] Rapport spécial sur la Technologie de RF, par. 42-48 : « Avant de recourir à un tel service, un corps policier devrait à tout le moins se demander si cela est nécessaire à l'enquête qu'il mène, et si l'intérêt public particulier qui est recherché est proportionnel à l'intrusion subie. »
[173] INFORMATION AND PRIVACY COMMISSIONER, *Investigation Report F12-01. Investigation into the use of facial recognition technology by the insurance corporation of British-Columbia*, [2012] B.C.I.P.C.D. No. 5, par. 113.
[174] *Id*., par. 42; *R.* c. *Spencer*, 2014 CSC 43, par. 43-47; *R.* c. *Jarvis*, 2019 CSC 10, par. 41.

s'appliquer aux informations partagées sur les médias sociaux et qui sont collectées par une technologie d'IA. En effet, cette protection accrue se justifierait en raison du caractère hautement intrusif de ces technologies, compte-tenu de leur haute performativité, de la nature intime rattachée aux renseignements biographiques, historiques et détaillés que l'on retrouve sur Internet et de l'importance que l'on accorde collectivement à leur protection – comme en témoignent nos lois en matière de renseignements personnels[175]. À ce sujet, le *Commissariat à la protection de la vie privée* avait jugé, dans le cadre de son *enquête sur Clearview AI., Inc*., qu'un outil de RF « fondé sur l'extraction systématique et le traitement de milliards d'images de personnes innocentes de tout crime, constitue une intrusion majeure et substantielle par l'État dans la vie privée des Canadiens »[176]. Ces enseignements, à notre avis, s'appliquent également aux autres technologies de surveillance fonctionnant par IA, comme l'expliquent les rapporteurs du *Citizen Lab* :

> « The aggregation and analysis of metadata and other open source electronic records without judicial oversight could provide questionable access to information that the Supreme Court has said cannot be obtained through direct means. Privacy safeguards, including prior judicial authorization, are therefore necessary when law enforcement agencies collect and analyze content and metadata that is captured from online platforms or other environments where individuals operate freely with relative anonymity. »[177]

Concevoir les données sur les médias sociaux comme des ressources « publiques » accessibles à tous les services de police aux fins de surveillance ou d'enquête constitue une inversion du raisonnement : *la criminalité se retrouve présumée* et la surveillance des médias sociaux vise *ex post facto* à justifier cette intrusion de la vie privée au nom d'une application efficace et préventive de la loi criminelle. Cette inversion du raisonnement contrevient également à l'esprit de la présomption d'innocence protégée par la *Charte* (art. 11d)). Ce qui doit être préservé par l'article 8 de la *Charte*, c'est la nécessité de fonder toute intrusion dans la vie privée sur des « motifs raisonnables de croire » qu'une infraction a été commise ou sera commise et que des informations liées au crime pourront bel et bien être trouvées à la suite de cette intrusion : « Pre-emptive fishing expeditions [sur les médias sociaux] could hardly satisfy that standard. »[178]

Interception de communication privée en temps réel. *A fortiori*, nous joignons nos inquiétudes à celles exprimées par le *Citizen Lab* concernant les pratiques permises par un système d'IA comme l'*ICAC Child On-line Protection System*. Selon Citizen Lab, ce système permettrait de scanner et de retenir *en temps réel* de l'information de certains *chats privés*, ce qui constituerait une infraction en vertu des articles 184 (interception des communications) et 193 C.cr. (divulgation de renseignements) du *Code Criminel* (ci-après « C.cr. »). Afin de respecter l'article 8 de la *Charte*, il semble nécessaire d'obtenir un mandat en vertu des articles 185 et 186 C.cr.[179] (autorisation). En ce qui concerne la collecte de conversations archivées ou historiques, nous en traiterons dans la Partie III sur la collecte de la preuve par un outil d'IA.

3.2. Droit à l'égalité et à la protection contre la discrimination : encodage, reconduction et prolifération de la discrimination systémique par les outils d'IA

L'article 15 de la *Charte* garantit à tous le droit à la même protection et au même bénéfice de la loi. La loi criminelle doit s'appliquer également à tous, et les personnes ne peuvent faire l'objet d'une discrimination

[175] Rapport spécial sur la Technologie de RF, par. 42; Pour les critères de l'expectative de vie privée, *R.* c. *Jarvis*, 2019 CSC 10, par. 41.
[176] Rapport spécial sur la Technologie de RF, par. 46.
[177] Citizen Lab, p. 77.
[178] Citizen Lab, p. 78.
[179] Citizen Lab, à la p. 60 et à la note de bas de page 248.

en raison d'un des motifs énumérés à l'article 15 ou pour des motifs qui y sont analogues. Selon les rapporteurs du *Citizen Lab*, la protection constitutionnelle contre la discrimination s'étendrait aux *actions*, à la *conduite* et même à la *manière* que les policiers fédéraux, provinciaux et municipaux mènent leur enquête : « For instance, section 15 and human rights legislation would likely apply where a policing policy relies on a biased algorithm, or where an algorithmic "prediction" contributes to an officer discriminating against a member of a marginalized community. »[180] Une personne victime de discrimination au sens de la *Charte* en raison de la conduite des policiers pourrait alors réclamer des dommages en vertu de l'art. 24(1) de la *Charte*[181]. Dans *Elmardy* v. *Toronto Police Services Board*, le fait qu'il n'existait pas d'autres explications à la détention et la fouille abusive d'une personne noire que les préjugés raciaux conscients ou inconscients des deux policiers lui aura permis de réclamer à l'État des dommages compensatoires et punitifs pour cette atteinte au droit à la protection et au bénéfice égal de la loi en vertu de l'article 15 de la *Charte*[182]. Dans *Doe* v. *Metropolitan Toronto (Municipality) Commissioners of Police*, une femme a pu bénéficier d'une compensation financière en raison d'une atteinte à ses droits garantis à l'art. 24(1) de la *Charte* puisque les policiers avaient utilisé leur discrétion dans la *manière* de mener leur enquête, qui était de manière négligente et discriminatoire car teintée de préjugés sexistes[183]. Dans cette affaire, on a jugé que son droit à la liberté et à la sécurité (art. 7 de la *Charte*) ainsi que son droit à la non-discrimination (art. 15 de la *Charte*) avaient été enfreints par la conduite des policiers.

Le recours par la police aux outils d'IA soulève plusieurs enjeux quant au droit à l'égalité, si bien qu'il pourraitdonner droit à une compensation en vertu de la *Charte*. Leur algorithme peut reconduire, normaliser et même proliférer, du fait de sa haute capacité de traitement, l'historique des discriminations antérieurement vécues par certaines communautés. Cela est d'autant plus vrai lorsque l'algorithme est alimenté ou conçu à l'aide de données *historiques* qui peuvent être biaisées ou qui ont été récoltées dans le contexte de pratiques discriminatoires[184]. Même lorsque la race, le genre, ou la caractéristique de distinction prohibée n'est pas expressément prise en compte par l'algorithme, ce dernier peut produire des résultats discriminatoires s'il prend en compte des « *proxies* », i.e. un facteur autre que le motif de discrimination prohibé mais qui, étant fortement liée à la caractéristique de distinction prohibée, s'y substitue et agit comme si elle était prise en compte[185]. Certaines communautés, faisant déjà l'objet d'une sur-représentation dans les données policières et judiciaires au Canada, sont donc susceptibles de redevenir la cible première des interventions policières en raison de ces outils de police prédictive; celles-ci seraient alors l'objet d'une surveillance accrue, déraisonnable et injustifiée[186]. En retour, plus de surveillance amènera à davantage d'arrestations, ce qui,

---

[180] Citizen Lab, p. 104 se référant à *Elmardy* v. *Toronto Police Services Board*, 2017 ONSC 2074 et *Doe* v. *Metropolitan Toronto (Municipality) Commissioners of Police*, 39 OR (3d) 487, 160 DLR (4th) 697, 126 CCC (3d) 12.

[181] *Vancouver (Ville)* c. *Ward*, 2010 CSC 27; Tout en reconnaissant cette possibilité, une autrice relève la complexité et la difficulté d'effectuer une telle demande, Gabriella JAMIESON, « Using Section 24(1) Charter Damages to Remedy Racial Discrimination in the Criminal Justice System », 22 *Appeal: Review of Current Law and Law Reform* 71, 2017, CanLIIDocs 87, <https://canlii.ca/t/r9>; cf. Ranjan AGARWAL et Joseph MARCUS, "Where There is no Remedy, There is No Right: Using Charter Damages to Compensate Victims of Racial Profling" (2015) 34-1 *NJCL* 75, p. 89.

[182] *Elmardy* v. *Toronto Police Services Board*, 2017 ONSC 2074, par. 20, 23 et 40.

[183] *Doe* v. *Metropolitan Toronto (Municipality) Commissioners of Police*, 39 OR (3d) 487, 160 DLR (4th) 697, 126 CCC (3d) 12.

[184] Citizen Lab, p. 104-106. Le recours aux outils d'IA soulève d'autres préoccupations liées au droit à l'égalité : **(i)** il existe une sur-représentation de certains groupes sociaux dans les données cumulées par certains services publics sociaux, ce qui soulève des craintes quant aux outils d'IA utilisant des données trans-institutionnelles (par exemple, le modèle HUB du SPPAL et le CSA de l'Edmonton police), *Id*., p. 122; **(ii)** le choix même des crimes visés par les outils d'IA peuvent être discriminatoire. On remarque, par exemple, que les outils utilisés par les services de police canadiens sont davantage axés sur les crimes de rue ou de propriété plutôt qu'envers les crimes environnementaux ou financiers, *Id*., p. 115; **(iii)** nous craignons également l'encodage, la reconduction ou la multiplication des biais et préjugés inconscients des *concepteurs*, *Id*., p. 120-121.

[185] A. CHRISTIN, "Predictive algorithms and criminal sentencing", préc., note 123, p. 280-281.

[186] Citizen Lab, p. 107 et 109.

une fois les données de ces arrestations rentrées dans l'algorithme, se traduira par encore plus de surveillance dirigée à l'endroit de ces populations ou de ces quartiers (*effet cliquet*)[187].

3.3. Droit contre la détention ou l'arrestation arbitraire : instauration d'un soupçon généralisé

Nous craignons les effets insidieux que peuvent avoir, en raison de leur aura de scientificité, les outils de prédiction de criminalité sur l'appréciation des policiers et sur leurs interventions auprès des suspects. Comment distinguer, dans les motifs qui ont mené à la détention ou l'arrestation d'un suspect, ceux qui relèvent du jugement indépendant du policier de ceux qui reposent sur la prédiction offerte par l'outil d'IA ? L'article 9 de la *Charte* garantit la protection de tous contre l'arrestation ou la détention arbitraire. Une intervention arbitraire serait une intervention effectuée en l'absence de *motifs raisonnables*. Nous nous demandons alors si la prédiction de l'outil d'IA peut être considérée comme un « motif raisonnable » d'intervenir et si elle peut légalement motiver, en partie, l'intervention du policier auprès d'un suspect.

La détention pour fins d'enquête. Tout d'abord, la détention pour fins d'enquête ne doit pas être « arbitraire ». Elle doit donc se fonder sur un « soupçon raisonnable ». Pour être raisonnable, le soupçon doit reposer sur des faits *vérifiables* et *objectifs*. Au contraire, une pratique policière serait arbitraire et déraisonnable lorsqu'elle résulte de données inexactes, biaisées ou collectées dans le contexte de pratiques discriminatoires : « par définition, la détention fondée sur un profilage racial ne repose pas sur des soupçons raisonnables. »[188] Sachant que les prédictions des outils d'IA peuvent être produites à partir de données inexactes[189], que les données historiques peuvent avoir été collectées dans le contexte de pratiques discriminatoires, qu'en raison de leur méthode de traitement les outils d'IA peuvent décupler ces biais, et que l'algorithme lui-même pourrait refléter les biais discriminatoires du concepteur, il semble difficile de concevoir comment un policier qui utilise un tel outil dans sa pratique quotidienne, même s'il dit ne pas se fier exclusivement à celui-ci, pourrait fonder ses interventions *autrement* qu'à partir des biais relatifs au fonctionnement de ces outils et aux biais relatifs aux données qui les alimentent[190].

Même si les motifs à la base du soupçon raisonnable n'ont qu'à éveiller objectivement la *possibilité* de criminalité, les facteurs justifiant un soupçon raisonnable ne peuvent pas être des « facteurs anodins ». Un « facteur anodin » est un facteur qui n'indique pas nécessairement la possibilité que la personne soit engagée dans une activité criminelle précise. Il a été avancé qu'il ne résultera pas, de la simple combinaison de plusieurs facteurs anodins, « une sorte d'alchimie » capable de supporter un soupçon raisonnable de criminalité si ces facteurs ne se renforcent pas au point d'indiquer une *possibilité* que la personne visée soit engagée dans une activité criminelle[191]. En combinant une variété de facteurs généraux, extérieurs à la personne soupçonnée ou extérieurs au contexte particulier de l'enquête menée par les policiers, et qui, parfois, sont ouvertement des facteurs anodins (comme le veut la promesse marketing que les outils d'IA ont la capacité de découvrir des *hidden patterns* à partir de faits qui n'entretiennent aucun lien logique

---

[187] A. CHRISTIN, "predictive algorithms and criminal sentencing", préc., note 123, p. 280; Bernard E Harcourt, *Against Prediction: Profling, Policing, and Punishing in an Actuarial Age*. Chicago, IL: University of Chicago Press, 2006, p. 3.
[188] *R.* c. *Le*, 2019 CSC 34, par. 77-78.
[189] *R.* c. *Bernshaw*, [1995] 1 RCS 254, qui précise qu'une détention ou une arrestation fondée des données inexactes ne peut être qu'arbitraire.
[190] Citizen Lab, p. 125, et à la p. 127 : "Police services that seek to use algorithmic policing technologies that depend on underlying police data will thus face challenges in demonstrating the reliability and lack of bias in such systems as well as demonstrating that generalized suspicion does not play a role in subsequent decisions to detain or arrest someone."
[191] *R.* v. *Urban*, 2017 ABCA 436

apparent pour l'observateur humain), nous pourrions argumenter que les algorithmes fonctionnent bien souvent à l'instar de cette « alchimie » de faits anodins. Il s'avère alors que la *possibilité* de criminalité indiquée par les outils d'IA n'a de sens qu'à l'intérieur de sa méthode complexe de calcul et ne dit rien au policier d'expérience. Bien souvent, la considération par le policier d'expérience des facteurs multiples et décontextualisés pris en compte par l'algorithme ne serait pas à même d'éveiller chez lui un soupçon raisonnable. La nécessité de se référer au *bon sens* du policier et à son expérience pratique afin d'établir la raisonnabilité du soupçon est d'ailleurs soulignée par la Cour suprême :

> « L'examen de la question de savoir si un ensemble particulier de faits donne lieu à des soupçons raisonnables ne saurait se muer en un exercice scientifique ou métaphysique. Le bon sens, la flexibilité et l'expérience pratique quotidienne sont les mots d'ordre qui doivent guider cette analyse qui s'effectue du point de vue d'une personne raisonnable munie des connaissances, de la formation et de l'expérience de l'enquêteur. » (nos soulignés)[192]

Plus substantiellement encore, l'évaluation de la raisonnabilité du soupçon doit se rapporter à « la mesure dans laquelle il est nécessaire au policier de porter atteinte à une liberté individuelle afin d'accomplir son devoir »[193], donc en prenant en considération que « les droits relatifs à la liberté individuelle constituent un élément fondamental de l'ordre constitutionnel canadien. »[194] Le concepteur qui s'aventurerait à paramétrer son algorithme en fonction de cet enseignement ne pourrait faire autrement que d'imposer au policier son jugement de valeur sur la manière d'équilibrer ces valeurs. Au fond, il revient à se demander si nous acceptons collectivement que des innocents soient détenus pour fin d'enquête, au nom de la lutte effective contre la criminalité, en fonction de méthodes algorithmiques complexes qui, notamment en raison de l'apprentissage-machine, finissent sans doute par échapper au policier qui l'utilise[195]. En sommes-nous rendus, dans notre imaginaire collectif, à ce niveau de confiance et de délégation envers la machine au point où le raisonnement de la machine puisse produire les conditions raisonnables pour entraver la liberté individuelle ? N'avons-nous pas besoin en tant que citoyen, justement, pour pouvoir accepter une telle entrave à la liberté de pouvoir *partager,* avec le décideur à qui l'on délègue une partie du pouvoir de détention et d'arrestation, une certaine forme *commune* de raisonnement ? À notre avis, l'une des conditions premières à la délégation du pouvoir de privation de liberté envers les policiers est spécifiquement cette possibilité de *partager* et de *comprendre* le raisonnement de celui à qui l'on délègue ce pouvoir. C'est d'ailleurs ce qui se dégage derrière l'idée que les facteurs à la base du soupçon doivent pouvoir être vérifiables et « objectivement discernables », c'est-à-dire qu'ils doivent *pouvoir* faire l'objet d'un examen judiciaire indépendant, d'une procédure contradictoire, d'un dialogue[196]. En raison des problématiques de transparence de ces technologies et de leur fonctionnement par apprentissage-machine, qui peut même échapper au concepteur et, sans aucun doute, au policier, nous voyons difficilement comment le raisonnement de l'outil d'IA peut posséder ces qualités. L'outil d'IA, du fait de son opacité, *ferme ce dialogue* sur les conditions raisonnables à la privation de liberté et sur la manière d'équilibrer ces valeurs.

---

[192] *R.* c. *Mackenzie*, 2013 CSC 50, par. 73.
[193] *R.* c. *Mann*, 2004 CSC 52, par. 34.
[194] *R.* c. *Mann*, 2004 CSC 52, par. 35.
[195] Citizen Lab, p. 130 : "For example, in discussing the Vancouver Police Department's (VPD) GeoDASH algorithmic forecasting system, S/Constable Ryan Prox shared that VPD officers run their "algorithm in its machine-learning retraining mode at 3-week intervals; every 3-week interval it rewrites its algorithmic code. That's why we do the independent audits. Because I can't tell you what factors it's weighting according to making the determinations for the boxes. I can tell you if it's doing it accurately, based on where the incidents are taking place, but I can't tell you the 'why', and what weighting it's putting on what factors."
[196] *R.* c. *Chehil*, 2013 CSC 49, par. 26.

Pour être raisonnable, le soupçon doit être également fondé sur les caractéristiques *spécifiques* d'un suspect et non se fonder sur ses *caractéristiques générales*, comme ses caractéristiques immuables, sur les *caractéristiques du lieu* dans lequel il se trouve ou sur les *préjugés du policier* à l'égard d'un groupe culturel. En somme, un lien clair entre la personne *précise* faisant l'objet de la détention et une infraction criminelle récente ou en cours doit motiver l'intervention du policier[197]. Par nature, les algorithmes ne peuvent produire que des inférences, c'est-à-dire des corrélations statistiques fondées sur une compilation de caractéristiques générales : « Algorithmic policing methods tend to rely on generalized inferences by definition. »[198] Pour cette raison, la simple prédiction par un outil d'IA concernant la probabilité de criminalité dans un lieu donné ne peut fonder ou servir à fonder un soupçon raisonnable[199]. Comme l'a reconnu la Cour suprême dans *R*. c. *Mann*, « le fait qu'un quartier possède un taux de criminalité élevé ne constitue pas en soi une raison de détenir quelqu'un »[200]. Conséquemment, le déploiement des forces policières en fonction d'une prédiction de criminalité par calcul algorithmique est susceptible de vicier, *à la source*, la raisonnabilité du soupçon et de conditionner, à cause de la suggestion de l'algorithme et de son aura de scientificité, l'existence du soupçon avant même sa formation naturelle chez le policier. Sur ce point, l'aura d' « objectivité » qui accompagne la suggestion de l'outil d'IA est susceptible d'empiéter sur le degré de discrétion nécessaire pour que le policier d'expérience puisse prendre la bonne décision dans le contexte. Cette *subjectivité* et cette *discrétion* dans la prise de décisions en matière policière, que cherchent à combattre les outils d'IA, se révèlent finalement nécessaires au bon fonctionnement de notre système d'application de la loi :

> « The Supreme Court recognizes that police discretion is an essential feature of the criminal justice system. As Justice La Forest wrote in R v Beare, eliminating police "discretion would be unworkably complex and rigid." (…) Police discretion requires both rational justification that is proportionate to the seriousness of the conduct and exercising discretion in the public interest. (…) Whether and to what degree police officers should maintain their discretion when relying on predictive technologies involves a host of policy considerations. (…) While predictive technologies are theoretically capable of injecting a degree of objectivity into crime-prevention and policing, they may also serve to amplify and perpetuate existing practices that further marginalize over-policed groups. »[201]

L'arrestation sans mandat. Pour que l'arrestation sans mandat ne soit pas arbitraire, le policier doit avoir des « motifs raisonnables de croire » que la personne a commis ou est sur le point de commettre un acte criminel (art. 495(1)(a) C.Cr.). Ces « motifs raisonnables de croire » renvoient à une croyance subjective qui doit se fonder sur des faits objectivement justifiables permettant à une personne raisonnable de croire que la personne est impliquée dans un acte criminel[202]. Plusieurs enjeux liés au fonctionnement particulier des outils d'IA de prédiction vus dans la précédente sous-partie s'appliquent également ici.

Substantiellement, nous devons nous demander ce qui justifie cette délégation du pouvoir d'arrestation aux policiers. Qu'est-ce qui justifie une telle délégation, si ce n'est leur capacité de *partager* avec le citoyen ordinaire une certaine forme commune de raisonnement. C'est précisément la capacité pour le policier de comprendre que « ce besoin [pour la société d'être protégée contre le crime] commande l'établissement d'un *équilibre raisonnable* entre le droit des particuliers à la liberté et la nécessité de protéger la société contre

---

[197] *R*. c. *Mann*, 2004 CSC 52, par. 34-35.
[198] Citizen Lab, p. 125. *Cf*. également S. DU PERRON et K. BENYEKHLEF, préc., note 86, p. 19.
[199] Citizen Lab, p. 125.
[200] *R*. c. *Mann*, 2004 CSC 52, par. 47.
[201] Michael PURCELL et Mathew ZAIA, "Prediction, Prevention And Proof: Artificial Intelligence And Peace Bonds In Canada", 98-3 *Canadian Bar Review* 515, 2020, CanLIIDocs 3308, https://canlii.ca/t/t07k, p. 541.
[202] *R.* c. *Storrey*, [1990] 1 R.C.S. 241.

le crime. »[203] En ce sens, le critère d' « objectivité » exigée pour arrêter un suspect sans mandat, ne doit pas se confondre avec la « scientificité » ou la « dé-subjectivation » du raisonnement. Le critère objectif réfère même plutôt à une forme d'intersubjectivité humaine (qui prend la forme d'un dialogue entre le policier et la volonté collective des citoyens); une intersubjectivité humaine qui serait fondée sur des faits objectifs, devant être comprise comme un réel partagé par d'autres humains, donc qui est également empreinte de références *sociales* communes. Justement, lorsqu'on dit que « l'existence de ces motifs raisonnables et probables doit être *objectivement établie* », on dit, « en d'autres termes, [qu']il faut établir qu'une *personne raisonnable, se trouvant à la place de l'agent de police*, aurait *cru* à l'existence de motifs raisonnables et probables de procéder à l'arrestation. »[204] (nos italiques)

À cette « intersubjectivité raisonnable » requise par la loi s'oppose l'objectivité dépersonnalisée, inflexible, mathématique et décontextualisée des outils d'IA et dont le raisonnement ne fait pas nécessairement référence à un réel partagé, au sens social commun (« la personne raisonnable »). Le droit policier et pénal ne peut faire l'économie d'une référence à la personne humaine et ce, malgré les limitations propres à la subjectivité humaine; c'est que la mesure raisonnable dans la mise en équilibre de la liberté et de la sécurité est proprement sociale et dialogique. Elle résulte d'un « dialogue » entre le policier et la volonté collective des citoyens. Cet exercice d'équilibration des valeurs doit résulter d'une *tentative* par le policier de saisir cette volonté et qui pourra, si elle est contestée, être soumise à la procédure contradictoire du procès. Ainsi, aucune formule mathématique ne pourra répondre d'avance à cet exercice d'équilibration des valeurs de manière satisfaisante, celle-ci étant même sujette à évoluer avec le temps.

3.4. Autres protections constitutionnelles : Égalité procédurale, défense pleine et entière, droit au remède et réduction de la peine

Comme nous l'avons vu, l'exercice de plusieurs droits constitutionnels et le bénéfice des garanties en matière criminelle dépendent intrinsèquement des enjeux de la *transparence* et de la communication du code source des algorithmes à la base des outils d'IA. Chaque accusé a le droit à une défense pleine et entière, ce qui oblige la Couronne à communiquer l'entièreté de la preuve à l'accusé (art. 7 de la *Charte*)[205]. Afin de pouvoir faire contrôler, par habeas corpus, la légalité de sa détention (art. 10c) de la *Charte*), de contester son arrestation ou sa détention (art. 9 de la *Charte*) ou de bénéficier de son droit à la réparation (art. 24(1)(2) de la *Charte*), il semble nécessaire que l'accusé ait accès à un certain éclairage quant au fonctionnement de l'algorithme qui est à la source de son arrestation, de sa détention et de la poursuite intentée contre lui.

Une conduite discriminatoire ou préjudiciable par les agents de la paix et, par extension, le recours à un outil algorithmique qui favoriserait une conduite discriminatoire pourraient également mener à une réduction de la peine sur la base du concept de proportionnalité *individualisée* promulgué par la Cour suprême dans *R.* c. *Nasogaluak*[206]. Cette interprétation du principe fondamental de la peine appelle à prendre en considération, afin de déterminer le degré de sévérité de la peine, la *souffrance* qui a déjà été infligée à l'accusé alors qu'il était entre les « mains de l'État » (la conduite d'un procureur ou d'un policier) : « A

[203] *R.* c. *Storrey*, p. 249-250
[204] *R.* c. *Storrey*, p. 250
[205] R. c. *Stichcombe*, [1991] 3 RCS 326
[206] *R.* c. *Nasogaluak*, 2010 CSC 6.

Charter breach indicates that the state has offended these values and concerns and a sentence can and should communicate society's resulting condemnation if the breach has a sufficient link to the circumstances of the offence or the offender. (…) His sentence is justifiably reduced because he has already suffered harm at the hands of the state in response to his misconduct. When a judge decides how much and what form of punishment to inflict on the accused, the ways in which he has already suffered is salient. »[207]

[207] Pour le concept de « proportionnalité individualisée » de la peine qui permet de prendre en compte l'ensemble des souffrances déjà infligées par l'État à l'endroit de l'accusé, *cf*. Benjamin L. BERGER, « Sentencing and the Salience of Pain and Hope » 70 *Supreme Court Law Rev* 2d 337 qui fonde son interprétation, notamment, sur l'arrêt *R*. c. *Ipeelee*, 2012 CSC 13, par. 86 : « à qui le tribunal impose-t-il une peine si ce n'est au délinquant qui se trouve devant lui? Si le délinquant est un Autochtone, le tribunal doit tenir compte de sa situation dans son ensemble, y compris les circonstances particulières décrites dans l'arrêt *Gladue*. »

PARTIE II. JUSTICE PRÉDICTIVE

1. Pratiques nationales

Certains indicateurs démontrent que le Canada est intéressé à suivre l'expérience américaine, c'est-à-dire à algorithmiser les outils d'évaluation du risque de récidive et à recourir à leurs résultats au cours de la procédure en matière criminelle **(sous-partie 1.1)**., Nous présenterons le cadre légal particulier en matière correctionnelle susceptible d'offrir une garantie minimale d'exactitude au fonctionnement des outils d'IA **(sous-partie 2)**. Toujours en matière correctionnelle, nous verrons que la protection de l'article 15 de la *Charte* contre la discrimination offre une garantie d'impartialité lorsque *Service correctionnel Canada* (SCC) a recours à un outil actuariel pour déterminer la dangerosité d'un détenu **(sous-partie 3.1)**. Compte tenu qu'il n'existe pas actuellement en droit criminel canadien de recours aux outils d'IA aux fins de la justice prédictive, nous vous proposerons une analyse critique des principaux reconditionnements qui seraient susceptibles de survenir en rapport avec l'idée même du *Juste* si ces nouvelles technologies étaient intégrées dans le processus judiciaire **(sous-partie 3.2)**. Le cadre normatif propre à chacune des étapes où un outil d'IA de prédiction de la récidive pourrait être introduit, c'est-à-dire **(i)** la peine, **(ii)** l'évaluation de la cote de dangerosité, la libération conditionnelle, **(iii)** l'enquête sur remise en liberté et **(iv)** l'audience sur l'engagement de ne pas troubler l'ordre, sera présentée successivement de manière à relever les références, à chacune de ces étapes, à la notion du « juste » **(sous-partie 3.3)**. Cela permettra de relativiser l'importance à accorder aux prédictions du risque formulées par ces outils.

Qu'entendons-nous par la « justice prédictive » et comment cette notion se matérialise-t-elle au Canada ? La notion de « justice prédictive » est susceptible d'incorporer, tout d'abord, la pratique visant à recourir aux outils algorithmiques afin de prédire les probabilités qu'une certaine décision soit rendue; ces outils pourraient être utilisés par le procureur de la Couronne ainsi que par les avocats de la défense pour connaitre leur chance de succès. La pratique ne semble pas courante ou bien documentée au Canada. Ensuite, la « Justice prédictive » incorpore diverses pratiques liées à la prise de décision allant du recours aux outils d'IA afin d'*assister* un juge dans la décision à rendre jusqu'à l'automatisation *complète* de la prise de décision. La prise de décision automatisée ou l'assistance dans la prise de décision par des outils d'IA seraient déjà des pratiques courantes en matière administrative et ces pratiques sont notamment présentes à un certain degré en matière d'immigration et de demande d'asile[208]. Dans ces cas, la notion plus large de « justice prédictive » se laisse alors définir à travers la pratique particulière des « décisions automatisées ». Dans sa *Directive sur la prise de décisions automatisée* (2019) adressée à l'administration publique, le gouvernement canadien offre une première définition officielle de ce versant de la justice prédictive. On y utilise le terme « système décisionnel automatisé », qu'on définit ainsi : « comprend toute technologie qui soit informe ou remplace le jugement des décideurs humains. Ces systèmes proviennent de domaines tels que les statistiques, la linguistique et les sciences informatiques, et utilisent des techniques telles que les systèmes basés sur des règles, la régression, l'analytique prédictive, l'apprentissage automatique,

[208] Teresa SCASSA, préc., note 86, p. 5; Makoto Hong CHENG et Hui CHOON KUEN, "Towards a Digital Government: Reflections on Automated Decision-Making and the Principles of Administrative Justice", 31 *Sing Acad of Law Jo* 875, 2019. Plusieurs limites techniques empêchent les chercheurs de faire état de l'étendue du recours aux outils d'IA par les décideurs publics, notamment l'absence de recensement public, CDO1, p.9-10. Encore là, plusieurs limites techniques empêcheraient un tel recensement, par exemple, « it's difficult to get an accurate measure since case workers or probation officers don't always declare the use of risk reports", Agnese SMITH, « automating justice », *National - CBA National - Canadian Legal Affairs* (13 mars 2018), en ligne : <https://nationalmagazine.ca/en-ca/articles/law/ethics/2018/automating-justice> (consulté le 15 mars 2022).

l'apprentissage en profondeur et les réseaux neuronaux. » (nos soulignés) Rappelons que la *Directive* n'est pas obligatoire, qu'elle a été écrite pour s'appliquer aux organisations *administratives fédérales* offrant un *service* au public, qu'elle n'a pas été pensée pour encadrer des procédures en matière criminelle, qu'elle ne s'applique pas aux services nationaux de sécurité (clause 5.4) ni aux services de détection de la fraude[209]. Il n'est pas clair non plus si la *Directive* devrait être suivie dans le cadre des décisions du SCC. À notre avis, cela ne cadrerait pas avec l'esprit sous-tendant la *Directive*; nous voyons difficilement comment le traitement offert dans un établissement carcéral pourrait être associé, par analogie, à un simple « service » externe rendue par une entité administrative à des « clients » (clause 5.2). Dans le contexte carcéral, des garanties supplémentaires devront être implantées.

En droit criminel et dans le système correctionnel, le recours aux outils algorithmiques s'inscrirait, comme dans les services policiers, à l'intérieur d'une réorientation plus profonde de la fonction de l'institution pénale vers la *prévention* du crime. L'institution pénale en général semble elle aussi se déplacer d'une fonction *réactive* - maintenir un ordre public en exprimant, après le fait, des valeurs par l'entremise d'une punition - vers une fonction *préventive* axée sur la protection et la sécurisation effective de la société et de ses membres vulnérables. Ce contexte aura permis de favoriser la prolifération des outils actuariels et statistiques au Canada portant sur le risque de récidive à travers toutes les étapes de la procédure en droit criminel[210] pour finalement permettre de *normaliser* cette fonction de prévention face aux risques. L'ajout éventuel d'une IA à ces outils actuariels du risque représenterait simplement une phase d'*intensification* de cette réorientation plus large de l'institution pénale vers la fonction préventive du droit criminel. Sur ce point, certains chercheurs remarquent d'ailleurs la parenté des arguments justifiant l'introduction des outils actuariels sur le risque au niveau de la peine avec ceux qui soutiennent le recours aux outils d'IA pour informer une décision[211]. En résumé, la quête d'objectivité dans la prise de décision, le besoin de cohérence et d'uniformité, notamment au niveau de la peine, ainsi que des attentes de plus en plus élevées au sein du public en matière de sécurisation effective de la société pourraient être à l'origine de l'intérêt du Canada envers les outils d'IA aux fins de justice prédictive.

---

[209] T. Scassa, préc., note 86, p. 12-13.

[210] Par exemple, *The Ontario Domestic Assault Risk Assessment* (ODARA) est un outil actuariel canadien de prédiction du risque créé par l'OPP pour prédire le risque de récidive de violence conjugale, Site web d'ODARA, en ligne : https://odara.waypointcentre.ca/. Il avait été utilisé dans la phase pilote pour déterminer l'utilité de la caution, Site Web du gouvernement de l'Ontario, « McGuinty government launches new tool to assess risk of domestic assault », 12 novembre 2004, en ligne : https://news.ontario.ca/en/release/93058/mcguinty-government-launches-new-tool-to-assess-risk-of-domestic-assault ■ *Static-99R* est un outil de prédiction de la récidive pour les délinquants sexuels qui est parfois utilisé lors de la peine, cf. PUBLIC SAFETY CANADA, « Static-99R Coding Rules », 21 décembre 2018, en ligne : https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/sttc-2016/index-en.aspx#h2.5-h3.2. ■ *Security Reclassification Scale* (SRS) est un outil actuariel non-algorithmique utilisé par les services correctionnels pour déterminer la cote de sécurité à attribuer à l'accusé, *May* c. *Établissement Ferndale*, 2005 CSC 82 ■ Dans *Ewert* c. *Canada*, 2018 CSC 30, par. 11 nous apprenions que le SCC aurait utilisé les outils actuariels du risque suivants : « *l'échelle de psychopathie de Hare - révisée* (« PCL-R »), un outil qui a été conçu pour évaluer la présence de psychopathie, mais qui est aussi utilisé pour mesurer le risque de récidive. M. Ewert a également contesté le recours au *Guide d'évaluation du risque de violence* (« GERV ») et au *Guide d'évaluation du risque chez les délinquants sexuels* (« GERDS »), deux outils actuariels conçus pour évaluer le risque de récidive violente; la *Statique-99*, un outil actuariel conçu pour estimer la probabilité de récidive sexuelle ou violente; *l'Échelle des risques de violence : Délinquants sexuels* (« ERVDS ») une échelle visant à évaluer le risque de récidive sexuelle qui est employée relativement à la prestation des traitements destinés aux délinquants sexuels. ». Pour une critique de l'impact du recours de plus en plus fréquent de ces outils d'évaluation du risque à l'étape de la peine, Kelly HANNAH-MOFFAT, "Actuarial Sentencing: An "Unsettled" Proposition", (2012) 1-27 *Justice Quarterly* 30.

[211] Danielle KEHL, Priscilla GUO et Samuel KESSLER, "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing", *Responsive Communities Initiative*, Berkman Klein Center for Internet & Society, Harvard Law School, 2017, p. 5.

1.1. Intérêt du Canada dans le recours aux outils algorithmiques pour prédire le risque de récidive dans le cadre d'une procédure en matière criminelle

Il semblerait qu'il n'y ait pas actuellement au Canada un usage notable des outils d'IA pour informer ou assister une décision en matière criminelle ou correctionnelle[212]. En revanche, il y aurait déjà aux États-Unis un recours à des outils algorithmiques et des logiciels pour évaluer le risque de récidive – dont certains fonctionnent par *apprentissage-machine* – lors de la mise en liberté sous caution, de la détermination de la peine et de la libération conditionnelle[213]. Nous pensons qu'en raison de la normalisation progressive, au sein de la procédure criminelle canadienne, des pratiques axées sur la prédiction du risque à l'aide d'outils actuariels, ces outils pourraient éventuellement être déployés et utilisés au Canada. Plusieurs acteurs privés et publics au Canada ont exprimé leur intérêt dans le développement de ces outils.

Uniquement au cours des quatre dernières années, il est possible d'identifier plusieurs centres d'entrepreneuriat, des centres de développement de logiciels et plusieurs chercheurs universitaires qui ont annoncé avoir envisagé de développer ou avoir déjà développé des outils fonctionnant avec un système d'IA destinés à assister un décideur dans sa prise de décision en droit criminel[214]. En 2018, des chercheurs étudiaient, à des fins académiques, la possibilité d'introduire en **Colombie-Britannique** un système fonctionnant par *apprentissage-machine* pour assister la prise de décision relative à la mise en liberté sous caution[215]. Les chercheurs derrière ce projet (Albert Yoon, Benjamin Alarie & Anthony Niblett - *Blue J Legal* et *Université de Toronto*) sont connus comme étant les créateurs d'un logiciel utilisant l'IA afin d'identifier des « hidden patterns » dans les décisions judiciaires en droit du travail[216]. En **Ontario**, un avocat et un ingénieur, en partenariat avec la *Legal Innovation Zone* de l'*Université Ryerson de Toronto* et l'accélérateur d'entreprise privée *The Accelerator Centre*, ont mis sur pied un logiciel (*SmartBail*) qui fonctionne par *apprentissage-machine* et qui est destiné à assister les juges dans leur prise de décision lors de l'enquête sur remise en liberté : « SmartBail accurately and objectively assesses pre-trial risk, giving decision-makers the insight to process individuals awaiting bail quickly, fairly, and at a significantly lower cost to taxpayers. »[217] En **Saskatchewan**, des chercheurs du *Centre for Forensic Behavioural Science and Justice Studies* de l'*Université de Saskatchewan* travaillent depuis 2020 sur le développement d'un outil d'IA susceptible de remplacer le *Level Service Inventory - Ontario Revised* (« LSI-OR »). Le LSI-OR est un outil d'évaluation des risques et des besoins des détenus utilisé dans notre système correctionnel[218]. Selon

---

[212] Citizen Lab, p. 54

[213] L'un des outils les plus notables aux États-Unis est le « Correctional Offender Management Profiling for Alternative Sanctions » (ci-après « COMPAS ») qui est utilisé pour assister le décideur lors de l'enquête sur la mise en liberté sous caution, la peine et la libération conditionnelle, *cf.* Citizen Lab, p. 53. *V.* également pour d'autres détails, Danielle KEHL, Priscilla GUO et Samuel KESSLER, préc., note 212, p. 9-12

[214] Citizen Lab, p. 53.

[215] Agnese SMITH, préc., note 209.

[216] Doug O'NEILL, « Toronto Startup Blue J Legal Uses AI to Help Predict How Courts will Rule on Employment Law Cases », *Startup here Toronto*, 5 juin 2018, en ligne : https://startupheretoronto.com/type/profiles/toronto-startup-blue-j-legal-uses-artificial-intelligence-to-help-predict-how-courts-will-rule-on-employment-law-cases/ ; Chris SORENSEN, « U of T startup Blue J Legal raises US$7 million, plans cross-border expansion », 29 Novembre 2018, *U of T News*, en ligne : https://www.utoronto.ca/news/u-t-startup-blue-j-legal-raises-us7-million-plans-cross-border-expansion

[217] Autrefois, le site Smartbail.ca nous menait vers une page Web décrivant le logiciel, mais désormais il s'agit d'un site vendant des chaussures! Nous vous référons plutôt au compte twitter de Smart Bail, https://twitter.com/smart_bail; Alyshah HASHAM, "Soon, intelligent machines could help decide whether to keep people in jail. It's time to prepare", *Toronto Star*, 19 Juillet 2019, en ligne : Soon, intelligent machines could help decide whether to keep people in jail. It's time to prepare | The Star

[218] Concernant le LSI-OR, voir CORRECTIONAL SERVICE CANADA**,** FORUM on Corrections Research, Volume 9, Numéro 1, en ligne : https://www.csc-scc.gc.ca/research/forum/e091/e091f-eng.shtml (page archivée, dernière modification 2015-03-05)

le site web de l'*Université*, ce nouvel outil permettrait de réévaluer, par un système automatisé fonctionnant par *apprentissage-machine*, la cote de sécurité des détenus et offrirait une évaluation « unique » de leurs risques et de leurs besoins : « The project mainly includes two parts as follows: 1) studying unique patterns, or in other words, the way offenders complete the LSI-OR; and 2) employing a machine learning approach, such as Naive Bayes classifier, as an alternative to the LSI-OR. »[219] S'il devait être mis en service, ce nouvel outil constituerait alors le premier outil algorithmique utilisé au Canada en matière correctionnelle servant à assister des autorités dans des décisions pouvant avoir un impact direct sur la liberté d'une personne[220].

Même si l'enthousiasme des décideurs publics face à ces innovations reste timide, certaines agences gouvernementales semblent avoir montré un certain intérêt envers le recours aux outils d'IA aux fins de justice prédictive[221]. En **Ontario**, un porte-parole du *Ministère du procureur général* avait affirmé que la province évaluait bel et bien la possibilité d'implanter un outil algorithmique pour soutenir des décisions reliées à la mise en liberté sous caution. En 2018, ces recherches étaient à la phase préliminaire[222]. En 2019, sans confirmer que le gouvernement de l'Ontario considérait encore le recours à l'IA à l'étape de la mise en liberté sous caution, un porte-parole du même *Ministère* affirmait que le gouvernement "remains committed to continuously improving and to investigating modern methods of case evaluation, such as risk assessment tools, while ensuring public safety and the rights of the accused."[223]

1.2. Inquiétudes exprimées quant à la fiabilité des outils d'IA à la lumière de l'expérience américaine

Ces outils d'IA n'ayant pas encore été déployés au Canada, il n'existerait pas d'étude visant à établir la fiabilité et l'impartialité d'un outil d'IA en particulier dans le contexte canadien, notamment en ce qui a trait à son impact sur les populations surreprésentées dans nos milieux carcéraux (autochtones et personnes atteintes de troubles mentaux)[224]. Des études aux États-Unis ont été menées pour examiner les potentiels biais et la fiabilité des outils actuariels, non-algorithmiques, utilisés pour prédire la récidive et pour assister les décideurs, comme le LSI-R (version antérieure du LSI-OR) qui était utilisé dans les milieux correctionnels au Canada et qui peut également être utilisé pour la peine, la libération conditionnelle et lors de l'enquête sur remise en liberté. Selon certaines études, le LSI-R aurait tendance à surévaluer le risque de récidive des personnes noires et à les surclasser dans le cadre de l'exercice d'attribution de la cote de sécurité; les facteurs de risque utilisés par l'outil actuariel pour évaluer le risque de récidive de la population

---

[219] UNIVERSITÉ DE SASKATCHEWAN, Centre for Forensic Behavioural Science and Justice Studies, "Current Projects", en ligne : Current Projects - The Centre for Forensic Behavioural Science and Justice Studies - University of Saskatchewan (usask.ca); Rahim ORAJI, « A Machine Learning Generalization of LSI-OR », A Thesis Submitted to the College of Graduate Studies and Research in Partial Fulfillment of the Requirements for the degree of Master of Science in the Department of Computer Science University of Saskatchewan Saskatoon, 2016, disponible en ligne : A Machine Learning Generalization of LSI-OR (usask.ca)

[220] Citizen Lab, p. 54. Concernant le fonctionnement d'un outil d'IA par inférence Bayésienne, voir Tomas ROJAS VAZQUEZ, « L'inférence bayésienne : l'intelligence artificielle par la statistique », Site web du Laboratoire de Cyberjustice, 2020, en ligne : https://www.cyberjustice.ca/2020/12/16/les-techniques-algorithmiques-de-lia-linference-bayesienne/ . L'inférence bayésienne pourrait se définir comme étant la capacité d'« apprentissage dans l'incertitude », pour plus de détails voir Sadie CREESE, « The threat from AI », p. 206, dans *Artificial Intelligence And The Law : Cybercrime And Criminal Liability*, Dennis J Baker et Paul H. Robinson (dir.), New York, Routledge, 2021.

[221] Citizen Lab, p. 53.

[222] Agnese SMITH, préc., note 298.

[223] Alyshah HASHAM, "Soon, intelligent machines could help decide whether to keep people in jail. It's time to prepare", préc., note 218.

[224] CDO1, p. 22.

générale pourraient ne pas être les mêmes pour des populations particulières, culturellement minoritaires[225]. Au Canada, des études semblables ont été menées à partir de notre population carcérale particulière afin de vérifier la prédictivité des modèles d'outils actuariels traditionnels, comme ceux de la famille LSI, sur la population spécifique des détenus autochtones. Selon une étude, il semblerait que ces outils offrent une grande fiabilité, même lorsqu'ils sont utilisés auprès des détenus autochtones[226]. Néanmoins, certaines disparités de traitement dans les résultats lorsque ces tests étaient effectués auprès des détenus autochtones ont justifié l'expression d'une mise en garde et d'un appel à la prudence lorsque l'on recourait à ces outils auprès de cette population particulière[227]. En règle générale, il semblerait que les outils actuariels non-algorithmiques soient efficaces pour prédire le risque général au sein d'un groupe, mais qu'ils le sont moins efficaces lorsqu'ils sont utilisés auprès de sous-communautés spécifiques ou auprès des individus[228]. Selon une étude réalisée au Canada, le recours fréquent par les outils actuariels traditionnels non-algorithmiques à certains *facteurs statiques*, comme les antécédents judiciaires, (par opposition aux *facteurs dynamiques* qui, eux, prennent en considération la progression spécifique de l'accusé dans son plan de réhabilitation) tendent à exagérer les facteurs de risque des détenus autochtones, comparativement aux détenus non-autochtones, et pourraient avoir moins d'efficacité prédictive lorsqu'ils sont utilisés auprès de cette population[229]. Aux États-Unis, le COMPAS de la compagnie *Equivante*/*Northpointe* a également fait l'objet d'études qui ont démontré de potentiels biais raciaux. Selon une étude, cet outil actuariel *algorithmique* générerait des prédictions disproportionnées quant au risque de récidive des détenus afro-américains[230].

À la lumière de l'étude sur l'outil COMPAS, des inquiétudes ont été émises par les rapporteurs du *Citizen Lab* notamment en ce qui a trait à la possible reconduction et prolifération des biais discriminatoires des concepteurs des outils d'IA de prédiction de risque de récidive. Ceux-ci étant appelés à faire des choix dans la programmation, dans la manière d'ordonner les différents paramètres, dans la sélection des données et des facteurs qui sont pris en compte, ils pourraient, consciemment ou non, construire une architecture technique qui, systématiquement, offrirait des résultats inégalitaires en fonction d'un motif prohibé de discrimination ou un motif analogue[231]. Les rapporteurs de la CDO ont également fait part de leurs inquiétudes particulières quant à la qualité et la fiabilité des données provenant de banques de données sur les détenus, comme la *Ontario's ICON* (*Integrated Courts Offenses Network*) et la CIPC, qui pourraient

---

[225] Daniel KONIKOFF et Akwasi OWUSU-BEMPHAH, "Big Data and Criminal Justice – What Canadians Should Know", Broadbent Institute, p. 6, en ligne : https://d3n8a8pro7vhmx.cloudfront.net/broadbent/pages/7751/attachments/original/1592504286/BigData_and_Criminal_Justice_Report.pdf?1592504286; Tracy l. FASS et al., "The LSI-R and the COMPAS validation data on two risk-needs tools", (2008) 35(9) *Criminal Justice and Behavior* 1095; Kevin W. WHITEACRE, "Testing the Level of Service Inventory–Revised (LSI-R) for Racial/Ethnic Bias", (2006) 17(3) *Criminal Justice Policy Review* 330, p. 341 conclura que le LSI-R avait une tendance consistante à commettre des erreurs de classification chez les personnes noires comparativement à lorsqu'il était utilisé chez les personnes caucasiennes ou hispaniques.

[226] J. Stephen WORMITH, Sarah M. HOGG et Lina GUZZO, "The Predictive Validity of the LS/CMI with Aboriginal Offenders in Canada", (2014) 42(5) *Criminal Justice and Behavior* 481, p. 504.

[227] *Id*.

[228] Kate ROBERTSON et Jill R. PRESSER, "Algorithmic Technology and Criminal Law in Canada", Chapitre 3, p.74, dans Jesse Beatson, Gerald Chan, Jill R. Presser (dir.), « Litigating Artificial Intelligence », *Emond Publishing,* Toronto, 2020 citant M. Shaw et K. Hannah-Moffat, "Gender, Diversity And Risk Assessment In Canadian Corrections", 47(3) *Probation Journal* 163, 2000.

[229] K. ROBERTSON et J. R. PRESSER, "Algorithmic Technology and Criminal Law in Canada", préc., note 229, p. 117 citant Perley-Robertson, Bronwen & Helmus, Leslie Maaike & Forth, Adelle, « Predictive accuracy of static risk factors for Canadian Indigenous offenders compared to non-Indigenous offenders: Implications for risk assessment scales », (2018) 25 *Psychology, Crime & Law*.

[230] Julia ANGWIN, Jeff LARSON, Surya MATTU et Lauren KIRCHNER, « Machine Bias », *ProPublica*, en ligne : Machine Bias — ProPublica; Tracy l. fass, préc., note 226; Julia DRESSEL et Hany FARID, "The accuracy, fairness, and limits of predicting recidivism", 4(1) *Science Advances*, 2018; Citizen Lab, p. 54, et notamment la note de bas de page 214 concernant les réponses face à ces critiques de la part du développeur et de certains criminologues.

[231] Citizen Lab, p. 120.

potentiellement être utilisées par un futur outil d'IA. Considérant l'impact direct sur la liberté du détenu et le traitement hautement performatif, extensif et invasif des technologies d'IA, le recours à ces banques de données exige des standards plus élevés en matière de qualité des données, de fiabilité et d'impartialité dans leur traitement. Les chercheurs de la CDO proposent de se référer aux standards établis dans le rapport de *Partnership on AI :* « There are many articles, reports and recommendations regarding best practices for data in pretrial risk assessments and similar tools, including several Partnership on AI proposals, addressing requirements on how measure for intended variables, bias, distinct predictions and data retention and reproducibility. It is not clear whether Canadian criminal databases, such as Ontario's ICON or CPIC [en français, "CIPC"], meet these standards. »[232]

2. Cadre légal particulier en matière correctionnelle et garanties d'impartialité, de fiabilité, d'efficacité et d'exactitude

Il n'existe pas de cadre légal visant l'encadrement des outils d'IA utilisés aux fins particulières de la justice prédictive. Nous vous renvoyons donc à l'exposé du cadre normatif général de la Partie I. Nous complèterons toutefois ce cadre normatif général qui traitait de certaines garanties en matière d'impartialité, de fiabilité, d'efficacité et d'exactitude, en présentant le cadre légal particulier qui s'applique en matière correctionnelle puisque celui-ci fait également référence à certaines de ces garanties.

En matière correctionnelle, la *Loi sur le système correctionnel et la mise en liberté sous condition* (LSCMLC) prévoit un cadre normatif particulier concernant la collecte, le partage et la protection des renseignements personnels des détenus dans les pénitenciers fédéraux. Le SCC doit collecter tous « les renseignements personnels pertinents, notamment les antécédents sociaux, économiques et criminels, y compris comme jeune contrevenant » (art. 23 LSCMLC). Ces renseignements pourront être ensuite communiqués « pour prendre la décision de les mettre en liberté » ou « pour leur surveillance » (étape de la libération conditionnelle). Ils pourront être communiqués à la « Commission des libérations conditionnelles du Canada, aux gouvernements provinciaux, aux commissions provinciales de libération conditionnelle, à la police et à tout organisme agréé par le Service en matière de surveillance de délinquants » et même, dans certaines situations, à la victime (art. 25-26 LSCMLC).

Ces renseignements pourraient éventuellement être utilisés et traités par un futur outil d'IA pour déterminer la cote de sécurité du détenu, pour accorder ou refuser la libération conditionnelle d'un détenu ou pour permettre à le SCC d'évaluer l'opportunité de soumettre un cas à la *Commission des libérations conditionnelles du Canada* (« CLC »), afin de le maintenir en incarcération et de l'empêcher de bénéficier de la libération d'office. On sait qu'actuellement, afin de prendre leurs décisions, « les commissaires se basent également sur des évaluations actuarielles et des instruments d'évaluation du risque » et se fondent sur « tous les renseignements pertinents » pour évaluer le risque de récidive dont « les renseignements fournis par la police, les tribunaux, les procureurs de la Couronne, les professionnels de la santé mentale, les autorités correctionnelles, les organismes privés et les victimes d'actes criminels »[233]. Afin d'assurer la

[232] CDO1, 25. Concernant les recommandations de PARTNERSHIP ON AI, "Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System", 23 avril 2019, en ligne : Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System - The Partnership on AI

[233] COMMISSION DES LIBÉRATIONS CONDITIONNELLES DU CANADA, « Commission des libérations conditionnelles du Canada : Pour la sécurité du public », Janvier 2011, en ligne : https://www.canada.ca/fr/commission-liberations-

fiabilité d'une décision fondée sur ces données, la *Loi sur le système correctionnel et mise en liberté sous condition* prévoit que le SCC « est tenu de veiller, dans la mesure du possible, à ce que les renseignements qu'il utilise concernant les délinquants soient à jour, exacts et complets. » (art. 24(1)) Compte tenu des potentiels biais discriminatoires et de l'opacité du traitement algorithmique, il semble important de renforcer cette garantie lorsqu'un outil d'IA est utilisé à des fins de prédiction de la récidive. Au soutien de notre préoccupation, nous citons la Cour suprême qui, dans *Ewert* c. *Canada* (2018), a jugé que l'obligation de vérification raisonnable de l'exactitude des données devait s'étendre, non seulement aux renseignements collectés, mais également *aux résultats produits* par un outil actuariel d'évaluation du risque et que, par conséquent, l'outil utilisé devait avoir une « capacité forte plutôt que faible à prédire les risques »[234]. De cette garantie d'efficacité découle une certaine protection contre *l'impartialité* et les *biais discriminatoires* des algorithmes, « le SCC devait à tout le moins prendre au sérieux les préoccupations crédibles, maintes fois soulevées, sur la validité douteuse des renseignements obtenus à partir des outils contestés concernant les détenus autochtones parce que ces outils ne tiennent pas compte des différences culturelles. »[235] La Cour suprême dans cette affaire critiquait l'approche adoptée par le SCC : « [e]n faisant fi de la possibilité que ces outils désavantagent systématiquement les délinquants autochtones et en omettant de prendre des mesures pour s'assurer qu'ils génèrent des renseignements exacts, le SCC a manqué à l'obligation qui lui incombe suivant le par. 24(1) de la LSCMLC. »[236] Cette décision pourrait avoir un impact important sur l'utilisation future d'outils d'IA pour prédire le risque par nos services correctionnels, car elle exige désormais un degré supplémentaire de précaution avant de les utiliser, voire de se retenir de les utiliser lorsque le détenu est issu d'un groupe culturel minoritaire[237].

3. Principes généraux du droit

3.1. Exercice du droit à l'égalité indissociable de la garantie de transparence des outils d'IA

L'article 15 de la *Charte canadienne* est susceptible d'offrir une protection aux personnes sujettes à une évaluation du risque lorsque l'utilisation d'un outil actuariel entraine un traitement carcéral disproportionnément sévère ou fondé sur des biais à l'égard d'un groupe vulnérable protégé par la *Charte*[238]. Comme nous l'avons vu, la Cour suprême révèle dans *Ewert* c. *Canada* que le recours par la SCC à des outils actuariels qui « surestiment effectivement le risque posé par les détenus » qui sont issus d'un des groupes protégés par l'article 15 ou qui « mènent à des conditions d'incarcération plus sévères ou à la privation de possibilités de réadaptation en raison d'une telle surévaluation » violerait la protection constitutionnelle contre la discrimination prévue à la *Charte*[239]. Selon des chercheurs, une preuve similaire à celle qui découlait de l'enquête menée par *ProPublica* à l'endroit de l'outil COMPAS aux États-Unis permettrait alors de soutenir la preuve d'un traitement discriminatoire par l'État[240]. Cependant, pour en arriver à une telle conclusion, il est nécessaire de prouver devant la cour que l'on a subi un traitement discriminatoire. En effet le fardeau de la preuve repose sur celui qui prétend être victime d'une telle

conditionnelles/organisation/publications-et-formulaires/commission-des-liberations-conditionnelles-du-canada-pour-la-securite-du-public.html
[234] *Ewert* c. *Canada*, 2018 CSC 30, par. 35 et 43.
[235] *Id.*, par. 66.
[236] *Id.*
[237] CDO1, p.21.
[238] *Id*., p. 22.
[239] *Ewert* c. *Canada*, 2018 CSC 30, par. 79.
[240] Citizen Lab, p.120-121; S. DU PERRON et K. BENYEKHLEF, préc., note 86, p. 52.

discrimination. Dans l'arrêt *Ewert*, le détenu ne détenait d'ailleurs pas assez de preuve afin de soutenir sa prétention. Selon des chercheurs, « la charge de prouver comment le fonctionnement d'outils d'évaluation du risque peut mener à des résultats discriminatoires constitue un obstacle majeur à la protection effective du droit à l'égalité à l'égard de tous »[241]. Cela est d'autant plus vrai lorsque ces outils sont associés à une IA en raison de leur fonctionnement opaque et des limites à l'accès au code source.

Par conséquent, la possibilité de bénéficier de la protection contre la discrimination demeure conditionnelle au respect de la transparence de l'outil d'IA et de l'accès au code source. L'exercice du droit à l'égalité est donc également menacé par les protections accordées au secret commercial. Pour cette raison, des avocats demandent aux décideurs publics d'éviter d'avoir recours à des outils d'IA de compagnies privées, qui sont protégés par le secret commercial, afin de prévenir de telles restrictions à la communication de la preuve[242]. Comme nous le verrons, les résultats issus d'un outil de prédiction du risque apparaissent issus d'un traitement purement *technique* et *objectif,* alors qu'en réalité, ceux-ci sont le résultat de choix normatifs et politiques effectués *en amont* par les concepteurs de l'outil. Ainsi, la manière d'ordonner certains paramètres de l'algorithme est susceptible de favoriser une décision plutôt qu'une autre. En ce sens, les paramètres mêmes devraient pouvoir faire l'objet d'un débat contradictoire – certains auteurs parlent d'un droit au « technological due process »[243]. Une plus grande transparence est nécessaire afin que les tribunaux et les chercheurs externes puissent évaluer si les facteurs ou le poids accordé à certains facteurs produisent un traitement inégalitaire[244]. En raison de la normativité qui découle de l'architecture technique, la CDO a également exprimé la nécessité d'améliorer, dans le système judiciaire canadien, les garanties actuelles en matière de transparence en garantissant à tous les intervenants judiciaires un *accès à un exposé intelligible* sur les choix effectués *en amont* de la prédiction statistique, notamment quant aux données utilisées et quant aux paramètres priorisés par le concepteur (« *data literacy* »)[245].

3.2. Atteintes à la notion du « Juste » en droit criminel et aux autres principes généraux du droit

Les outils d'IA utilisés pour assister ou rendre une décision sont susceptibles d'influencer considérablement notre conception du « juste » en droit criminel. Ils proposent et normalisent une certaine conception de la décision *juste* à partir d'une méthode délibérative qui diffère substantiellement de l'exercice décisionnel traditionnel mené par un humain dans sa quête de la décision *juste*. Nous proposons une réflexion sur l'épistémologie propre à la quête de justice afin de voir en quoi les outils d'IA détournent celle-ci au profit d'une réponse *prédictible*, guidée par les paramètres de l'utilité et de l'efficacité. La décision de l'outil d'IA est valide uniquement du fait de sa conformité à une réalité statistique et quantifiable, au détriment de l'imaginaire collectif, social et symbolique qui soutient le droit. Nous avançons que la décision *juste* peut difficilement être prédéterminée et produite positivement par un outil d'IA, qu'elle découle plutôt d'une expérience humaine – une quête visant à modérer et limiter sa propre faillibilité – et du rituel du procès assuré par une procédure contradictoire et équitable. Après cette réflexion générale, nous présenterons les

[241] S. DU PERRON et K. BENYEKHLEF, préc., note 86, p. 52
[242] CDO1, p. 32 se référant à Taylor R. MOORE, "Trade Secrets and Algorithms as Barriers to Social Justice", *Center for Democracy and Technology*, 3 août 2017, en ligne : Trade Secrets and Algorithms as Barriers to Social Justice - Center for Democracy and Technology (cdt.org)
[243] Danielle KEHL, Priscilla GUO et Samuel KESSLER, préc., note 212, p. 32 se référant à Danielle KEATS CITRON et Frank PASQUALE, "The Scored Society: Due Process for Automated Predictions", 2014, 90 *Wash. U. L. Rev.* 1, p. 20 et Danielle KEATS CITRON, "Technological Due Process", (2008) 85 *Wash. U. L. Rev.* 1249, p.1254.
[244] *Id*; CDO1, p. 32-33.
[245] CDO1, p. 26.

différentes références à la notion du « juste » en droit pénal canadien. Cette présentation nous permettra de relativiser l'importance à accorder aux prédictions des outils d'IA lorsqu'ils sont utilisés dans une procédure en matière criminelle ou correctionnelle.

Y'a-t-il une idée du « juste » proposée par les outils d'IA? Si plusieurs auteurs craignent l'introduction de ces outils dans le droit, c'est que ceux-ci seraient à même d'injecter leur propre conception du *juste* à l'intérieur de notre système de droit. Pour Lawrence Lessig, les artéfacts technologiques ne sont pas axiologiquement neutres puisque leur architecture technique est à même de proposer une nouvelle conception du *juste* (« Code is Law »)[246] : « Just like any other technological artifact, code is not neutral, but inherently political: it has important societal implications, insofar as it might support certain political structures or facilitate certain actions and behaviors over others »[247]. La création d'un algorithme requiert inévitablement une réflexion humaine afin de faire un choix dans la manière d'ordonner les paramètres, cela impliquera nécessairement un jugement moral et normatif. Par exemple, « [t]hose who develop an algorithm such as COMPAS may have to choose (or may unknowingly decide) between, for example, prioritizing equal predictive accuracy for all groups, prioritizing minimizing false negatives over minimizing false positives, or prioritizing the maximization of true positives at the expense of increasing false positives. »[248] Le fait de recourir à un outil d'IA pour assister une décision ne devrait pas être un moyen de contourner l'exigence d'une procédure contradictoire équitable; l'outil devrait pouvoir faire l'objet d'un débat sur les choix politiques et moraux qui résident derrière la conception de l'outil. Autrement, il existe un risque réel de voir la réflexion des concepteurs de l'algorithme sur ce qui est « juste » s'imposer à celle du juge.

Créer un algorithme à partir d'une certaine conception du *juste*, qu'elle soit consciente ou non, puis la fixer indéfiniment par le codage a comme inconvénient de clore le débat, d'évacuer et de rigidifier la quête de justice et de solutions : « in contrast to traditional legal rules, which must be appreciated by a judge and applied on a case-by-cases basis, code-based rules are written in the rigid and formalized language of code, which does not benefit from the flexibility and ambiguity of natural language. »[249] Afin d'illustrer la problématique de la polysémie de la notion de « justice », le scientifique Arvind Narayanan s'est exercé à recenser au moins 21 formules *différentes* de ce qui est *juste* pouvant être utilisé par un outil d'IA[250]. Cette polysémie, qui est menacée par le recours aux outils d'IA, nous semble pourtant essentielle à l'exercice même de la quête de la décision *juste*, en ce sens qu'elle permet l'adaptabilité, qu'elle encourage le débat contradictoire entre des conceptions différentes du *juste* et qu'elle empêche de se rabattre sur des réponses préétablies, sans lien avec l'expérience vécue des sujets lors du procès.

Corrélation statistique vs. causalité juridique. Les outils d'IA, une fois introduits dans le procès en droit criminel, viennent proposer une nouvelle conception de la décision *juste* à rendre; une conception du *juste* qui met davantage l'emphase sur la prévisibilité, la continuité d'un certain état de fait, sur l'efficacité de la décision et dont la validité juridique se limite à son adéquation à une réalité *statistique*, qui plus est, à une

---

[246] L. LESSIG, *Code and Other Laws of Cyberspace*, New York, Basic Books, 1999, p. 59; K. BENYEKHLEF, *Une possible histoire de la norme : les normativités émergentes de la mondialisation* 2e éd., Éditions Thémis, Montréal, p. 98; Samer HASSAN et Primavera DE FILIPPI, « The Expansion of Algorithmic Governance: From Code is Law to Law is Code », 2017, *Field Actions Science Reports*, Special Issue 17.
[247] Samer HASSAN et Primavera DE FILIPPI, préc., note 247.
[248] Citizen Lab, p. 120.
[249] Samer HASSAN et Primavera DE FILIPPI, préc., note 247.
[250] Citizen Lab, p. 120 se référant à Arvind NARAYANAN, "Tutorial: 21 fairness definitions and their politics" en ligne : Tutorial: 21 fairness definitions and their politics - YouTube

réalité *quantifiable*. Avec les outils d'IA, est *juste*, est valide en droit, ce qui est *prédit*, ce qui *a été prédit*, ce qui *peut se prédire*. Ils influencent donc la décision de justice à rendre à partir d'une méthode de raisonnement qui diffère de celle qui a cours traditionnellement dans nos tribunaux. En effet, le fonctionnement des outils algorithmiques repose sur une agréation de données et un traitement statistique dont le produit (la suggestion) est une *corrélation*, c'est-à-dire « une parenté statistique entre des nombres » sans pour autant « s'interroger sur la raison qui permet de mettre en rapport les deux phénomènes »; tout fonctionne comme si « les corrélations entre deux phénomènes dans le passé suffisent pour prédire leur continuité dans l'avenir »[251]. Ainsi, la décision *juste* à rendre devient celle qui, statistiquement, est la décision prévisible ou prédite, et ce, même si la prédiction a peu à voir avec la décision *juste* à rendre dans un cas particulier. Ici, la vérité « interprétative » qui soutient traditionnellement le droit, propre à la procédure contradictoire – et qui permet de conduire la quête vers le *juste* – est évacuée au profit d'une vérité « prédictive », qui est simplement une conformité à la réalité statistique quantifiable[252]. Cette analyse est également partagée par la professeure Caroline Lequesne Roth : « [d]ans le cas d'une évaluation algorithmique, le logiciel identifie des concordances de faits, discriminantes au regard des dossiers déjà instruits ; il délivre un score de probabilité d'après une masse de données, appréciant la situation de chacun au travers de paramètres quantitatifs. La normativité juridique « sociale et herméneutique » est ainsi décontextualisée au bénéfice d'une vérité scientifique. »[253] En ce sens, la corrélation statistique s'oppose alors à la *causalité juridique* qui est la méthode de raisonnement propre au système de droit et dont le domaine de vérité est davantage soutenu par notre « imaginaire social »[254] ou collectif que par une quantification factuelle traduite par les lois de la statistique.

Choix vs. compilation (« mass data » et « hidden patterns »). Le domaine de vérité des outils d'IA utilisés aux fins de la justice prédictive ainsi que leur méthode de raisonnement reposent sur le *nombre* de précédents plutôt que sur leur application valide au cas en l'espèce. Par conséquent, ce qui a été décidé dans le plus grand nombre des cas deviendra ce qui sera la décision *juste* à rendre dans le futur, car « la norme issue d'un logiciel prédictif ne procède plus du raisonnement juridique mais de la masse des décisions »[255]. La décision de justice est alors motivée par la mise en corrélation de *tous* les faits disponibles et traités par l'algorithme, plutôt que par une sélection minutieuse de certaines décisions et par une application juridiquement valide établie à l'issue d'un débat contradictoire et équitable. Comme en témoigne la promesse commerciale des concepteurs d'outils d'IA voulant que l'outil soit en mesure de découvrir des *hidden patterns* au sein de la jurisprudence, la décision informée par ces outils résulte, au fond, de l'amalgame de *tous les indices factuels quantifiables* trouvés dans l'*ensemble* des décisions antérieures recensées : « Le droit se transforme donc en corrélation de faits et les faits, corrélés, deviennent, indépendamment de leur légitimité, normatifs »[256]. Pourtant selon le philosophe Paul Ricoeur, l'application *valide* des règles à un cas individuel « ne saurait être réduite à une procédure mécanique », car elle suppose plutôt l'interprétation, la discrimination et le débat contradictoire équitable : **(i)** tout d'abord, un *choix* doit être fait entre des descriptions narratives contradictoires présentées devant toutes les parties concernées dans le cadre du rituel du procès

[251] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 231.
[252] *Id*., p. 101.
[253] Caroline Lequesne Roth, « La science des données numériques au service du contrôle fiscal français », dans Alain PARIENTE (dir.), *Les chiffres en finances publiques*, Mare & Martin, Paris, 2019, par. 13(a).
[254] Nous reprendrons à plusieurs reprises ce concept au cours du rapport. Nous le définissons comme étant la manière dont nous concevons l'espace public, l'ordre commun en société, et la manière dont nous souhaitons l'organiser. Nous l'empruntons au philosophe canadien Charles TAYLOR, *Modern Social Imaginaries*, Duke University Press, 2003.
[255] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 227
[256] Caroline Lequesne Roth, préc., note 254, par. 13.

(*interprétation narrative*)[257]. À notre avis, le modèle épistémologique du droit se distingue ici de la suggestion de l'outil d'IA qui, elle, émet une corrélation statistique générale qui est totalement *extérieure* au procès particulier *en cours* et qui échappe à l'expérience vécue par les parties lors du procès. En plus, **(ii)** la décision de justice implique qu'un *choix* soit fait entre les différentes interprétations données par le passé afin de les harmoniser (*interprétation légale*)[258]. Le modèle épistémologique propre au droit se distingue alors de la compilation statistique de « *toutes* les décisions (ou le maximum) », car « [cette potentialité technique des outils IA] affaiblit le pouvoir de sélection et de distinction exercé par les professionnels. »[259] Face au savoir juridique qui est fondé sur la sélection et la discrimination des faits et des décisions, les outils d'IA proposent un mode de raisonnement nouveau qui vient imposer le « nombre des décisions », la quantité, les *hidden patterns* et la répétition factuelle décontextualisée, en tant que nouvelles conditions à la décision *juste*, à la validité juridique : « Data functions as a mythology: it comes with a "widespread belief that large data sets offer a higher form of intelligence and knowledge." »[260] Les outils d'IA sont alors portés par une tout *autre* forme d'imaginaire social, ce que les auteurs Antoine Garapon et Jean Lassègue appellent le « mythe de l'automatisation ou de la délégation aux machines »[261].

Le juste en tant qu'expérience humaine et symbolique. L'idée du *juste* ferait nécessairement appel à la protection de cette part d'***indéterminé*** au sein de la problématique juridique : cette incertitude qui précède nécessairement toute décision de justice. À l'opposé de cet *indéterminé*, on retrouve les « certitudes » proposées par les outils d'IA. Ces outils seraient en mesure, *avant* même le procès, de déterminer la décision à prononcer. En prédéterminant l'issue finale du procès, les outils d'IA réduisent l'importance de la procédure contradictoire dans la détermination de la décision *juste* à rendre. Les algorithmes prédictifs modifient alors notre rapport spatio-temporel au procès en interchangeant les étapes classiques de sa procédure. « Alors que nous pensions que tout le procès est tourné vers un dénouement final qui dit le droit au terme d'un long processus, la justice prédictive fournit, sinon la solution, du moins une idée très précise de l'issue, avant même de commencer une affaire »[262], ce qui n'est pas sans pervertir la logique derrière la **présomption d'innocence** protégée par l'art. 11d) de la *Charte*. La présomption d'innocence est nécessairement liée à un processus temporel bien défini dictant que l'accusé ne pourra être reconnu coupable qu'*au terme* d'un procès impartial et équitable. De l'*indéterminé* qui caractérisait notre conception de la justice, voilà que la justice se retrouve *prédéterminé* par le traitement algorithmique. Pour la juriste Mireille Delmas-Marty, ce serait « cette **incertitude** même [qui provient de la notion du *juste*] [qui peut] bien corriger, tout compte fait, ce qu'il y a de trop logique ailleurs »[263] comme dans les certitudes « utiles » qui soutiennent la philosophie utilitariste, qu'on retrouve d'ailleurs dans les propositions des outils algorithmiques de prédiction du risque. L'incertitude propre au *juste* permet de corriger ce qu'il y a de « trop logique ailleurs », comme les prétentions issues de l'imaginaire commercial voulant que les outils d'IA soient en mesure de prédire *plus efficacement* la récidive et de *maximiser* la protection de la société. Elle permet de *limiter* certains « excès », notamment ceux liés à la radicalité technique des outils d'IA. C'est là

[257] Paul RICOEUR, *Le juste II*, Esprit, 2001, p. 262
[258] Id., p. 262
[259] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 227.
[260]Angèle CHRISTIN, "Predictive algorithms and criminal sentencing", préc., note 123, p. 274, citant Danah BOYD et Kate CRAWFORD, "Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon," 2012, 15-5 *Information, Communication, & Society*, 662-679.
[261] A. GARAPON et J. LASSÈGUE, préc., note 13.
[262] *Id*., p. 115.
[263] Mireille DELMAS-MARTY, *Les chemins de la répression - Lectures du code pénal,* Presses universitaires de France, Paris, 1980, p. 78.

l' « efficacité propre » à la notion du « juste » en droit criminel. Le concept du *juste*, une fois introduit dans un système de droit, agirait alors « négativement », en ce sens qu'il n'est pas une chose que l'on peut *produire* « positivement »[264], mais à laquelle on arrive par la modération des tendances qui portent vers l'excès, la certitude absolue. Sur ce point, nous dirions que, compte tenu de la polysémie du « juste », nous pourrions difficilement nous satisfaire en droit criminel d'un algorithme programmé en vue de produire « positivement », de manière usinière, des décisions « justes ». Cette part **d'indéterminé** permet la décision *juste* en droit criminel, car elle commande au décideur de faire preuve de **précaution**, de **modération**, de **souplesse** et d'**adaptabilité** aux faits particuliers présentés devant lui et de tempérer les prétentions contradictoires des parties.

Pour les auteurs Antoine Garapon et Jean Lassègue, « **l'imperfection**, aussi paradoxal cela peut-il paraître est une condition du droit. Pour assurer une juste application du droit, il faut garder cette *distance entre l'énonciation de la loi sous forme d'idéalité*, et *sa rencontre avec la réalité*. »[265] Compte tenu que l'indéterminé requiert une juste distance entre l'idéalité à poursuivre (société « juste » et « pacifiée ») et ce qui est suggéré par la réalité statistique (la société « réelle », « scientifisée »), il faudra, pour s'assurer d'arriver à matérialiser la décision *juste*, aménager et protéger « un espace de jeu pour soupeser le pour et le contre et trouver au cas par cas la solution la plus juste »[266]. Cet espace de jeu, cette *distance*, Paul Ricoeur dira qu'il est permis par l'équité de la procédure : « dans [le] cadre cérémoniel [du procès], se déploie un jeu complexe de langage, régi par les règles de procédure qui assurent l'équité requise du procès »[267] et dont le respect permettrait ce « pragmatisme transcendantal » à même de rendre *justice* dans un cas donné[268].

En plus de la nécessité de protéger le rituel contradictoire s'ajoutent d'autres conditions à la formation d'une décision *juste*, notamment la position symbolique du juge humain en tant que tiers : « Juste distance [entre les parties par la procédure], médiation d'un tiers, impartialité s'énoncent comme les grands synonymes du sens de la justice »[269]. Ainsi, la décision *juste* doit être **informée par le rituel du procès**, et plus substantiellement par **l'expérience vécue par les parties présentes** lors du procès. En effet, « la fonction primaire du procès est de transférer les conflits [entre les parties] du niveau de la violence à celui du langage et du discours. »[270] Ainsi le procès, et le sentiment de justice qui en découlera, ne peut se passer du langage, du discours, de la procédure contradictoire entre les parties et de la formation de *sens* qui découle de ce rituel judiciaire. Les parties ne sauraient se contenter des réponses prédéterminées calculées par l'outil d'IA; elles ne sauraient se satisfaire de celles-ci en raison de l'opacité « quasi-inaccessible » et désymbolisé du langage informatique[271]. Elle prive les parties de ce que Paul Ricoeur décrit comme étant un transfert de la violence vers le langage, le discours, l'échange.

Puisqu'on ne peut déterminer à quel point la prédiction de l'outil d'IA a pu influencer le juge dans sa décision, et en raison de l'opacité d'un décideur algorithmique et de l'impossibilité technique d'identifier tous les biais lorsqu'il fonctionne par apprentissage-machine, les caractéristiques mêmes du raisonnement par un outil d'IA pourraient attenter au *principe de droit* voulant qu'en droit criminel l'accusé et le public

[264] *Id.*
[265] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 163
[266] *Id.*
[267] P. RICOEUR, préc., note 258, p. 261.
[268] *Id.*
[269] P. RICOEUR, préc., note 258, p. 13.
[270] *Id.*, p. 261
[271] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 43-47.

ont le droit d'obtenir des **décisions motivées** qui sont « raisonnablement intelligibles pour les parties » et qui permettent à la Cour d'appel d'évaluer la « *justesse* » du raisonnement[272]. Contrairement au *sens* produit lors du procès, la prédiction produite par l'outil IA est une décision qui est purement extérieure à l'expérience vécue des parties lors du procès, elle est une « pure corrélation mathématique qui n'est pas fondée sur ce que les sujets vivent dans les interactions sociales auxquelles ils participent. »[273]

Ensuite, pour avoir le sentiment que la justice a été rendue, la décision de justice doit, non seulement provenir d'un tiers symbolique qui préside le rituel, mais également d'un juge qui est un **être *humain***, un être « semblable » aux parties. Afin de s'assurer que les parties puissent reconnaitre la légitimité du décideur et de sa décision, ils doivent pouvoir se reconnaitre en lui; ils doivent pouvoir *partager* avec lui une même expérience du réel et un même mode de pensée. La *pensée* humaine et le *calculé* de l'IA diffère encore aujourd'hui en plusieurs points[274]. L'IA n'aurait pas non plus la même facilité à produire des jugements moraux que les humains. Leurs jugements moraux diffèrent de ceux des humains, autant au niveau de la forme (abstraction humaine vs. Règles encodées sous forme de oui/non) que de la qualité du jugement : « Moreover, it is currently impossible to capture human values like empathy and compassion in algorithms. Ethics are too complex to transfer to a computer code (Moor 2006). »[275] Un autre facteur essentiel à la légitimité de l'autorité du tribunal en droit criminel serait précisément cette capacité pour le décideur de maintenir un corridor communicationnel avec les différents acteurs au procès à travers lequel il peut s'exprimer en termes symboliques et moraux et, ainsi, traiter les parties comme des agents moraux à part entière. Pour assurer la pérennité de la fonction symbolique du droit criminel, il est important que le procès donne cette occasion à un juge humain de communiquer avec les parties, *d'agent moral à agent moral* : « That process—of one robed judge and one convicted defendant in conversation—has moral value in and of itself, and the addition of an interloping machine may undermine that function (Donohue 2019). (…) In criminal justice, the lack of human interaction and the "dehumanizing" effect this could bring be an insurmountable problem (for a different view, see Bagaric, Hunter, and Stobbs (2019))»[276]. Pour le philosophe Paul Ricoeur, le juge est nécessairement « un citoyen ordinaire », « un être humain comme nous »[277], mais qui bénéficie d'une *élévation symbolique* qui lui permet de prononcer la décision *juste*, de *rendre justice*. C'est cette position de *tiers/semblable* qui permet de « donner chair à la justice », en effet, « [les juges] sont la *bouche* de la justice. »[278] Ce qu'exprime ici le philosophe, c'est que la quête du *juste* ne peut se passer d'un rapport humain et de sa dimension symbolique. Le *juste* n'évoque rien à la machine : « Que serait en effet une justice qui ne s'adresserait pas au sentiment de justice, à ce qu'il y a d'humain dans les

---

[272] *R*. c. *Sheppard*, 2002 CSC 26; Sadie CREESE, « The threat from AI », préc., note 221, p. 208 : « [C]urrently there is no practical way to establish sophisticated properties of ML [machine-learning] decisions, such as absence of bias or influence from specific mindsets, as it is incredibly hard (impossible, some would say) to even specify those requirements in terms of the data-points and features being reasoned about (and so interrogated when considering explicability of automated decisions). »

[273] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 225.

[274] John L.M. MCDANIEL et Ken G. PEASE, « Introduction » dans John L.M. McDaniel et Ken G. Pease (Dir.), *Predictive Policing and Artificial Intelligence*, Routledge, 2021, p. 18, et à la p. 16 : « The feedback could cause a machine learning algorithm to change so much that the new rules and variables it develops may render it entirely different, in look and function, to the form that it originally took. Machine learning algorithms are celebrated precisely because of this unpredictability. One of the things that separates current machine learning algorithms from human thinking is that the algorithmic process often bears no comparison to human-like deductive reasoning. ». L'IA ne serait pas non plus en mesure de produire des jugements moraux comme ceux des humains, Sigrid VAN WINGERDEN et Mojca M. PLESNIČAR, « Artificial Intelligence and Sentencing Humans against Machines », Chapitre 12, p. 238 dans J. Ryberg et J. V. Roberts (Dir.), *Principled Sentencing and Artificial Intelligence*, Oxford University Press, New York, 2022.

[275] Sigrid van WINGERDEN et Mojca M. PLESNIČAR, préc., note 275, p. 238

[276] *Id*., p. 245.

[277] P. RICOEUR, préc., note 258, p. 261.

[278] *Id*.

humains ? »[279] Le juge doit *décider*, malgré sa **faillibilité**. C'est là que se trouve toute la quête de « justice » en droit criminel; la résistance aux excès, la précaution face à sa propre subjectivité et à sa propre émotivité ainsi qu'un appel public à la modération et à la relativisation des attentes vindicatives. La délégation aux machines nous prive de cet exercice, cette ascèse, ce « pragmatisme transcendantal », et, même, nous en éloigne, car elle ne fait, finalement, que compiler et multiplier statistiquement la faillibilité individuelle de tous les autres juges qui ont précédé.

Face à la faillibilité humaine individuelle qui caractérise la décision de justice, à cette approche imparfaite, mais humaine, équilibrée et informée par le débat contradictoire et agissant davantage sur le plan symbolique s'oppose la « radicalité de la norme technique » et l'aura d' « infaillibilité » des outils d'IA : « On peut parler de radicalité de la technique lorsque cette **indétermination** ou la possibilité de modifier la règle sont interdites par la technique, c'est-à-dire par une raison étrangère au droit. »[280] Les outils d'IA par leur promesse d'objectivité, nous offre une « hyper-perception du réel » (promesse commerciale de découvrir des *hidden patterns*), mais tout en demeurant très limités, car ceux-ci se limitent au « réel quantifiable »[281]. Cette hyper-proximité avec *une partie* du réel coupe le degré nécessaire de distance et d'abstraction pour permettre à l'institution pénale de parler en termes **symboliques** : « Les logiciels ne pourront jamais prédire de manière déterministe la totalité de ce qui peut se produire en raison de la limitation interne propre à la notion même de calculabilité. Et l'intelligence artificielle n'échappe pas à la règle. »[282] L'idée du *juste* échappe à la vérité statistique, car elle requiert une certaine part d'indéterminé. C'est en ce sens que certains juges aux États-Unis délaissent le recours à la statistique pour préserver leur intuition formée par leur expérience (« foot-dragging »)[283]. À ce sujet, la professeure et sociologue Angèle Christin rapporte les propos d'une juge sénior aux États-Unis, qui avance qu'au nombre des choses qu'on ne peut quantifier figure la décision *juste* : « You can take the same case, with the same defendant, the same criminal record, the same judge, the same attorney, the same prosecutor, and get two different decisions in different courts. Or you can take the same case, with the same defendant, the same judge, etc., at a two-week interval and have completely different decision. *Is that justice? I think it is.* » (nos italiques) Ce que les outils algorithmiques ne peuvent produire positivement, comme il s'agit d'une quête symbolique et humaine, c'est justement la décision *juste*, comme l'expliquent Antoine Garapon et Jean Lassègue : « la préservation d'un indéterminé non calculable est nécessaire à la construction d'un sens juridique. Cet élément est en effet indispensable à la constitution du droit comme forme symbolique. Le droit n'est pas réductible à la somme des prescriptions positives, car il offre également la capacité de transformer ces dernières tout en gardant son pouvoir prescriptif. Or, le calcul prédictif l'enferme et le prive de cette capacité. »[284]

---

[279] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 135.
[280] *Id*., p. 161
[281] *Id*., p. 133
[282] *Id*., p. 235.
[283] Angèle CHRISTIN, « Algorithms in Practice : Comparing Web Journalism and Criminal Justice », (2017) 1-14 *Big Data & Society* 1, p. 9. Loin d'effacer la discrétion et la subjectivité dans la prise de décision, les outils algorithmiques "déplacent" cette discrétion vers des parties moins visibles du processus judiciaire, nous pensons aux policiers, aux avocats, aux experts cliniques ou aux agents de probation, *cf*. Sarah BRAYNE et Angèle CHRISTIN, "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts", *Social Problems*, Oxford university, 2020, p. 13. On pourrait éventuellement observer le même phénomène de "résistance" face aux décisions algorithmiques chez les différents agents qui y auront recours, par exemple les agents de probation, *cf*. T. Scassa, préc., note 86, à la p. 9 se référant à Jennifer RASO, "The In-Between Space of Administrative Justice: Reconciling Norms at the Front-Lines of Social Assistance Agencies," dans Jason NE Varuhas et Shona Wilson Stark (dir.), *Frontiers of Public Law*, Hart Publishing, Oxford, 2020, p. 482. Cette étude cherchait à évaluer comment s'exerce la discrétion décisionnelle ou interprétative des agents de l'assistance sociale chez *Ontario Works* en face d'indications normatives concurrentes.
[284]A. GARAPON et J. LASSÈGUE, préc., note 13, p. 237.

Impartialité et déresponsabilisation. Les promesses qui entourent la faculté « prédictive » de ces nouveaux outils menacent également l'impartialité du juge protégée à l'art. 11d) de la *Charte*. Le juge subit la pression de ses pairs par l'entremise de la proposition de l'algorithme qui, elle, est fondée sur ce que les autres juges ont décidé par le passé. Nous pouvons bien concevoir qu'il serait difficile pour un juge, *seul*, de déroger à la prédiction de l'algorithme qui lui indique, non seulement qu'un grand nombre de juges ont par le passé tranché de telle manière, mais que ceux-ci auraient également jugé de telle manière dans le cas qui se présente devant lui : « Cette « éclipse du jugement », empiétant sur notre autonomie délibérative, repose pourtant sur un prérequis qui soulève son lot de difficultés : la dépendance des outils intelligents sur les données massives (Big Data). »[285] En plus, le juge ou le professionnel peuvent ressentir une pression qui provient de l'aura de scientificité et d'objectivité de la machine qui le contraindrait à décider conformément à sa prédiction : « The quantitative assessment provided by a software program always seems more reliable, "objective," and scientific than other sources of information, including one's feelings about an offender »[286]. Dépourvus d'une formation en statistique, le juge ou l'expert témoin peuvent alors ne pas se sentir outillés pour contester les formules à la base de la suggestion de l'algorithme et pour motiver une décision qui serait contraire à l'outil d'IA, même si son expérience ou la réalité du cas particulier lui commandent d'y déroger[287]. On parle alors d'un phénomène de « déresponsabilisation des décideurs » : « Sous l'effet performatif des algorithmes, la recommandation se substitue à la décision humaine; la prétendue objectivité algorithmique à l'État de droit. »[288] Le problème résultant de cette forme de dépendance aux réponses de l'outil d'IA, c'est que son *code source* a tendance à figer pour l'avenir une seule conception du *juste* qui, elle, peut ne pas être applicable au cas particulier présenté devant le juge : « Cette normativité algorithmique conduit à raisonner, non plus en fonction de la situation singulière des individus, mais plutôt à partir d'inférences, de corrélations et de prédictions (…) »[289] Les décisions prises en droit criminel commandent pourtant l'**individualisation** des décisions, surtout au niveau de la peine. Pour ces raisons, on craint alors que les outils d'IA ne créent une « uniformisation des pratiques » et n'entrainent un « effet moutonnier » face à la contrainte normative qu'offre la décision *juste prédictible*[290] : « algorithms tend to have a performative quality: they contribute to create the situation they describe. »[291] Les outils d'IA utilisés à des fins prédictives tendent vers le *statu quo* : l'état du droit et notre conception du *juste* se retrouvent ainsi figés, aucune évolution et aucune dérogation ne sera permise. Déléguer complètement, ou en partie, la prise de décision en matière criminelle à un algorithme aurait pour effets néfastes de déresponsabiliser en tout ou en partie le juge par rapport aux conséquences directes de sa décision sur l'accusé et de « déshumaniser » le processus pénal, ce qui soulève un sérieux problème éthique, particulièrement dans le cas de la détermination et du prononcé de la peine :

> « Moreover, it might be too easy for humans to shift such responsibility to AI. Relieving humans from serious consideration of the morality of punishment would allow us to inflict pain (legitimate and legal, but still pain) without feeling in any way responsible (cf. Floridi et al. 2018). (…) There are crucial features of sentencing that are so inherently *human* that we cannot imagine them being successfully replaced by AI. Making moral judgments not only requires us to consider various

[285] *Id*., p. 280; Karim BENYEKHLEF et Jie ZHU, « Intelligence artificielle et justice : Justice prédictive, conflits de basse intensité et données massives », (2018) 30(3) *CPI* 789, p.809-810, en ligne : https://www.lescpi.ca/articles/v30/,
[286] A. Christin, "Predictive Algorithms and Criminal Sentencing", préc., note 123, p.287
[287] *Id*.
[288] S. DU PERRON et K. BENYEKHLEF, préc., note 86, p. 42.
[289] *Id*., p. 19.
[290] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 239.
[291] Angèle Christin, "Predictive Algorithms And Criminal Sentencing", préc., note 123, p. 279

> justifications for punishing people *but also makes us take responsibility for our actions. If we leave sentencing to AI—are we not losing the very essence of what makes sentencing a human process*?»[292] (nos italiques)

Il y a effectivement une énorme contradiction, susceptible même d'annihiler la valeur symbolique du processus d'imputabilité et du châtiment, dans le fait de responsabiliser un accusé pour ses actions alors que le juge, lui-même, n'est pas prêt à prendre la responsabilité du châtiment imposé et des conséquences sociales, psychologiques et physiques qui en découleront. Il est crucial, pour prononcer une peine qui soit dignement humaine, qu'un décideur *humain* soit pleinement impliqué dans la détermination du degré de sévérité de la peine. C'est uniquement de cette manière que nous pouvons *espérer* que la peine soit mesurée et modérée, puisque le décideur humain pourrait se sentir davantage concerné par la peine à prononcer - étant lui-même humain, sujet à la souffrance, et pouvant lui-même un jour se retrouver à la place de l'accusé, étant sujet au châtiment, ce qui n'est pas le cas de la machine dans notre droit actuel. Ainsi, nous assurons la possibilité que le décideur soit soucieux et préoccupé par les conséquences qui découleront de sa décision et qu'il les comprenne plus *profondément et substantiellement* que ne peuvent le faire les outils d'IA actuels. Nous conserverons alors la possibilité que le décideur *limitera* la sévérité de la peine de manière qu'il arrive à être *en paix*, dans son âme et conscience, avec le degré de souffrance qui sera infligé à l'accusé.

Privatisation de l'expérience de la justice. Dans les situations où l'outil d'IA servirait non pas à informer le juge de la décision à rendre, mais à informer un accusé dans sa décision d'aller à procès ou non en fonction des probabilités qu'un juge le reconnaisse coupable, la relation entre le justiciable et la justice pénale se retrouve également complètement modifiée. D'une quête de justice, nous passons à un exercice visant à peser les *pour* et les *contre* d'aller en procès à la lumière de ses chances de succès. Cela a pour conséquence d'encourager une pratique déjà bien répandue de négociation sur le plaidoyer (*plea bargaining*) – reconnue et dont la protection a été renforcée par la Cour suprême[293] – ce qui contribue à une privatisation de l'expérience de la justice : « Avec les outils prédictifs, l'enjeu n'est plus la décision juridique, mais la résolution sociale de l'affaire par le biais d'une négociation. (…) dès lors, puisqu'on sait comment le futur va se dérouler, autant considérer qu'il existe déjà ici et maintenant, ce qui transforme du même coup la prédiction en *injonction*. »[294] L'expérience de la justice ne se fait plus dans l'espace public du tribunal, mais dans les échanges privés entre avocat et procureur (*justice de couloir*) : « En se concentrant sur l'efficacité et en reversant la solution des conflits du côté du privé, elle manque la nécessaire participation collective à la résolution des conflits ou de désordres sociaux. Elle manque également l'occasion donnée au peuple de remettre en jeu la légitimité des institutions et l'occasion pour elles de la regagner le cas échéant. » La justice prédictive nous prive de l'expérience symbolique de la « maitris[e] collective de la violence »[295], de la « tentative institutionnelle de surmonter la violence par le discours. »[296] L'introduction d'outils actuariels algorithmiques servant à prédire le risque de récidive à l'étape de la remise en liberté pourrait avoir pour conséquence de « privatiser » l'étape de l'enquête sur remise en liberté, d'exacerber son caractère « négocié » et de réduire l'accès à un cautionnement raisonnable auquel l'accusé a le droit en vertu de l'article 11e) de la *Charte*. En effet, en raison de la complexité inhérente liée à la contestation des résultats

[292] Sigrid van WINGERDEN et Mojca M. PLESNIČAR, préc., note 275, p. 247 se référant à Estcourt, A., et K. Marr. 2019. "Thinking Machines and Smiley Faces." Australian Law Journal 93 (10): p. 855-865, p. 856 et à Floridi, L., et al., 2018. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." Minds and Machines 28 (4): p. 689-707.
[293] *R.* c. *Anthony-Cook*, 2016 CSC 43.
[294] A. GARAPON et J. LASSÈGUE, préc., note 13, p. 240-241.
[295] *Id*., p. 190-191
[296] P. RICOEUR, préc., note 258, p. 262.

d'un outil d'IA, du coût qui pourrait en découler et du temps supplémentaire que l'accusé devra purger en détention durant ces procédures, l'accusé pourrait être dissuadé de contester des conditions déraisonnables de remise en liberté proposées par la Couronne à la lumière des résultats prédits par un outil actuariel d'IA. L'introduction de ces outils à cette étape exacerberait les problématiques actuelles de cette procédure; celle-ci étant déjà critiquée actuellement au Canada comme conférant trop de pouvoir à la poursuite et comme étant trop discrétionnaire, machinale et expéditive[297].

3.3. Application des considérations sur le *juste* aux outils d'IA lors des différentes étapes du procès

Plusieurs principes du droit que nous avons dégagés de la notion du *juste* dans la précédente partie sont susceptibles d'être protégés par la *Charte canadienne* en vertu de l'article 11d) qui prévoit, outre la présomption d'innocence, le droit d'être jugé par un tribunal indépendant et impartial, ainsi qu'en vertu de l'article 7 qui prévoit le droit à une procédure contradictoire et équitable. Nous avons avancé plus généralement que le droit à une décision rendue par un tiers humain et impartial, à la suite d'une procédure contradictoire, et le droit à une décision qui n'est pas prédéterminée d'avance, donc qui est adaptée aux circonstances particulières de l'accusé, se retrouvent également dans les nombreuses références, dans notre droit criminel, au droit à une peine « juste », à un cautionnement « raisonnable » et à des mesures correctionnelles ou des conditions de libération qui sont « raisonnables » et « humaines ». Dans la présente partie, nous nous intéresserons à la résonnance particulière que prend la notion d'une décision *juste* en droit criminel au cours des différentes étapes où un outil d'IA pourrait être introduit et comment ces références à la notion du *juste* devraient venir tempérer l'enthousiasme qui entoure le recours aux outils d'IA aux fins de la justice prédictive.

3.3.1. Introduction générale sur la fonction de l'institution pénale au Canada

La fonction normative qui est propre à l'institution pénale n'est ni de gérer, ni d'administrer les risques liés à la criminalité en société, ni encore de protéger avec effectivité ses membres (*fonction managériale et administrative du risque*). Au Canada, on lui réserve généralement une fonction beaucoup plus limitée. En raison des moyens limités dont elle dispose – par essence, elle se limite à exprimer des valeurs par l'infliction mesurée de souffrance[298]. Sa fonction normative propre est principalement symbolique. En raison de l'incertitude liée à l'« efficacité » de ses moyens symboliques, les interventions de l'institution pénale sont généralement régies par un principe de modération et de restriction[299]. L'institution pénale a comme fonction normative propre de « *contribuer* », par ses interventions, au maintien d'une société paisible, sûre *et* juste[300]. Ce que l'institution pénale poursuit, ce serait ce « point de balance spécifique » entre la *Justice* et la *Sécurité* : ainsi, la gestion des risques et le contrôle du taux de criminalité ne devrait jamais être à même de

---

[297] M. SYLVESTRE, N. BLOMLEY, et C. BELLOT, *Red Zones: Criminal Law and the Territorial Governance of Marginalized People*. Cambridge: Cambridge University Press, 2020. Concernant l'impact de l'introduction de ces outils sur le droit de l'accusé à accéder à un cautionnement raisonnable, voir Kate Robertson et Jill R. Presser, "Algorithmic Technology and Criminal Law in Canada", préc., note 229, p. 101.

[298] *R.* c. *M. (C.A.)*, [1996] 1 RCS 500, par. 81; CANADA, Department of Justice, *The Criminal Law in Canadian Society*, Ottawa, 1982, p. 39.

[299] COMITÉ CANADIEN DE LA RÉFORME PÉNALE ET CORRECTIONNELLE, *Rapport du Comité canadien de la réforme pénale et correctionnelle : justice pénale et correction : un lien à forger*, Imprimeur de la Reine pour le Canada, Ottawa, 1969, par. 11 (ci-après « Rapport Ouimet »); COMMISSION DE RÉFORME DU DROIT DU CANADA, *Les principes de la détermination de la peine et du prononcé de la sentence*, document de travail 3, Ottawa, Mars 1974, p. 1 (ci-après « Rapport Hartt »)

[300] Art. 718 C.cr.; CANADA, *The Criminal Law in Canadian Society*, préc., note 299, p. 5.

compromettre la mission de justice plus large poursuivie par l'institution pénale. Cette dernière cherche à *renforcer* l'autorité morale de la loi; elle doit maintenir un ordre *symbolique* de justice, de paix et de sécurité *par l'expression de décisions* « *justes* ». La notion du *juste* agit d'ailleurs comme un « frein » aux interventions extensives de l'institution pénale qui viseraient principalement à assurer la sécurité (philosophie utilitariste)[301]. Depuis 1996, cette fonction normative particulière de l'institution pénale se retrouve à l'article 718 *C.cr.* et dans les objectifs de la peine. Elle confirme l'idée que le moyen symbolique de l'institution pour instaurer l'ordre, *la peine*, demeure limité et ne peut, à lui seul, garantir la sécurité des membres de la société. Déjà en 1972, la *Commission Hart* prêtait à l'institution pénale une fonction essentiellement symbolique : « Est-il réaliste de s'attendre à ce que la loi fasse plus que souligner la gravité de l'infraction, et, au moyen d'une variété de peines, affirmer, confirmer et protéger les valeurs fondamentales de la communauté? »[302]

Néanmoins, le développement d'un nouvel arsenal des peines axées sur le contrôle du risque, les attentes de plus en plus élevées de la population en matière de sécurité (comme le révèle le développement de politiques pénales populistes[303]) et la croyance renouvelée en la capacité de prédire le risque de récidive grâce au développement de nouvelles technologies actuarielles sont susceptibles de modifier la fonction propre de l'institution pénale. C'est dans ce contexte que s'inscrirait l'introduction des outils d'IA aux fins de prédiction du risque dans les diverses étapes du processus pénal. Pour certains auteurs, même si les juges ne sont pas les mieux placés pour prédire et administrer les risques en société, et même s'il n'existe pas encore d'études scientifiques significatives nous permettant d'établir que ces outils sont en mesure de réellement les prédire, l'engouement actuel envers ces outils d'IA témoigne de l'importance grandissante que prend désormais le « contrôle des risques » au sein de l'exercice décisionnel en matière criminelle, notamment au niveau de la peine[304]. Le *Département de la Justice des États-Unis* s'inquiétait d'ailleurs de ce changement de fonction en 2014; la peine autrefois imposée en raison d'un acte passé serait maintenant imposée en raison des risques de préjudices futurs[305]. Le Canada n'est pas à l'abri de cette reconceptualisation de la fonction de l'institution pénale. Nous pensons que la prédiction du risque suggérée par les outils actuariels, algorithmiques ou non, ne peut jouer qu'un rôle limité et ne devrait pas être déterminante dans la quête d'une décision *juste*. C'est qu'il existe une différence fondamentale entre la promesse d' « efficacité » des outils d'IA et la quête essentielle et plus large visant à rendre une décision *juste* :

> « Even if a perfectly accurate algorithm does exist, the *fairness-as-accuracy* definition might still come up short in the event that an algorithm leads to generalizations about particular groups. Consider an accurate algorithm that comes to the blanket conclusion that men tend to deserve higher risk scores than

[301] *Id*., p. 40-41.
[302] Rapport Hartt, p. 6.
[303] Hélène DUMONT, « Chronique canadienne - Une décennie de populisme pénal et de contre-réformes en matière punitive au Canada », 1-1 *Revue de science criminelle et de droit pénal comparé*, 2011, pp. 239-252.
[304] Danielle Kehl, Priscilla Guo et Samuel Kessler, préc., note 212, p. 27 : "Although it has long been clear that managing risk is a part of the sentencing consideration, the use of these algorithms almost certainly increases the prominence of risk assessments in the decision-making process. Yet judges might not be the best or most appropriate actors to try to manage these risks. Nor is there a significant body of evidence at this point that suggests we are actually good at predicting or managing risk—or that longer sentences, for example, might decrease the risk of recidivism."
[305] *Id*. p. 7, se référant à Jonathan WROBLEWSKI, *Letter from Jonathan Wroblewski, Director, Office of Policy and Legislation, U.S. Department of Justice's Criminal Division, to the Honorable Patti B. Saris, Chair, U.S. Sentencing Commission,* 29 Juillet 2014, p.8, en ligne : https://www.justice.gov/sites/default/files/criminal/legacy/2014/08/01/2014annual-letter-final-072814.pdf : "Determining imprisonment terms should be primarily about accountability for past criminal behavior. (…) As analytics evolve, we are concerned about the implications of sentencing policy moving away from this precept."

> women. Whether or not the algorithm is ***accurate***, would it be ***fair*** for individuals to be judged based on immutable characteristics such as gender? »[306]

La prédiction du risque par les outils actuariels, algorithmiques ou non, ne dit pas tout. Elle ne répond pas à l'exercice particulier qu'est la détermination de la décision *juste*. Les outils actuariels répondent à une logique *autre* que celle propre au *juste*, ils visent uniquement à permettre de contrôler un risque statistique, quantifiable et pouvant faire l'objet d'une prédiction : « Risk-assessment tools, in this view, function as a technology of governmentality: they operate at a distance, through statistical analysis, defining classes of individuals who are more or less "risky" and should be controlled more or less forcefully. »[307] En somme, nous craignons que leur intégration massive au sein du processus pénal ne transforme pour de bon la fonction de l'institution pénale. L'outil d'IA de prédiction du risque, en mettant l'emphase sur les risques prédits que la société court, est alors à même de venir imposer au juge une certaine conception de son rôle lors de la détermination de la peine : « risk-assessment tools are not the value-neutral objects that advocates paint them to be: they crystallize specific political ideas about the role of punishment. »[308] Selon Angèle Christin, l'extension du recours à ces outils participe à la consécration d'une *culture du contrôle du risque* au sein de l'institution pénale; ces outils mettraient de l'avant un recours massif à l'emprisonnement et prioriseraient systématiquement l'objectif pénologique de la neutralisation des personnes à risque de récidiver[309]. Nous verrons comment la quête d'une décision *juste* en droit canadien devrait venir tempérer le poids des prédictions algorithmiques sur le risque de récidive lors des différentes procédures où elles pourraient potentiellement être admises.

3.3.2. À l'étape du prononcé de la peine

En droit pénal canadien, le prononcé d'une peine « juste » et « proportionnelle » constitue le principe « fondamental » de l'exercice de détermination de la peine[310]. Face à la radicalité et l'unilatéralité des réponses offertes par l'outil actuariel algorithmique, l'approche canadienne en matière de peine se veut historiquement précautionneuse et empreinte de modération en raison de l'incertitude liée à l'efficacité symbolique de la peine[311]. Le juge bénéficie historiquement d'un large pouvoir discrétionnaire afin de lui donner toute la latitude nécessaire pour effectuer cet exercice difficile de la détermination de la peine, où la peine devra être proportionnelle à la « gravité du crime » commis par l'accusé et à sa « responsabilité morale » : « Le juge ne dispose pas d'un critère ou d'une formule d'application simple à cet égard. Il faut s'en remettre au jugement et à la sagesse du juge qui détermine la peine, que le législateur a investi d'un pouvoir discrétionnaire considérable à cet égard à l'art. 718.3. »[312] Ainsi, la prédiction du risque futur ne devrait

---

[306] *Id*., p. 30. Cette réflexion fait écho au problème éthique de déresponsabilisation des juges au regard de la souffrance infligée qu'il y aurait dans l'éventualité où ils seraient remplacés par une IA, tel que soulevé dans la partie précédente (II.3.2.). Même si un jour l'IA devenait plus fiable ou plus efficace pour prédire les risques de récidive que le jugement par un humain, il demeure que l'on perdrait cette composante humaine indispensable au prononcé de la peine et l'apport symbolique au maintien d'un ordre *juste* qui, d'ailleurs, se révèle comme étant la fonction normative propre à l'institution pénale, v. Sigrid VAN WINGERDEN et Mojca M. PLESNIČAR, préc., note 275, p. 247.
[307] A. Christin, "predictive Predictive Algorithms and Criminal Sentencing", préc., note 123, p. 286.
[308] *Id*., p.287.
[309] *Id*., p.286.
[310] Art. 718 et 718.1 *C.cr.*; *R.* c. *Ipeelee*, 2012 CSC 13, par. 37 : « La proportionnalité représente la condition *sine qua non* d'une sanction juste. »
[311] COMMISSION CANADIENNE SUR LA DÉTERMINATION DE LA PEINE, *Reformer la sentence: une approche canadienne*, Ottawa, Justice Canada, 1987, p. 160-168 (« Rapport Archambault »); *R.* c. *M. (C.A.)*, [1996] 1 RCS 500, par. 78; *R.* c. *Nasogaluak*, 2010 CSC 6, par. 42; *R.* c. *Ipeelee*, 2012 CSC 13, par. 37.
[312] R. c. Proulx, 2000 CSC 5, par. 116.

avoir qu'une portée négligeable lors de la détermination de la peine, puisque le risque futur n'est pas pris en compte dans la mesure fondamentale de la sévérité de la peine en droit canadien : « Imposing a severe sentence of imprisonment on such an offender because he is considered to be likely to commit serious crimes in the future runs the risk of violating the principle of proportionality in sentencing »[313]. Contrairement aux outils d'IA dont le fonctionnement répond uniquement à l'objectif de neutralisation des risques et de la protection effective de la société, le prononcé de la peine est régi par le principe fondamental de proportionnalité et le juge, tout en devant le respecter, peut également poursuivre d'autres objectifs comme la dénonciation, la dissuasion, la réhabilitation, la réparation et la conscientisation (art. 718.2 C.cr.) : « most predictive instruments emphasize one major justification at the detriment of the others: incapacitation, that is, a view of justice based on estimating the risk to society posed by the offender when deciding on a sentence designed to incapacitate dangerous individuals. »[314] Si les prédictions des outils actuariels peuvent être une donnée intéressante dans les cas où un juge souhaiterait poursuivre l'objectif de neutralisation, le recours aux outils de prédiction du risque s'avère généralement peu pertinent pour informer le quantum de la peine dénonciatrice ou dissuasive[315]. D'ailleurs, en amenant systématiquement le juge à s'engager dans la réalisation de l'objectif de neutralisation des risques en fonction des risques prédits, ces prédictions peuvent avoir pour effet normatif de faire perdre de vue le principe de retenue stipulant que l'incarcération doit être la mesure de *dernier recours* (718.2(d) C.cr.).

Contrairement aux outils d'IA qui fonctionnent par des généralisations et des inférences statistiques en prenant en considération des données qui sont extérieures à l'accusé, le prononcé de la peine en droit canadien commande, lui, une évaluation individualisée. Le principe d'individualisation de la peine a été reconnu par la Cour suprême comme étant nécessaire afin d'atteindre la proportionnalité de la peine[316]. En face de ces principes fondamentaux, nous avons une « normativité algorithmique [qui] conduit à raisonner, non plus en fonction de la situation singulière des individus, mais plutôt à partir d'inférences, de corrélations et de prédictions (…) »[317]. En effet, l'outil d'IA est programmé pour dresser des corrélations à partir des caractéristiques *généralisables* de l'accusé, donc des traits de l'accusé qui sont partagés par des groupes sociaux définis : « Following this statistical line of reasoning, defendants are then sentenced based on their *belonging to a specific group with "risky" characteristics* rather than because of their *individual actions*, which goes against the jurisprudential value of individualism. »[318] Dans *State* v. *Loomis* aux États-Unis[319], la Cour a émis quelques directives visant à encadrer et limiter le recours aux logiciels de prédiction du risque, comme COMPAS, pour déterminer la peine, notamment en raison de son impertinence pour déterminer la peine « juste » à appliquer à un accusé particulier[320]. Il a été jugé qu'un tel outil ne devait ni

[313] Julian V. ROBERTS et Simon VERDUN-JONES, "Directing Traffic at the Crossroads of Criminal Justice and Mental Health: Conditional Sentencing after the Judgment in Knoblauch", (2002) 39-4 *Alta L Rev* 788, p. 793.
[314] A. CHRISTIN, "Predictive Algorithms and Criminal Sentencing", préc., note 123, p. 286.
[315] Danielle KEHL, Priscilla GUO et Samuel KESSLER, préc., note 212, p. 13-14
[316] R. c. Proulx, 2000 CSC 5, par. 82; *R.* c. *Nasogaluak*, 2010 CSC 6; *R.* c. *Ipeelee*, 2012 CSC 13.
[317] S. DU PERRON et K. BENYEKHLEF, préc., note 86, p. 19; Kate Robertson et Jill R. Presser, "Algorithmic Technology and Criminal Law in Canada", préc., note 229, p. 116 : "Algorithmic methods tend to be generalized inferences by definition. (…) Although algorithmic risk prediction technologies generate estimates in relation to a particular individual, the estimate of risk is generated on the basis of a fixed list of predetermined criteria for comparison, which is then compared to historical bulk data from past cases. The algorithm has no opportunity to hear submissions, context, or specific facts of the case, including mitigating circumstances or an extensive wealth of information about the individual that would be relevant to an individualized assessment of risk."
[318] A. CHRISTIN, "Predictive Algorithms and Criminal Sentencing", préc., note 123, p. 281.
[319] 881 N.W.2d 749 (Wisc. 2016)
[320] Danielle KEHL, Priscilla GUO et Samuel KESSLER, préc., note 212, 2017 : "While the risk score can help a judge better understand a defendant's unique situation and relevant factors, the court held that it should not be used to determine the length or severity of the punishment, and certainly should not be counted as an official aggravating or mitigating factor in a sentencing decision. (…) Its

être le seul facteur pour déterminer la peine, ni même un facteur « déterminant » de la peine[321]. Néanmoins, plusieurs critiques ont été formulées contre ce jugement, ces directives n'étant pas à même de restreindre les effets néfastes de l'utilisation d'un tel système. Tout d'abord, les juges ont permis le recours à ce logiciel même s'il prend en considération des données, susceptibles d'amoindrir l'égalité des protections accordées par la loi, comme le *statut socio-économique* et le *genre*, puisque ces facteurs permettent de mieux prédire le risque[322]. Ensuite, il est difficilement concevable que le juge reste hermétique à la prédiction du risque lorsque le risque prédit est élevé; la prédiction a un effet normatif et deviendrait alors un facteur déterminant de la décision[323].

Sous le couvert d'en arriver à une « prédiction » du risque de récidive plus précise et de permettre un contrôle plus efficace des risques, les outils de prédiction du risque intègrent à leur algorithme un large éventail de caractéristiques *générales* et *immuables* (âge, genre, éducation, revenus, emploi, mode de vie, etc.) susceptibles de reconduire systématiquement des désavantages préexistants et de résulter en des peines plus sévères en raison de la seule appartenance à un groupe social[324] : « As former US Atorney General Eric Holder points out, "By basing sentencing decisions on *static factors* and immutable characteristics—like the defendant's education level, socioeconomic background, or neighborhood—they may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society." »[325] (nos italiques) Les outils de prédiction du risque, et *a fortiori* lorsqu'ils sont dotés d'une IA, en raison de leur opacité et de leur aura d'objectivité, conduisent à un « *camouflage technologique* » en inscrivant des caractéristiques qui mériteraient d'être considérées comme des facteurs atténuants de la peine à titre de facteurs de risque susceptibles d'aggraver la peine, détournant ainsi l'exercice propre au prononcé de la peine qui est de déterminer une peine *juste* et proportionnelle[326]. En effet, l'introduction de ces outils à l'étape de la peine créerait une situation paradoxale : « A young, poor, or uneducated defendant might be at a higher risk for recidivism, but those same circumstances might also diminish his culpability and justify a more lenient sentence, rather than a harsher one. »[327] Nous observerons le même paradoxe à l'égard des accusés atteints de troubles mentaux : s'ils représentent un défi quant à la gestion de leurs risques dans la communauté, ils doivent recevoir une peine moindre en raison de leur responsabilité morale amoindrie[328]. Même si ces caractéristiques pouvaient « efficacement » établir le risque de récidive, il n'en demeure pas moins qu'ils ne sont pas en mesure d'indiquer la peine *juste* à rendre dans un cas donné. Il s'agit d'une autre illustration du clivage fonctionnel qui existe entre une institution pénale qui poursuit essentiellement la décision *juste* à rendre, qui est de *l'ordre du symbolique*, et des outils d'IA qui poursuivent la gestion effective des risques en communauté et reposent sur une *réalité statistique et quantifiable*.

---

lack of relevance to other important sentencing aims like retribution (which is a backward-looking assessment of an individual's blameworthiness) (…) In order to ensure that these limitations are being followed, the court mandated that a judge must explain at sentencing "the factors in addition to a COMPAS risk assessment that independently support the sentence imposed.""

[321] *Id*., p. 22.
[322] *Id*., p. 21.
[323] *Id*.
[324] *Id*., p. 25.
[325] A. CHRISTIN, "Predictive Algorithms and Criminal Sentencing", préc., note 123, p. 281.
[326] Danielle KEHL, Priscilla GUO et Samuel KESSLER, préc., note 212, p. 25
[327] *Id*.
[328] *R*. v. *Tremblay*, 2006 ABCA 252; *R*. v *Batisse*, 2009 ONCA 114; *R*. v. *Belcourt*, 2010 ABCA 319; *R*. v. *Edmunds*, 2012 NLCA 26; *R*. v. *Peters*, 2000 NFCA 55; *R*. v. *Ayorech*, 2012 ABCA 82; *R*. v. *Dedeckere*, 2017 ONCA 799; *R*. v. *Adamo*, 2013 MBQB 225

3.3.3. Dans le Service correctionnel et à l'étape de la libération conditionnelle

Dans le même esprit, le SCC doit prendre ses décisions à la lumière de son « but » mentionné à l'article 3 de la LSCMLC. S'il doit assurer *en priorité* la « protection de la société » (art. 3.1 LSCMLC), il doit également « contribuer au maintien d'une société juste » par des mesures de garde et de surveillance qui sont sécuritaires, mais aussi « humaines » (art. 3 LSCMLC). Le système correctionnel vise la protection du public et à assurer la sécurité dans la société, certes, mais d'une manière particulière, car il a également pour but d' « aider » les délinquants, « au moyen de programmes appropriés dans les pénitenciers ou dans la collectivité, à leur réadaptation et à leur réinsertion sociale à titre de citoyens respectueux des lois ». Par conséquent, les références au « maintien d'une société juste », aux mesures qui doivent être « humaines » et au moyen particulier de la *réhabilitation* pour assurer la paix et la sécurité dans la société viennent tempérer l'idée qu'une décision pourrait être simplement fondée sur la prédiction du risque de récidive par un outil d'IA et sur ses suggestions pour le contenir. Le système correctionnel vise, *in fine*, la protection de la société *par la réhabilitation éventuelle du délinquant*. À la lumière de cette fonction particulière, l'exercice d'attribution d'une cote de sécurité aux détenus nous apparait comme allant au-delà de la simple prédiction objective des risques. En effet, elle commande de poursuivre un idéal particulier de justice axé sur notre *croyance* collective envers la réhabilitation qui s'incarne à travers les différentes références au « juste » et à l'« humanité » des mesures - des exigences complexes qui pourraient échapper au simple calcul mathématique de la *machine*.

De plus, la qualification même des « catégories de risque », donc de ce qui « constitue une grande menace pour la sécurité du public » et de ce qui « exige un degré élevé de surveillance et de contrôle » (art. 30 LSCMLC et l'art. 18 du règlement d'application[329]), relève davantage de notre imaginaire collectif du risque, de notre tolérance aux risques et de notre manière de concilier le risque avec l'objectif de réhabilitation que d'un simple calcul statistique :

> « It is important to understand, however, that the determination of what constitutes a low, medium or high score is an *explicit policy choice*, *not a statistical or technical outcome*. An algorithm will produce *a numerical prediction*. Human decision-makers decide *whether* or *how* to label a statistical prediction as a low, medium or high risk. Nor does scoring always correspond to our common sense understanding of risk. As a result, risk scoring can be extremely misleading or prejudicial unless users are literate about what the score really means and how it was determined. »[330] (nos italiques)

En ce sens, la prédiction du risque ne peut éviter de se prêter au débat contradictoire, car le risque dépend de considérations sociales, d'un certain sens de la *justice* et doit faire l'objet d'une délibération particulière. Pour illustrer le genre de considérations sociales susceptibles d'intégrer l'exercice de prédiction du risque, citons l'article 17 du règlement d'application qui prévoit que, pour attribuer une cote de sécurité d'un individu, le SCC doit notamment prendre en compte « toute maladie physique ou mentale ou tout trouble mental dont il souffre ». Selon notre conception de la fonction du SCC, cet indice pourrait à la fois indiquer un certain niveau de risque, mais pourrait aussi indiquer que le délinquant aura besoin d'un cadre de surveillance particulier, peut-être même moins restrictif, pour permettre un traitement thérapeutique et sa réhabilitation. L'outil d'IA n'est pas en mesure de répondre *à quel moment* il est nécessaire de prioriser une aile de surveillance maximale pour assurer la sécurité au détriment d'un accès à des soins adaptés et à un environnement carcéral propice à la guérison. De plus, le SCC est guidé par d'autres considérations que le

[329] *Règlement sur le système correctionnel et la mise en liberté sous condition*, DORS/92-620 (« règlement d'application »)
[330] CDO1, p. 25.

risque. Il est ainsi guidé par un « principe de respect » des « différences ethniques, culturelles, religieuses et linguistiques, ainsi qu'entre les sexes, l'orientation sexuelle, l'identité et l'expression de genre, et doit tenir compte des besoins propres aux femmes, aux Autochtones, aux minorités visibles, aux personnes nécessitant des soins de santé mentale et à d'autres groupes » (art. 4 LSCMLC). Depuis 2019, toute décision du SCC doit également prendre en compte, lorsqu'ils « pourraient abaisser le niveau de risque », « les facteurs systémiques et historiques touchant les peuples autochtones du Canada » (art. 79.1(1)(2) LSCMLC). En raison des craintes liées à la partialité des outils d'IA, cet exercice particulier pourrait difficilement leur être confié.

La CLC est, elle aussi, guidée par des principes du droit qui se rapportent à la notion du *juste* et qui sont susceptibles d'échapper aux outils d'IA utilisés pour prédire le risque de récidive. Ici aussi, même si « la protection de la société est le critère prépondérant » depuis 2012, ce dernier doit se concilier avec d'autres considérations propres au *juste,* puisque « la mise en liberté sous condition vise à contribuer au maintien d'une société juste, paisible et sûre *en favorisant*, par la prise de décisions appropriées quant au moment et aux conditions de leur mise en liberté, la réadaptation et la réinsertion sociale des délinquants en tant que citoyens respectueux des lois. » (art. 100 et 100.1 LSCMLC) La CLC doit réaliser un exercice complexe d'équilibration et de conciliation entre certains principes du *juste* et le *risque*, ce qui comprend plus largement notre tolérance collective au risque et notre conception sociale du risque, et ce, tout en conservant en tête le but ultime qui est la réalisation de la protection de la société *par la réinsertion sociale*. Même si l'outil d'IA prédit « une récidive avant l'expiration légale de la peine », les commissaires doivent ensuite *qualifier* ce risque : il ne doit pas être « inacceptable pour la société » et la libération devra « contribuer » à la « protection » de la société « *en favorisant* sa réinsertion sociale » (art. 102 LSCMLC)[331]. La conception sociale du risque et notre tolérance face au risque sont susceptibles de fluctuer avec le temps[332]; l'outil d'IA pourrait ne pas évoluer au même rythme que notre imaginaire collectif. De plus, les facteurs pris en compte par l'outil d'IA peuvent ne pas correspondre aux facteurs pertinents qu'aurait sélectionnés le décideur pour évaluer le risque de récidive d'un détenu en particulier. Les choix du concepteur de l'outil se substitueraient alors au décideur. Si le décideur peut en théorie user de discrétion, se fier à son expérience et se référer à une énorme variété de facteurs[333] pour individualiser au mieux la décision, le recours généralisé aux outils actuariels - et a *fortiori lorsqu'ils sont dotés d'une IA* - pourrait amener à une déresponsabilisation des décideurs, une désindividualisation, une simplification et une automatisation de la décision[334]. Selon le *Manuel des politiques décisionnelles à l'intention des commissaires* (2021), les outils actuariels - ce qui pourrait comprendre ceux dotés d'une IA - doivent être utilisés uniquement pour les fins pour lesquelles ils ont été conçus et uniquement sur des groupes populationnels sur lesquels ils ont été validés[335]. Si la prédiction actuarielle diffère du jugement clinique de l'agent de libération conditionnelle, les commissaires

[331] La commission peut également refuser la liberté d'office prévu par la loi et maintenir un détenu en incarcération jusqu'à la fin légale de sa peine si elle a « des motifs raisonnables de croire qu'un délinquant commettra, s'il est mis en liberté avant l'expiration légale de sa peine, soit une infraction causant la mort ou un dommage grave à une autre personne, soit une infraction d'ordre sexuel à l'égard d'un enfant, soit une infraction grave en matière de drogue » (art. 129(3) LSCMLC) Là encore, il y a lieu de se demander si la seule prédiction de l'outil d'IA peut équivaloir à des motifs raisonnables de croire.

[332] Comme en témoigne l'historique au Canada dressé par Fernades PRATES, « La libération conditionnelle », Chapitre 5, p. 79-95 dans Jimenez E. et Vacheret M. (dir.) *La pénologie. Réflexions juridiques et criminologiques autour de la peine*, Montréal, Presses de l'Université de Montréal, 2013.

[333] COMMISSION DES LIBÉRATIONS CONDITIONNELLES DU CANADA, *Manuel des politiques décisionnelles à l'intention des commissaires - Deuxième édition no. 20, 2021-10-20,* Pol. 2.1, p. 1 et ss.

[334] Fernades PRATES, « La libération conditionnelle », préc., note 333, p. 92-94.

[335] COMMISSION DES LIBÉRATIONS CONDITIONNELLES DU CANADA, *Manuel des politiques décisionnelles à l'intention des commissaires - Deuxième édition no. 20, 2021-10-20,* Pol. 2.1, aux articles 6 à la p.1.

doivent prendre en compte les conclusions discordantes[336]. Compte tenu de l'aura de scientificité qui entoure l'outil actuariel - et *a fortiori* l'IA - nous craignons que l'une finisse par éclipser l'autre.

Finalement, les détenus ainsi que les potentiels candidats à la libération conditionnelle bénéficient également de *certaines* garanties d'équité procédurale qui comprennent, malgré l'inapplicabilité des garanties criminelles de l'arrêt *Stinchcombe* dans le contexte administratif[337], le droit à une décision motivée, écrite et détaillée et à la communication sommaire des renseignements soutenant la décision pour pouvoir y répondre (art. 7 de la *Charte*, art. 4 et 27 LSCMLC en matière correctionnelle et art. 101b), e) et 144 LSCMLC en matière de libération conditionnelle). Le respect de ces exigences peut être plus difficile si la décision est confiée à un outil d'IA qui fonctionnerait par apprentissage-machine. Là encore, l'outil ne saurait rendre la décision *juste* recherchée en raison de son fonctionnement opaque. Dans *May* c. *Établissement Ferndale*, la Cour suprême indiquait que la SCC avait privé illégalement des détenus de leur liberté en ne leur communiquant pas la manière dont le logiciel informatisé non-algorithmique de réévaluation de la cote de sécurité, *Security Reclassification Scale*, avait pondéré certains facteurs[338].

3.3.4. À l'étape de l'enquête sur la remise en liberté sous caution

L'enquête sur la remise en liberté semble tout à fait désignée, à première vue, pour accueillir les technologies d'IA visant la prédiction de la récidive puisque cette enquête requiert spécifiquement un exercice de prédiction[339]. En effet, la détention préventive peut être justifiée lorsque le procureur démontre qu'elle serait « nécessaire pour la protection ou la sécurité du public », « eu égard aux circonstances, *y compris toute probabilité marquée* que le prévenu, s'il est mis en liberté, commettra une infraction criminelle ou nuira à l'administration de la justice » (art. 515(10) *C.cr.*). À première vue, il semblerait alors qu'une évaluation des risques de récidive ou une prédiction de la récidive réalisée par un outil d'IA pourraient *contribuer*, en prenant en compte d'autres indices, à justifier la détention préventive. Le juge peut en effet considérer toute preuve « plausible ou digne de foi » (art. 518(1)e); reste à voir si les outils actuariels fonctionnant par IA respecteront cette exigence. De plus, même si la prédiction par un outil d'IA est jugée pertinente, elle pourrait être exclue par le juge en vertu de la *common law* pour des motifs de principe, qui se rapportent à la notion du *juste* en droit criminel[340]. En somme, au-delà de son admissibilité en preuve, il nous semble que les dispositions concernant la remise en liberté sous caution font appel à des considérations sociales plus larges que la simple prédiction efficace du risque; des considérations qui nécessitent une réflexion sur ce qui est *juste*, et qui doivent parfois prendre le dessus sur la simple analyse statistique. La prédiction par un outil actuariel ne pourrait remplacer le juge humain, ni ne devrait occuper une trop grande place dans la décision.

La *Charte* a consacré un principe de justice à titre de garantie en matière criminelle lors de la mise en liberté sous caution : « Tout inculpé a le droit de ne pas être privé sans *juste* cause d'une mise en liberté assortie

[336] *Id*., Pol. 2.1, aux articles 7 à la p.2.
[337] *R*. c. *Stinchcombe*, [1991] 3 RCS 326
[338] *May* c. *Établissement Ferndale*, 2005 CSC 82, Par. 117. V. également *Établissement de Mission* c. *Khela*, 2014 CSC 24
[339] Bernard E. HARCOURT, "Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age ", (2005) *University of Chicago Public Law & Legal Theory*, *Working Paper No. 94*, p. 35-37.
[340] Kate ROBERTSON et Jill R. PRESSER, "Algorithmic Technology and Criminal Law in Canada", p.82, préc., note 229, citant *R*. c. *Corbett*, [1988] 1 RCS 670, par. 99 : "Tout élément de preuve pertinent est admissible, sous réserve du pouvoir discrétionnaire d'exclure tout ce qui risque de causer un préjudice indu, d'induire en erreur ou d'embrouiller le juge des faits, de prolonger démesurément les procédures, ou ce qui devrait par ailleurs être exclu pour des motifs clairs de droit ou de principe. »

d'un cautionnement *raisonnable* »[341] (art. 11e)). Ce critère de la « juste cause » est un « élément essentiel d'un système de justice pénale éclairé » qui « consacre l'effet de la présomption d'innocence à l'étape préalable au procès criminel » et protège le droit à la liberté de l'accusé[342]. Reste à voir, et c'est précisément la question ici, à quel point la prédiction de la récidive suggérée par un outil d'IA peut servir de « juste cause » à la privation de liberté et si ses suggestions peuvent constituer un « cautionnement raisonnable ». Comme nous le savons, les notions de *juste* et de *raisonnable* font nécessairement référence à des considérations d'ordre social plus larges que la simple prédiction statistique du risque.

Dans le *Code criminel*, il existe tout d'abord une présomption de remise en liberté du prévenu (art. 515(1) *C.cr.*), et ce, « aux conditions les moins sévères possibles ». Il s'agit du *principe de retenue* qui est désormais codifié à l'article 493.1 *C. cr.* : « la détention avant le procès est l'exception et non la règle. »[343] Dans les situations où l'outil viserait à *remplacer* le décideur humain, nous pouvons nous demander comment le concepteur arrivera à encoder ce principe de retenue. Si l'outil de prédiction ne sert qu'à assister le juge, nous pouvons craindre la perte de force de ce principe face à l'aura de scientificité qui entoure la prédiction de l'outil d'IA. Quelle force le juge pourra-t-il donner au principe de retenue face à un algorithme qui prédit un risque moyen ou élevé de récidive ? À partir de quel *pourcentage* faut-il prioriser la sécurité du public au détriment de la présomption d'innocence ? Ensuite, le décideur doit désormais porter « une attention particulière » au fait de la vulnérabilité du prévenu, notamment lorsqu'il fait partie d'un groupe sur-représenté dans le système de justice pénale (493.2 *C.cr.*), comme les autochtones ou les personnes atteintes de maladies mentales. Le décideur doit également appliquer l'analyse socio-culturelle de l'arrêt *Gladue* qui porte sur les facteurs de discrimination systémique à l'encontre des prévenus autochtones[344]. Nous pouvons nous questionner sur la manière dont cette réalité sera codifiée à l'intérieur de l'algorithme et sur la capacité de l'algorithme à apprécier les facteurs fluides, fluctuants et évolutifs qui entourent cet exercice de correction des inégalités systémiques.

On retrouve cette idée que l'égalité est un principe inhéremment variable et fluctuant dans la pensée du philosophe canadien Charles Taylor : « Exactly what is meant by equality will vary »[345]. Notre conception de l'égalité est supportée, plus profondément, par notre imaginaire social et par la façon, à une époque donnée, dont on conçoit l'ordre moral qui nous entoure. Notre conception d'un ordre moral *égalitaire* peut se déplacer, selon Taylor, en fonction de trois axes : l'ordre moral moderne aurait varié dans son *étendue* (« more people live by it ; it has become dominant »), dans son degré d'*intensité* (« the demands it makes are heavier and more ramified ») et dans le degré de *concrétisation effective attendu* par le public (de l'herméneutique au prescriptif) : « The presumption of equality, implicit in the stating point of the state of Nature, where people stand outside all relations of superiority and inferiority, has been applied *in more and more* contexts, ending with the multiple equal treatment or nondiscrimination provisions, which are an integral part of most entrenched charters. »[346] Le désavantage avec l'outil d'IA est que celui-ci vient « fixer » une conception de l'égalité pour les décisions futures en fonction des décisions passées ou en fonction d'un ordonnancement de paramètres qui a été fixé dans le passé. Même si l'outil est programmé pour apprendre

[341] Voir également l'al. 2f) de la *Déclaration canadienne des droits,* S.C. 1960, ch. 44.
[342] *R.* c. *Antic*, 2017 CSC 27, par. 1. La présomption d'innocence est protégée par l'art. 11d) et le droit à la liberté est protégé par l'article 7 de la *Charte*.
[343] *R.* c. *Myers*, 2019 CSC 18, par. 1; R. c. St-Cloud, 2015 CSC 27 par. 70.
[344] Il a été jugé dans *R.* c. *Louie*, 2019 BCCA 257 que les facteurs de l'arrêt *R.* c. *Gladue,* [1999] 1 RCS 688, sont pertinents lors de l'enquête sur la remise en liberté.
[345] C. TAYLOR, préc., note 255, p.22.
[346] *Id*., p. 5. Voir également, pp. 4-9 et 16-17.

de lui-même (*machine-learning*), rien ne garantit qu'il évoluera en synchronicité avec notre imaginaire social[347]. En somme, la simple probabilité statistique du risque de récidive ne nous renseigne pas non plus sur le « juste équilibre » qui doit exister entre les différents impératifs sociaux qui soutiennent la détermination de la caution. Comment ordonnancer la présomption d'innocence, la protection de la société et la correction des inégalités systémiques? De cette manière, une trop grande importance accordée à la prédiction du risque de récidive pourrait possiblement venir éroder le principe de retenue, et plus important encore, l'effet de la présomption d'innocence. Cet exercice d'équilibration se retrouverait grandement influencé par les choix des ingénieurs lors de la construction du code de l'outil d'IA; le poids à accorder à chacun des impératifs devrait revenir au décideur, présent au procès, plutôt qu'à un ingénieur qui n'a pas assisté à la présentation des faits et à la procédure contradictoire.

De plus, nous pourrions argumenter que l'enquête sur la remise en liberté ne porte pas *précisément* et *uniquement* sur la prédiction de la récidive. La simple « probabilité statistique » de la récidive ne clôt pas à elle seule le débat. L'enquête semble faire appel à des considérations sociales plus larges : « The risk cannot be based on *conjecture* or speculation, or be a mere possibility ». D'un autre côté on reconnait que « it is *impossible* to make exact predictions about recidivism and future dangerousness. »[348] La prédiction exacte et incontestable sur la récidive future d'un criminel étant « impossible », c'est bel et bien *autre chose* qu'on recherche à cette étape; quelque chose de plus modeste, et qui fait appel à un *certain* imaginaire collectif sur la « dangerosité ». On évalue le « danger », y compris celui provenant du risque statistique, une fois réinscrit à l'intérieur de la dogmatique juridique et de l'imaginaire collectif qui le sous-tend.

Plusieurs considérations dans la « prédiction du risque de récidive » nécessitent donc une délibération indépendante, un débat contradictoire et une marge de manœuvre propre à la *juste* équilibration des différents impératifs sociaux et des garanties en matière criminelle. Les motifs du juge doivent se référer à ce qui est « nécessaire » afin de protéger le public, ce qui nécessite une certaine réflexion axiologique. Le juge bénéficie également d'une marge discrétionnaire dans l'évaluation de ce qui constitue une « probabilité marquée » (« *substantial likelihood* »). Par conséquent, le juge doit résister à un outil d'IA qui aurait tendance à lui suggérer une mise en détention pour le simple motif que celle-ci serait « pratique ou avantageuse » en raison d'un calcul statistique qui indiquerait que la récidive est « probante »[349]. De plus, cette probabilité marquée ne dit pas tout puisque le juge doit, ensuite, être en mesure de qualifier celle-ci comme étant « *susceptible de compromettre* la sécurité du public » et il doit, finalement, juger que l'incarcération avant procès serait « nécessaire » pour assurer la sécurité du public[350]. Le juge doit ensuite fonder sa décision sur les « circonstances » particulières au cas présenté devant lui, et non simplement sur les généralisations statistiques ou des liens corrélatifs inusités (« *hidden patterns* ») proposés par l'outil d'IA. Pour réaliser cet exercice, le juge peut prendre en considération une panoplie de facteurs pour évaluer la dangerosité[351]; les facteurs qui sont considérés par le juge pour évaluer la dangerosité peuvent donc aller

[347] John L.M. McDaniel et Ken G. Pease, préc., note 275, p.22 : "The law evolves over time and, when laws change, practitioners are expected to work through principled analysis rather than simply relying upon past cases (Law Society, 2019). Since most predictive and machine learning systems are trained on past data and are incapable of human-level principled analysis in a comparable fashion, reliance on them can lead to stagnation and conservative outcomes that hold the evolution of justice "anchored in the past" (Ibid: 23)"

[348] *R.* v. *Le*, 2006 MBCA 68, par. 30. *R.* c. *Morales*, (1992) 3 RCS 711 : « l'impossibilité de faire des prédictions exactes n'exclut pas un système de mise en liberté sous caution qui vise à priver de liberté sous caution ceux qui *risquent d'être dangereux*. »

[349] *R.* c. *Morales*, (1992) 3 RCS 711, p. 737.

[350] *Id*.

[351] *R.* c. *Rondeau*, (1996) RJQ 1155 (CA)

au-delà de ceux qui sont choisis par le concepteur de l'outil. Il y a donc un risque pour le juge à se fier uniquement à l'outil de prédiction, car celui-ci est susceptible de limiter considérablement l'examen qu'il aurait entrepris autrement. Dans ces situations, la défense doit pouvoir remettre en question la conception de la « dangerosité » proposée par l'informaticien et les facteurs qui ont été pris en compte. Ce n'est donc pas du simple risque statistique dont il est question, mais bien de la « dangerosité » de l'accusé, qui, elle, fait appel à des considérations d'ordre social et moral et qui méritent de faire l'objet d'un débat plus approfondi.

La Cour suprême a également jugé à deux reprises que le refus de la remise en liberté ne devait pas se fonder sur des « fins extérieures au bon fonctionnement du système de mise en liberté sous caution », comme l'émotivité ou la discrétion non-structurée. Nous craignons que la « réaction émotive » d'un juge, du procureur ou de l'expert clinique face à la prédiction d'un outil d'IA (ces professionnels pouvant être « mal informés » quant au fonctionnement des outils d'IA) ne finisse par contribuer à refuser une remise en liberté[352]. Nous craignons également que le recours à un outil d'IA pour prédire la récidive ne constitue finalement un moyen trop « arbitraire », conférant un trop grand pouvoir discrétionnaire[353] à celui qui le conçoit ou l'utilise, ce qui constituerait également une « fin extérieure » au bon fonctionnement du système de mise en liberté selon la Cour suprême.

Malgré l'apparente similarité des outils actuariels de prédiction du risque et de la tâche qui incombe au juge lors de l'enquête sur la remise en liberté, ces outils court-circuiteraient, selon le professeur Bernard Harcourt, la volonté initiale qui sous-tend notre système de droit criminel. En soi, le fait de fonder une décision de justice, même la détention avant procès, à partir des caractéristiques générales et immuables d'un individu et de leur corrélation avec la propension à la récidive serait problématique, puisque la décision de priver ou non la personne de sa liberté se retrouverait alors fondée sur des motifs *exogènes* aux prohibitions écrites dans le *Code criminel*. En effet, ces outils actuariels favorisent indirectement la criminalisation et la privation de liberté à partir de traits liées à la personne même de l'accusé, de son mode de vie et de ses comportements, ce que le Parlement n'a pas voulu criminaliser directement :

> "First, even when we are using prediction with regard to a criminal justice outcome that calls for prediction, we are likely to stigmatize other categories. (…) Profiling even where it seems most necessary may have the effect of marginalizing anyone who deviates from the norm and thereby may impose pressure on them to conform. Second, even when we use an innocuous trait—a category that does not bother us as much as race, gender or class— we are still derivatively, rather than directly, creating stigma. (…) If the troopers do indeed target speeders, this will likely produce a ratchet along the lines of speeders and non-speeders. It will produce a disproportionate correctional population that will likely communicate that "speeders" are, for instance, affiliated with the drug trade. (…) We may indeed want to criminalize speeding. But if so, it should be a decision about speeding, and not the product of a desire to criminalize drug trafficking. Not because speeding predicts drug trafficking. We need to reach the decision *independently*. We should not allow our actuarial methods to reshape or distort our law enforcement decisions"[354]

Pour Bernard Harcourt, la décision de justice doit se faire en fonction d'un support plus stable et qui serait « indépendant » des caractéristiques ou du mode de vie de la personne; par exemple, celui-ci et d'autres auteurs proposent que la décision de la caution soit prise sur la base de la gravité que l'on impute à certains

[352] *R.* c. *Hall*, 2002 CSC 64, par. 108.
[353] *R.* c. *Morales*, [1992] 3 RCS 711
[354] Bernard E. HARCOURT, "Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age ", (2005) *University of Chicago Public Law & Legal Theory*, *Working Paper No. 94*, p. 37.

crimes[355]. La décision de remise en liberté sous caution doit donc se faire *indépendamment* des prédictions fondées sur des corrélations, ceux-ci étant susceptibles de proposer des liens inusités (ex. personnes enclines au vagabondage ou nomade = plus grand risque de récidive ou de fuir l'administration = à détenir avant procès)[356]. Ces corrélations *dérivées* et *inusitées* proposées par les outils d'IA sont extérieures à la fois à la réflexion du magistrat renseigné par le procès, mais elles sont aussi extérieures au *Code criminel* et à la volonté initiale du législateur. Les outils actuariels permettent une extension du champ pénal. Le haut niveau de traitement des outils d'IA permet de suggérer des liens qui sont exogènes à notre volonté collective, qui est de criminaliser et de poursuivre un individu en raison de ses *acte*s et non en raison de ce qu'il *est*.

3.3.5. Lors d'une audience sur un engagement de ne pas troubler l'ordre

Une personne peut déposer une dénonciation lorsqu'elle a des « motifs raisonnables de craindre » qu'une autre personne commette certaines infractions ou cause certains préjudices[357]. Le juge peut alors tenir une audience. Il devra faire y comparaitre les parties et, s'il est démontré par prépondérance des probabilités que la crainte est justifiée par des motifs raisonnables, il pourra ordonner à la personne de contracter un engagement de ne pas troubler l'ordre public. La prédiction d'un outil d'IA peut-elle *équivaloir* à une « crainte raisonnable » qu'un crime soit commis ou la *motiver en partie* au point de justifier la dénonciation de cette personne et de lui ordonner de contracter un engagement de ne pas troubler l'ordre public ?[358]

Le libellé de l'article 810 *C.cr*., qui prévoit les critères généraux pour émettre cette ordonnance préventive, fait bel et bien mention de l'existence d'une « crainte » raisonnable d'une « personne », ce qui fait nécessairement référence à une personne humaine et à un ressenti humain. Dans l'état actuel du droit, nous ne pourrions admettre que le dénonciateur puisse être *substitué* par la machine : « informants who derive their fear exclusively from a predictive technology become mere conduits for statistics, potentially undermining the subjectivity of their own beliefs. In other words, the informant's role becomes redundant and perhaps dispensable (…) »[359] Dans les faits, l'un et l'autre ne devraient pas être considérés comme étant interchangeables. Les capacités d'appréhension, d'analyse et de perception du réel diffèrent fondamentalement chez l'homme et la machine. Par conséquent, si l'outil d'IA pouvait fonder une dénonciation, il faudrait des garanties plus élevées, comme un seuil au-delà de la prépondérance des probabilités, afin d'assurer l'équité de cette procédure. De plus, nous pouvons difficilement parler d'une « crainte » provenant d'une « personne » lorsque la crainte réelle du dénonciateur humain émane, est *causée* ou est, en partie, *influencée* par la prédiction de l'outil d'IA. Dans un tel scénario, il pourrait être difficile de prouver où la crainte a véritablement pris racine : "how would informants articulate or explain the basis for their fear, especially in a court of law that is governed by the rules of evidence?" [360] Il faut savoir également que la prédiction de la machine bénéficie généralement d'une autorité dérivée de son aura de

[355] *Id*. Voir également M. SYLVESTRE, N. BLOMLEY, et C. BELLOT, préc., note 298, p. 221 pour d'autres raisons : "At the moment, legal actors tend to focus on the likelihood that the accused, if released, might be committing *any criminal offence* or interfere with the administration of justice, *regardless of the gravity or seriousness of the future offence*, or *of any kind of proportionality between the nature of the restrictions imposed and that of the interference*. The importance legal actors attach to *crime prevention regardless of its gravity* is particularly problematic, as courts often do not have adequate evidence to sustain any risk analysis or base their decisions on evidence that would not be sufficient to justify a conviction (Ashworth and Zedner, 2014: 70)."
[356] Bernard E. HARCOURT, "Against Prediction: Sentencing, Policing, and Punishing in an Actuarial Age ", (2005) *University of Chicago Public Law & Legal Theory*, *Working Paper No. 94*, p. 37.
[357] *R*. c. *Penunsi*, 2019 CSC 39
[358] M. PURCELL et M. ZAIA, préc., note 202, p. 517.
[359] *Id*., p. 535.
[360] *Id*., p. 536-537.

scientificité; les recherches suggèrent qu'en cas de désaccord entre l'expert humain et l'expert robotisé, l'humain tend à déléguer la décision à la machine[361]. En somme, l'outil d'IA ne pourrait pas *remplacer* le dénonciateur et pourrait difficilement *contribuer et inciter* à la dénonciation - autrement, nous ne serions plus en présence d'une « crainte raisonnable » provenant d'une « personne ».

D'ailleurs, dans l'arrêt *R.* c. *Penunsi* la Cour Suprême a dénoncé le fait d'entreprendre un examen prospectif du risque dès qu'un contrevenant sortait de prison en raison du crime qu'il a commis, *donc avant même d'avoir pu développer naturellement une crainte raisonnable en raison d'un évènement précis et nouveau*, puisqu'autrement cela reviendrait à une itération indéfinie de la peine jugée proportionnelle par le tribunal:

> « Entreprendre une procédure d'engagement de ne pas troubler l'ordre public en vertu de l'art. 810.2 dès qu'une personne est libérée de prison risque d'entraîner une privation de liberté qui s'ajouterait au fait d'avoir purgé une peine déjà considérée comme proportionnée. Si *aucune autre preuve* ne permet de conclure que la crainte se concrétisera (par exemple, le fait pour le défendeur d'avoir proféré des menaces ou adopté un comportement violent pendant sa détention), une crainte fondée uniquement sur l'infraction pour laquelle le défendeur purge sa peine ne sera pas suffisante. Il ne serait pas approprié d'ordonner un engagement de ne pas troubler l'ordre public au titre de l'art. 810.2 sur cette base. L'engagement ferait alors office d'ordonnance de probation *de facto* et non d'outil prospectif visant à favoriser la justice préventive. »[362]

Nous voyons une similitude dans cette manière de procéder, vivement critiquée par la Cour suprême, avec le fait de conduire *automatiquement* des évaluations du risque à l'aide d'un outil d'IA dans le but de justifier une dénonciation de l'article 810 avant même le *développement naturel d'une crainte raisonnable* chez une personne, en raison *d'un fait nouveau*. Le fait pour la police d'avoir la possibilité de procéder de manière systématique à une évaluation du risque à l'aide d'un outil actuariel d'IA auprès des populations « à risque », par exemple, les personnes ayant un lourd casier judiciaire, sans même qu'une personne ait développé subjectivement des craintes de subir un préjudice, serait donc, selon ce point de vue, tout aussi problématique. La crainte ne doit pas uniquement reposer sur des « motifs raisonnables », elle doit exister, elle doit naitre subjectivement chez un dénonciateur. Permettre à l'institution pénale d'intervenir pro-activement avant même la naissance d'une crainte, en automatisant la conduite d'un examen de prédiction du risque pour justifier une ordonnance, reviendrait à étendre de manière déraisonnable le champ de la justice préventive et outrepasserait l'objectif d'apaisement des craintes des personnes vulnérables et de l'engagement envers le public qui est visé par la justice préventive. À quoi bon s'engager envers le public et l'apaiser, si la crainte n'émane que de la machine?

Dans l'arrêt *Budreo*[363], la Cour d'appel de l'Ontario a expliqué que notre système de justice pénale pouvait s'avancer dans des affirmations quant au risque futur de récidive de l'accusé malgré l'impossibilité d'atteindre une véritable objectivité. Cela démontre que la possibilité d'émettre validement et légalement une telle ordonnance repose, en réalité, sur *quelque chose de plus profond* qu'un calcul mathématique. L'importance de la contribution des outils d'IA de prédiction serait donc à relativiser même s'ils devaient un jour arriver à un plus haut niveau d'efficacité. C'est que l'exercice de prédiction du risque permis par le droit doit avant tout faire appel à l'expérience personnelle du juge lors d'une audience et à d'autres considérations d'ordre social : « as Holmes has reminded us, the life of the law has not been logic : it has been *experience*. The criminal law must operate in a world governed by practical considerations rather than abstract logic and, as a matter of practicality, the most that can be established in a future context is a

[361] *Id.*, p. 535 à la note 141 citant Jason MILLAR & Ian KERR, "Delegation, relinquishment, and responsibility: The prospect of expert robots" in Ryan Calo, A Michael Froomkin & Ian Kerr (dir.), *Robot Law*, Edward Elgar Publishing, UK, 2016, p. 126.
[362] *R.* c. *Penunsi* 2019 CSC 39, par. 63.
[363] *R.* v. *Budreo*, 2000 CanLII 5628 (ON CA).

likelihood of certain events occurring. »[364] Les prédictions sont tolérées tant qu'elles résultent de différents facteurs de risque qui sont *juridiquement reconnus* et qui ont été, et peuvent être, *examinés* par un juge. Pour y arriver, cela nécessite, depuis 1954, d' entendre les prétentions des deux parties dans le cadre d'une audience et de laisser le défendeur contester la crainte alléguée[365]. La prédiction sur le risque en droit n'est pas simplement mathématique, elle se prête à un débat contradictoire plus large et à une décision qui relève essentiellement des représentations faites devant le juge. Encore une fois, nous avançons que le « risque », malgré la promesse d'une prédiction objective de la part de l'outil d'IA n'échappe pas à un certain imaginaire collectif qui, lui, est fondé sur des considérations sociales plus larges; ces considérations sociales sont « relatives », elles doivent pouvoir faire l'objet d'un débat contradictoire, et devront, finalement, faire l'objet d'une délibération impartiale et indépendante par un juge en fonction de son expérience et de sa subjectivité.

Cette prédiction sur le comportement futur d'un défendeur se révèle alors être, plus largement, un exercice dialectique, permis par une procédure équitable, fondé à la fois sur des critères « objectifs », certes, mais également sur l'appréhension subjective du dénonciateur[366]. Au final, ce qu'il faut déterminer c'est s'il existe, par prépondérance des probabilités, « a reasonably based *sense of apprehension* about a future event, or as Then J. put it, "(…) a *belief*, objectively established, that the individual will commit an offence" »[367]. Pourrait-on argumenter que l'existence même d'une crainte subjective est tout aussi importante que sa raisonnabilité afin d'assurer l'équilibre actuel, qui se situe au fondement de cette pratique préventive? Actuellement, l'équilibre trouvé entre liberté et sécurité est assuré, non pas seulement par l'exigence découlant du critère objectif (*raisonnably based, objectively established*), mais par son association avec une appréhension subjective face à un évènement jugé désagréable (*belief, sens of apprehension*). Nous reprenons donc ici le questionnement émis par Michael Purcell et Mathew Zaia : qu'en est-il lorsque cette appréhension du préjudice provient d'une machine ?[368] Perdons-nous alors, dans notre quête d'objectivité, le fameux *sens humain subjectif* que suppose la justice ? En justifiant l'intervention au nom d'un contrôle plus efficace du taux de criminalité, et non simplement dans le but d'apaiser les craintes d'une victime humaine potentielle, perdrons-nous les justificatifs d'ordre humanitaire qui permettait de justifier en toute légitimité le développement de ce cadre pénal *préventif* ? Quel impact aura l'hyper-perceptibilité de la machine d'IA sur cet « équilibre raisonnable » que nous avions trouvé entre le droit à la liberté du défendeur et l'intérêt de l'État de protéger les plus vulnérables ?[369]

En somme, l'analyse du risque ne peut être déléguée complètement à une machine, car l'appréciation du risque justifiant l'ordonnance se réfère nécessairement à l'expérience humaine du risque, à ses nombreuses considérations sociales, à une certaine subjectivité, éventuellement une intersubjectivité pour établir sa raisonnabilité (débat), qui, elle seule, rend le processus « juste » et « équitable ». Le *juste* réclame donc l'indétermination, c'est-à-dire que l'on refasse *l'exercice* du jugement et de la délibération à chaque fois et

[364] *R.* v. *Lyons*, 1987 CanLII 25 (SCC), [1987] 2 S.C.R. 309, pp. 364-65 repris dans *R.* v. *Budreo*, 2000 CanLII 5628 (ON CA).
[365] *R.* v. *Riad*, 2014 ONSC 3407; *R.* c. *Henry*, REJB 1997-01177, (1997) AQ n 2071 (CS); *R.* c. *Penunsi*, 2019 CSC 39, par. 19.
[366] *R.* c. *Henry*, REJB 1997-01177, (1997) AQ n 2071 (CS).
[367] *R.* v. *Budreo*, 2000 CanLII 5628 (ON CA), par. 51 citant le juge Then dans *R.* v. *Budreo*, 1996 CanLII 11800 (ON SC), p.381 : « I do not accept the appellant's argument. The word "fear" or "fears" should not be considered in isolation but together with the modifying words in s. 810.1(1) "on reasonable grounds". Fear alone connotes a state of belief or an apprehension that a future event, thought to be undesirable, may or will occur. But "on reasonable grounds" lends objectivity to the apprehension. In other words, the phrase "fears on reasonable grounds" in s. 810.1(1) connotes a reasonably based sense of apprehension about a future event, or as Then J. put it, it "equates to a *belief*, objectively established, that the individual will commit an offence" (at p. 381). »
[368] M. PURCELL et M. ZAIA, préc., note 202, p. 517.
[369] *R.* v. *Budreo*, 2000 CanLII 5628 (CA), par. 48.

que l'on puisse remettre en question notre conception sociale du risque. La décision *juste* provient d'un juge humain, de son sens intime de la justice, de son expertise et de son expérience, une fois enrichi du rituel contradictoire et de la réalité partagée lors de l'audience.

PARTIE III. DROIT DE LA PREUVE

Au Canada, les principaux outils utilisés aux fins de la *collecte* et de la *production* de preuve, et qui fonctionnent à l'aide d'une IA ou qui pourraient y recourir, sont les technologies de RF, présentées dans la Partie I, les outils de triage et de détection de matériel d'exploitation sexuelle d'enfants à partir d'un appareil ou sur le web, les outils d'extraction automatisée de données à partir d'un appareil et les outils de production de preuve d'ADN par génotypage probabiliste.

Nous avons choisi d'aborder principalement deux familles d'outils automatisés : des outils liés à la *collecte*, et plus largement au triage et à la détection de preuves incriminantes sur le web ou sur un appareil, ainsi que des outils liés à la *production* de preuves. Comme il n'existe pas d'outils d'IA qui ont été intégrés à la pratique courante des tribunaux pour *évaluer* la crédibilité d'un témoin ou la qualité de la preuve, ni même de référence ou de préoccupations notables exprimées dans la doctrine concernant l'arrivée éventuelle de tels outils au Canada, nous priorisons les thèmes de la *Collecte* et de la *Production* de la preuve. Nous aborderons donc **(1)** la *collecte* par des outils de détection, d'extraction et de triage de preuves incriminantes, précisément, les images d'abus sexuels, et **(2)** la *production* de preuve d'ADN par génotypage probabiliste. Ces outils fonctionnent à l'aide d'une IA ou sont automatisés et pourraient éventuellement recourir à une IA Nous préciserons, lorsque cela sera nécessaire, les normes et les principes particuliers qui sont susceptibles de guider ces pratiques liées à la constitution du dossier de preuve.

1. Collecte de preuves par des outils automatisés : technologies automatisées de détection et de triage d'images d'abus sexuel, visualisation artificielle et extraction automatisée de données

1.1. Pratiques actuelles

Technologies automatisées de détection et de signalement. En 2004, *Sécurité publique Canada*, qui regroupe tous les ministères liés à la sécurité nationale, a mis sur pied la *Stratégie nationale pour la protection des enfants contre l'exploitation sexuelle sur Internet.* L'application de la *Stratégie* est rendue possible grâce à un partenariat avec la GRC, le *Ministère de la Justice* et le *Centre canadien de protection de l'enfance* (CCPE), un organisme sans but lucratif désigné pour gérer la ligne nationale de signalement d'exploitation sexuelle des enfants (Cyberaide.ca). L'un des objectifs officiels de la *Stratégie* est de « collaborer avec l'industrie numérique pour trouver de nouvelles façons de lutter contre [l'exploitation sexuelle des enfants en ligne]. »[370] Grâce au financement soutenu et répété de cette *Stratégie* par le gouvernement fédéral[371], le CCPE est en mesure de gérer le *Projet Arachnid*[372]*,* qui concerne « un robot

[370] Site web de SÉCURITÉ PUBLIQUE CANADA, « L'exploitation sexuelle des enfants sur Internet », en ligne : https://www.securitepublique.gc.ca/cnt/cntrng-crm/chld-sxl-xplttn-ntrnt/index-fr.aspx#:~:text=las (mis à jour le 2022-03-17)

[371] *Id*., « Le budget 2017 avait prévu 6 millions de dollars sur cinq ans, et 1,3 million de dollars par année par la suite, pour Sécurité publique Canada, via la Stratégie contre la violence fondée sur le sexe (Femmes et Égalité des genres Canada), afin de sensibiliser le public, d'améliorer la coordination des politiques et la recherche, et *de soutenir le projet Arachnid du Centre canadien de protection de l'enfance, qui est un outil permettant de repérer et de retirer le matériel d'abus pédosexuels en ligne*. » (nos italiques) Le budget fédéral de 2019 prévoyait 2,09 millions de dollar canadien attribué à la Sécurité publique du Canada pour « [c]ollaborer davantage avec le milieu du numérique au développement de nouveaux outils pour lutter contre l'exploitation sexuelle des enfants en ligne » et le « Budget 2021 propose d'accorder 20,7 millions de dollars sur cinq ans, à compter de 2021-2022 » pour faciliter les enquêtes en ligne de la GRC.

[372] Site Web du CENTRE CANADIEN DE PROTECTION DE L'ENFANCE, « Projet Arachnid », en ligne : https://protegeonsnosenfants.ca/fr/programmes-et-initiatives/projet-arachnid/ ; Site web du Projet Arachnid, en ligne : https://projetarachnid.ca/fr/#sommaire

Web qui détecte et traite des dizaines de milliers d'images par seconde et envoi des avis de retrait de matériel d'abus sexuels aux fournisseurs de services web à l'échelle mondiale. »[373] Ce processus, qui est automatisé, se déploierait « des milliers de fois par jour. »[374] Pour y arriver, *Arachnid* a recours à la technologie *PhotoDNA* de *Microsoft* et se réfère à une banque de données d'empreintes numériques - associées à chacune des photos prohibées - qui a été obtenue auprès de la GRC et d'Interpol. L'algorithme de *PhotoDNA* permettrait d'établir des correspondances entre les empreintes numériques trouvées sur le web et celles de la banque de données, et ce, sans utiliser la RF ni la reconnaissance d'objets[375]. On affirme que la « majeure partie des empreintes numériques » dans leur banque de données « ont été prélevées sur des photos et des vidéos examinées par des analystes du CCPE, par des équipes d'analystes d'autres centrales de signalement vouées à la protection des enfants et par des forces policières canadiennes et internationales. »[376]Ainsi, elle n'utiliserait pas encore l'IA, mais nous pensons qu'elle pourrait éventuellement en bénéficier, notamment pour découvrir d'elle-même des nouveaux documents n'ayant pas été préalablement hachés. La technologie serait, toutefois, déjà en mesure d'établir d'elle-même des correspondances entre des images non-identiques, même si elles ont été modifiées, et ce, de « quelconque façon »[377].

La *Sûreté du Québec*[378] et d'autres services de police au Canada[379] utilisent un logiciel-espion sensiblement similaire, nommé *Child Protection System*[380]*,* et qui a été conçu par la *Child Rescue Coalition* (CRC), un organisme à but non lucratif américain. L'outil automatisé fonctionne également en effectuant des concordances avec les hachages d'images déjà collectées sans utiliser une IA. Ce logiciel géolocalise les adresses IP et les ordinateurs personnels des personnes qui ont téléchargé ou partagé une image d'abus sexuel; il vise les réseaux de fichiers pair-à-pair (P2P networks), c'est-à-dire des réseaux de partage de fichiers similaires aux sites de piratage[381]. Le système permettrait également de monitorer des chambres de clavardage qui sont utilisées pour échanger du matériel illégal[382]. Le logiciel semble reposer sur une logique de prévention du risque et ciblerait les individus qui, selon les concepteurs ou les utilisateurs de la technologie, représentent un *plus grand risque* envers les enfants. En effet, si le site donne peu d'informations sur le fonctionnement même de la technologie, on y affirme simplement que cette

---

[373] Site web de SÉCURITÉ PUBLIQUE CANADA, « L'exploitation sexuelle des enfants sur Internet », en ligne : https://www.securitepublique.gc.ca/cnt/cntrng-crm/chld-sxl-xplttn-ntrnt/index-fr.aspx#:~:text=las (dernière modification 2021-08-12)

[374] CENTRE CANADIEN DE PROTECTION DE L'ENFANCE, *Projet Arachnid : l'accessibilité des images d'abus pédosexuels sur internet*, p. 7, disponible en ligne : https://protectchildren.ca/fr/ressources-et-recherche/projet-arachnid-accessibilite-images-abus-pedosexuels/

[375] Site Web de Microsof, "PhotoDNA", en ligne : https://www.microsoft.com/en-us/photodna

[376] CENTRE CANADIEN DE PROTECTION DE L'ENFANCE, *Projet Arachnid : l'accessibilité des images d'abus pédosexuels sur internet*, préc., note 375, p.7.

[377] Site web du Projet Arachnid, en ligne : https://projetarachnid.ca/fr/#sommaire

[378] Gabrielle DUCHAINE et Caroline TOUZIN, « Traquer les pédophiles en direct », *La Presse*, 10 janvier 2021, en ligne : https://www.lapresse.ca/actualites/enquetes/2022-01-10/l-autre-epidemie/traquer-les-pedophiles-en-direct.php

[379] Site web de CHILD RESCUE COALITION, « Ottawa Man Faces Charges Relating To Child Sexual Abuse Material », 22 Novembre 2019, en ligne : https://childrescuecoalition.org/ottawa-man-facing-child-pornography-charges/

[380] Site Web de CHILD RESCUE COALITION, en ligne : https://childrescuecoalition.org/ ; Pour l'historique de CHILD RESCUE COALITION, "Technology Powers fight against Child Predators", 20 avril 2020, en ligne : Technology powers fight against child predators - Child Rescue Coalition

[381] Olivia SOLON, "Inside the surveillance software tracking child porn offenders across the globe », *NBC NEWS*, 17 juillet 2020, en ligne : https://www.nbcnews.com/tech/internet/inside-surveillance-software-tracking-child-porn-offenders-across-globe-n1234019 : "The tool has a growing database of more than a million hashed images and videos, which it uses to find computers that have downloaded them. The software is able to track IP addresses — which are shared by people connected to the same Wi-Fi network — as well as individual devices. The system can follow devices even if the owners move or use virtual private networks, or VPNs, to mask the IP addresses"

[382] *Id*.

« [l]eading-Edge Technology » fonctionne par « [a]nalytics targeting the offenders at *greatest risk* of presently abusing children. »[383] (notre italique) Le système serait également capable d'identifier des documents légaux, mais qui sont décrits comme étant « suspicieux » lorsque ceux-ci sont téléchargés parallèlement à d'autres images illégales, par exemple des guides visant à gagner la confiance des enfants ou des parents, des romans sur l'inceste ou des dessins animés pornographiques[384]. Les programmeurs de CRC utilisent la technologie pour cibler les personnes représentant un « risque plus important » pour les enfants; ils seraient en mesure de déterminer si la personne ciblée est en position d'autorité envers des enfants ou s'elle représente un risque plus important envers les enfants pour une tout autre raison[385]. La CRC cherchait récemment à développer des partenariats avec des compagnies pour étendre leur technologie de prévention du risque sur des plateformes en ligne plus traditionnelles, comme Facebook, ou des sites de garderie ou d'école. En offrant une liste d'adresses IPs problématiques à ces compagnies, celles-ci pourraient effectuer elles-mêmes une vérification en cas de doute selon NBC news: "Additional screening on a babysitting app could include checking an account for "abnormal" characteristics, such as logging in much more frequently than a typical user, or checking whether it is attached to a profile indicating that the person is willing to travel long distances for a job or is offering a rate that is well below the average."[386]

Si ces logiciels ne semblent pas fonctionner à l'aide d'une IA à proprement parler, des chercheurs ont étudié les potentiels bénéfices qui découleraient du fait d'associer une technologie d'IA à un logiciel de concordance, comme *Child Protection System*. Une fois le logiciel associé à une technologie d'IA, il serait alors en mesure de comprendre par lui-même qu'elle est en présence d'un document d'abus sexuel et de repérer automatiquement sur le net les documents d'abus sexuels « nouvellement créés », donc qui n'ont pas encore été recensés dans la banque de données ou qui n'ont pas été « identifiés » comme tels dans le passé[387]. La technologie de détection de la CRC a été dénoncée aux États-Unis en 2019 par *Human Rights Watch* dans une lettre ouverte adressée au *Département de Justice américain*[388]. L'organisme soulevait alors le manque d'évaluation indépendante de cette technologie et les possibles atteintes aux droits et libertés des citoyens américains. Le fonctionnement exact de cette technologie ne semble avoir été très bien détaillé publiquement et ne semble pas avoir fait l'objet d'une étude indépendante au Canada sur les droits et libertés possiblement touchés par cette technologie. Rappelons qu'en droit canadien, comme nous l'avons vu dans la Partie I, les internautes ont une attente raisonnable de protection en matière de vie privée à l'égard de leur adresse IP une fois que celle-ci est associée à leur nom, leur numéro de téléphone ou leur adresse; une

---

[383] Site Web de CHILD RESCUE COALITION, en ligne : https://childrescuecoalition.org/law-enforcement/

[384] Olivia SOLON, préc., note 382 ; Site Web de CHILD RESCUE COALITION, "Technology Powers fight against Child Predators", 20 avril 2020, en ligne : Technology powers fight against child predators - Child Rescue Coalition : "Especially high-risk file sharers can be spotlighted, such as those who not only have illegal images, but "grooming" manuals on how to gain the trust of parents and children. While the manuals are actually protected under the First Amendment, the combination with illicit videos and photos is a major red flag, Yoost says."

[385] Site Web de CHILD RESCUE COALITION, "Boca Raton-Based Child Rescue Coalition Works "To Find Those That Victimize Kids", 18 novembre 2015, en ligne : BOCA RATON-BASED CHILD RESCUE COALITION WORKS "TO FIND THOSE THAT VICTIMIZE KIDS" - Child Rescue Coalition : "Wiltse said his organization has no idea who is viewing the child porn. In other words, they don't have names of the people behind the keyboard. Wiltse added the only thing that's put in the agency's database is the information that can locate the computer. However, Child Rescue Coalition programmers do use technology to figure out if the computer belongs to someone with access to kids. "These are the individuals that we believe are in positions of trust or other areas where we believe they pose a greater risk than someone else to children," Wiltse said."

[386] Olivia SOLON, préc., note 382.

[387] Claudia PEERSMAN, Christian SCHULZE, Awais RASHID, Margaret BRENNAN et Carl FISCHER, « iCOP: Live forensics to reveal previously unknown criminal media on P2P networks », (2016) 18 *Digit. Investig.* 50-64.

[388] HUMAN RIGHTS WATCH, « Letter to US Department of Justice About Child Protection System Software », 3 avril 2019, disponible en ligne : https://www.hrw.org/news/2019/04/03/letter-us-department-justice-about-child-protection-system-software#

entreprise privée ne pourrait pas divulguer *volontairement* à la police ces renseignements personnels sans que cette dernière obtienne un mandat judiciaire[389].

Nous nous inquiétons plus généralement de la logique du risque qu'insufflent ces technologies à l'intérieur de la justice pénale - car elles facilitent et favorisent la criminalisation des comportements et la pénalisation des individus à partir de la perception de risques par les concepteurs ou des utilisateurs de la technologie. Cette logique managériale du risque offre alors une grande discrétion à l'État dans la criminalisation des comportements et la pénalisation des individus, et par le fait même, amène à une plus grande incertitude juridique. Compte tenu que certains de ces outils technologiques permettent d'être utilisés ou sont conçus de manière à classer et prioriser les cas détectés en fonction du risque perçu, et, sachant que, dans l'optique d'économiser les ressources, les policiers pourraient avoir tendance à prioriser les cas identifiés comme étant à haut risque[390], ces technologies auront inévitablement un effet normatif important sur notre conception de la fonction de la justice pénale et son orientation future, ainsi que sur la composition des nouvelles cohortes de criminels. Nous nous posons les questions suivantes : les personnes « à haut risque » sont-elles nécessairement des personnes avec une plus grande *culpabilité morale* ? Sont-elles nécessairement des personnes dont le châtiment renforcera l'autorité morale de la loi ? Ne peuvent-elles pas être également des personnes socialement vulnérables et déjà stigmatisées, sans qu'il soit nécessaire d'y rajouter le stigmate pénal ? Recourir à ces outils, de manière à prioriser les cas « à risque », c'est attribuer présomptueusement une fonction précise - et restreinte - à l'institution pénale; c'est prendre le pari qu'elle est la mieux placée pour gérer et contrôler le taux de criminalité, qu'il s'agit là de *sa* fonction propre et qu'elle est véritablement capable de *contrôler* les risques en société par l'entremise d'un *châtiment*. Plus particulièrement, nous nous inquiétons du fait que, si c'est la machine qui établit elle-même le classement de priorisation, il ne faut pas oublier que sa conception du risque pourrait ne pas être celle que nous partageons actuellement et pourrait ne pas évoluer en synchronicité avec cette conception.

Technologie de triage des images par IA. Une fois un appareil, dûment saisi, les policiers canadiens peuvent avoir recours à un logiciel d'IA afin d'effectuer l'identification de matériel prohibé à travers l'ensemble des données contenues dans le disque dur. Depuis 2011, la GRC aurait elle aussi utilisé le logiciel *PhotoDNA* afin d'avertir l'enquêteur de l'existence de matériel prohibé. Le logiciel permettrait d'effectuer des associations entre le code de hachage d'une photo qui apparait « en ligne ou sur un disque dur » avec le code de hachage qui existe déjà dans leur banque de données[391].

En 2017, la GRC en partenariat avec l'*Université du Manitoba*, *Two Hat Security Ltd.* et *Mitacs* a cherché à développer, en plus de *PhotoDNA*, un nouvel outil qui, lui, fonctionnerait plutôt par « visualisation artificielle » pour identifier les photos « qui présentent des *probabilités élevées* de constituer de l'exploitation sexuelle d'enfants ». Désormais, « [c]e qui exigerait des semaines pour un enquêteur ne prendra que quelques minutes ou heures pour un algorithme, explique Brad Leitch, chef du développement de produits chez Two Hat Security. L'algorithme peut éliminer les photos d'arbres, de beignes et de la tour

---

[389] *R.* c. *Spencer*, (2014) 2 R.C.S. 212

[390] Concernant la pratique policière consistant à créer des listes des délinquants en fonction de leur risque de commettre un acte criminel, cf. John L.M. MCDANIEL et Ken G. PEASE, « Introduction » préc., note 275, p. 10.

[391] Amelia THATCHER, « Une démarche visionnaire : La GRC cherche un logiciel reconnaissant les images d'exploitation d'enfants », *Site Web de la Gendarmerie Royale*, Vol. 79, Nº 3 — Nouvelle technologie, 4 juillet 2017, en ligne : Une démarche visionnaire | Gendarmerie royale du Canada (rcmp-grc.gc.ca) (dernière modification 2017-06-12)

Eiffel avec facilité et prioriser les photos d'exploitation sexuelle vraisemblables (…) »[392] L'organisme souhaitait alors appliquer la nouvelle technologie au « triage des cas » lors de la saisie d'un disque dur. En 2019, le *Centre national des crimes d'exploitation d'enfants* (CNCEE) de la GRC annonçait, dans une rubrique sur le bien-être des employés, avoir recours à l'intelligence artificielle « depuis des années »[393]. Il serait ici question d'un « logiciel d'apprentissage machine », qui est « programmé pour accomplir des tâches précises sans instructions codées », afin de départager, à l'étape de la collecte de la preuve, les images de nature légale des images d'abus d'enfant dans l'ensemble des données du dossier saisi. L'enquêteur est ensuite appelé à vérifier les images classées comme illégales. Le logiciel serait en mesure de reconnaitre des images que les enquêteurs du CNCEE ou d'autres organismes canadiens ou internationaux ont déjà classées comme des images d'abus d'enfants. Cela évite à l'enquêteur de devoir consulter à nouveau ces images[394] afin de préserver son bien-être,.

Technologie d'extraction automatisée. La WRP et la GPS auraient utilisé au cours des dernières années un appareil, nommé *GrayKey*, développé par l'entreprise américaine *Grayshif*, qui permet l'extraction automatique de *toutes* les données de téléphones intelligents verrouillés[395]. Ce n'est qu'en août 2021 que l'appareil aurait été mis à jour afin de permettre une extraction ciblée de données afin de permettre aux policiers de respecter les conditions strictes prévues dans leur mandat de fouille :

> « In August, we released the Category-based extraction feature for iOS devices initially because of customer feedback where limited warrants are used, or urgent circumstances where a device is only available to an investigator for a brief period of time. This can be useful where specific categories of data are extracted without having to pull the entire full file system. Category-based extractions will allow you to survey the device once connected to GrayKey and select categories of data for extraction that might be of importance to your specific investigation. »[396]

Cela laisse présager que les extractions effectuées avant cette mise à jour ne pouvaient pas être strictement délimitées. Néanmoins, selon la WRP et la GPS, les extractions automatisées auraient été conduites en conformité avec les mandats obtenus auprès d'un juge ou avec le consentement du prévenu[397]. Le *Commissaire à l'information et à la protection de la vie de l'Ontario* n'aurait toutefois pas été consulté par les services de police avant de recourir à cette *nouvelle* technologie[398]. Ainsi, la régulation de cette technologie semblait laissée à la discrétion du bureau interne des services de police. Cette technologie automatisée ne semble pas recourir à une IA pour l'instant, mais pourrait éventuellement en bénéficier.

---

[392] *Id.*

[393] Paul NORTHCOTT, « Face aux « Les criminels les plus odieux » : l'impératif du bien-être des employés » *Site Web de la Gendarmerie Royale*, Vol. 81, Nº 3 — Reportages, 2 juillet 2019, en ligne : Face aux « Les criminels les plus odieux » | Gendarmerie royale du Canada (rcmp-grc.gc.ca)

[394] *Id.*

[395] Chris Seto, « Waterloo Regional Police have a device that can crack locked phones », *The Record*, 3 janvier 2021, en ligne : https://www.therecord.com/news/waterloo-region/2021/01/03/waterloo-regional-police-have-device-that-can-crack-locked-phones.html; Paula DUHATSCHEK, « Waterloo regional police in talks with privacy commissioner about cellphone unlocking software », *CBC News*, 19 mars 2021, en ligne : https://www.cbc.ca/news/canada/kitchener-waterloo/waterloo-regional-police-privacy-cellphone-unlock-1.5950922 ; Graeme MCNAUGHTON, « Guelph police have used GrayKey device to unlock, extract data from 47 iPhones », *Guelph Mercury*, 9 décembre 2020, en ligne : https://www.guelphmercury.com/news-story/10285518-guelph-police-have-used-graykey-device-to-unlock-extract-data-from-47-iphones/ ; Graeme MCNAUGHTON, « Guelph police have tool to unlock iPhones and copy contents, with no policy on when or how to use it », *Guelph Mercury*, 23 novembre 2020, en ligne : https://www.guelphmercury.com/news-story/10272853-guelph-police-have-tool-to-unlock-iphones-and-copy-contents-with-no-policy-on-when-or-how-to-use-it/

[396] Site Web de la compagnie Grayshift, en ligne : https://www.grayshift.com/groundbreaking-insights-into-graykey/

[397] Graeme MCNAUGHTON, « Guelph police have used GrayKey device to unlock, extract data from 47 iPhones », préc., note 396.; Paula DUHATSCHEK, préc., note 396.

[398] Chris SETO, préc., note 396; Graeme MCNAUGHTON, « Guelph police have tool to unlock iPhones and copy contents, with no policy on when or how to use it », préc., note 396; Paula DUHATSCHEK, préc., note 396.

Actuellement, le vendeur affirme que la technologie ne se contente pas d'extraire des données, mais permet aussi de « [i]dentify patterns and obtain key insights through expedited data extraction »[399].

1.2. Cadre normatif

L'adaptation du cadre normatif actuel en matière de respect de la vie privée, de protection contre les fouilles, saisies et perquisitions abusives (art. 8 de la *Charte*) représente un défi face à la potentialité technique de ces nouveaux outils. Nous ferons ici ressortir les principales normes touchées par ces outils.

Normes déterritorialisées. Des normes internationales ou régionales peuvent s'appliquer aux compagnies canadiennes ou américaines qui produisent ces technologies. Elles peuvent influencer leurs pratiques lorsqu'elles font affaire avec les services policiers canadiens. Par exemple, la compagnie *Two Hat Security Ltd.* de la Colombie-Britannique, qui travaille en collaboration avec la GRC sur un projet de « visualisation artificielle », se reconnait ouvertement comme étant liée par le *Règlement général sur la protection des données* de l'Union européenne[400]. Ce règlement permet de sanctionner des compagnies hors Union européenne lorsqu'elles collectent des données sur des citoyens européens sans se conformer à ses dispositions, notamment en matière de transparence (notification en cas de violation de données, droit à l'accès, droit à l'oubli)[401]. Leur projet de « visualisation artificielle », comportant des activités dans l'espace déterritorialisé du cyberespace, ne pourrait donc se passer d'une certaine conformité aux dispositions européennes et d'une collaboration avec les autres autorités extraterritoriales.

Normativité entourant l'obtention d'un mandat de perquisition, de saisie ou de fouille. Nous suggérons que le type de technologie utilisé pour motiver la saisie d'un appareil ou pour extraire ses données devrait être pris en compte par le juge lorsqu'on lui demande une autorisation judiciaire. Actuellement, la *common law* encadre l'application par les tribunaux de l'article 8 de la *Charte* sur les saisies, fouilles et perquisitions abusives de plusieurs manières. ***Détection et signalement par un logiciel automatisé ou une IA*** - Lorsque la demande pour obtenir un mandat de saisie concerne principalement un logiciel automatisé, comme celui de la *Child Rescue Coalition* - *a fortiori* lorsqu'il fonctionne par IA -, le juge devrait s'assurer que le fonctionnement du logiciel permet véritablement de soutenir une « cause probable », aux États-Unis[402], ou des « motifs raisonnables de croire », au Canada[403], que du matériel illégal pourra être découvert par une telle entrave au droit à la vie privée. Des signalements effectués par des logiciels, comme celui de la CRC, ne sont pas suffisants à eux seuls pour justifier une arrestation aux États-Unis, mais, selon le président de la CRC, ils pourraient « contribuer » à établir la « cause probable » permettant d'obtenir un mandat de saisie[404]. Au Canada, pour respecter l'article 8 de la *Charte* contre les fouilles, saisies et perquisitions abusives, les policiers qui veulent extraire des données d'un ordinateur doivent obtenir un mandat auprès du juge et lui démontrer qu'ils possèdent « des motifs raisonnables de croire que les ordinateurs qu'ils pourraient

---

[399] Site Web de la compagnie Grayshift, en ligne : https://www.grayshift.com/graykey/
[400] Site Web de la compagnie Two Hat, en ligne : https://www.twohat.com/solutions/gdpr-coppa-data-protection/ : « GDPR applies to any company that collects data from EU citizens, regardless of their physical presence in the EU. The regulation increases accountabilities for both "data controllers" (companies that collect personal data) and "data processors" (companies like Two Hat that process personal data). »; *Règlement général sur la protection des données,* 2016/679 (EU)
[401] Site Web de la COMMISSION EUROPÉENNE, en ligne : https://ec.europa.eu/info/law/law-topic/data-protection_en
[402] IVe amendement de la *Constitution américaine*
[403] *R*. c. *Morelli*, 2010 CSC 8, par. 127-128.
[404]Olivia SOLON, préc., note 382.

découvrir contiendront les choses qu'ils recherchent. »[405] Ces motifs raisonnables doivent provenir d'informations « détaillées » fournies par un indicateur « fiable » - il s'agit alors d'évaluer la probabilité fondée principalement sur la « crédibilité » de la source de ces motifs[406]. Comme il s'agit d'une question de droit, il reviendra au juge de déterminer si un tel logiciel (dont le fonctionnement pourrait éventuellement être assuré par une IA) est à même de constituer des « motifs raisonnables de croire » que les policiers trouveront ce qu'ils recherchent[407]. ***Extraction de preuves par une IA*** - Le juge devrait également être mis au courant du type de technologie qui sera utilisé pour extraire les preuves afin de dresser clairement, dans le mandat, les limites appropriées à cette forme de collecte de preuve. Les limites prévues à même le mandat permettent de respecter le principe de **proportionnalité** qui sous-tend tout empiètement sur la liberté des citoyens (art. 1 *Charte*). Lorsque la fouille, qui pourrait prendre la forme d'une extraction *automatisée*, concerne des messages électroniques *archivés* ou *historiques* sur un téléphone intelligent, celle-ci serait assujettie à l'article 8 de la *Charte*. En effet, la Cour suprême a jugé que les messages-textos envoyés par un tiers, qui sont enregistrés dans le téléphone du suspect et sur lesquels ce tiers n'a aucun contrôle, sont protégés par la *Charte* - toute personne a une attente raisonnable de protection à l'égard de sa vie privée concernant les messages qu'elle a envoyés[408]. Pour y accéder, la police devra donc obtenir une ordonnance de communication prévue à l'article 487.014 *C.cr.*[409] À ce sujet, les rapporteurs du *Citizen Lab* ont émis certaines inquiétudes concernant les biais culturels qui peuvent survenir lors de l'extraction automatisée de messages *historiques* :

> « The use of automated tools is difficult when analyzing language that is highly context-dependent and culturally specific, which may not be adequately reflected in the data sets used to train the algorithms. For example, Professor Andrew Ferguson notes that keywords related to gangs and shootings ("hit", "run", "strike", "cap", and "park") can also relate to baseball. An algorithm that cannot distinguish between the two contexts may draw police attention to individuals, based on innocuous communications, and subject those individuals to privacy violations. »[410]

Normativité découlant des directives internes des services de police. Actuellement, la régulation spécifique des nouvelles technologies utilisées pour mener des enquêtes, pour trier des documents ou pour extraire automatiquement des données semble relever de la discrétion des services de police. Ces services peuvent constituer un bureau d'examen à l'interne qui est chargé d'effectuer une vérification des risques liés à l'utilisation d'une nouvelle technologie aux fins de collecte de preuve. Cette forme d'autorégulation normative peut être problématique, car le degré de rigueur dans les critères d'examen, la force dans l'élaboration et dans l'application des normes peut grandement varier d'un service de police à l'autre et être circonstancielle à l'attention portée par les médias ou par le *Commissaire à la vie privée*. Par exemple, la GPS aurait utilisé la technologie de *Graykey* sans même avoir préalablement établi de politiques ou de directives internes pour encadrer les risques d'intrusion dans la vie privée liés à cette nouvelle technologie. Ce n'est qu'après avoir fait l'objet d'une attention particulière par les médias locaux que la GPS a annoncé l'élaboration d'une politique interne concernant l'examen préalable au recours de nouvelles technologies d'extraction[411]. La WRP avait déjà mis en place un bureau d'examen à l'interne qui était chargé d'effectuer un examen lorsque les policiers souhaitaient utiliser une nouvelle technologie aux fins de la collecte de

[405] *R.* c. *Vu*, 2013 CSC 60, par. 48; *R.* c. *Morelli*, 2010 CSC 8; *Hunter* c. *Southam Inc.*, [1984] 2 R.C.S. 145.
[406] *R.* c. *Debot*, [1989] 2 RCS 1140; *Baron* c. *Canada* [1993] 1 RCS 416.
[407] *R. c. Shepherd*, [2009] 2 R.C.S. 527
[408] *R.* c. *Marakah*, (2017) 2 R.C.S. 608
[409] Benyekhlef et Déziel, p. 234.
[410] Citizen Lab, p. 88.
[411] Graeme MCNAUGHTON, « Guelph police have used GrayKey device to unlock, extract data from 47 iPhones », préc., note 396.

preuve. La WRP a présenté le 15 avril 2020, une directive interne pour encadrer l'introduction de nouvelles technologies dans les méthodes policières :

> "The order says before any new technology that has the potential to violate a person's privacy is implemented by an officer, that member's leader will need to be advised of the device and its proposed use. A preliminary business case is to be prepared in writing for senior leadership. If preliminary approval is given, the director overseeing the access to information unit is informed to provided guidance regarding privacy considerations. If the new tech is going to be used for investigative/intelligence purposes, the Crown attorney and/or the Public Prosecution Service of Canada shall be consulted. A privacy impact assessment (PIA) is to be prepared and the decision on whether or not to consult the Information Privacy Commission will be made by the chief. Only after the PIA is signed off by the chief can the technology be used."[412]

On ne connait pas le degré de rigueur des critères de ces examens internes. Comme nous l'avons vu, il semblerait que les services de police n'ont pas eu besoin de consulter systématiquement le *Commissaire à la vie privée* de leur province avant d'utiliser cette nouvelle technologie. Généralement, les politiques internes établies les policiers sont appelées, après un examen du *Commissaire à la vie privée*, à être renforcées. Par exemple, à la suite du *Rapport spécial sur l'utilisation de la technologie de reconnaissance faciale* concernant l'utilisation d'AI Clearview par la GRC, cette dernière avait été appelée par les *Commissaires à la vie privée* à modifier ses pratiques internes. En mars 2021, elle a donc créé le *National Technologies Onboarding Program* « to centralize and bring more transparency to the processes that govern how the RCMP identifies, evaluates, tracks and approves the use of new and emerging technologies and investigative tools that involve the collection and use of personal information. This program will establish standardized processes for the assessment of developed and/or procured technologies and services, and evaluate compliance of collection techniques with privacy legislation. »[413] La GRC devait mettre en place les statuts opérationnels de cet organisme à l'automne 2021 et devait produire des directives internes aux employés de la GRC pour encadrer leur recours aux nouvelles technologies à des fins d'enquête. Nous n'avons pas pu trouver ces informations; elles ne semblent pas avoir été partagées publiquement.

Normes fondées sur des principes émanant de la société civile. Plusieurs directives fondées sur des principes et qui émanent des acteurs de la société civile sont susceptibles d'encadrer les pratiques d'enquête et de collecte de preuves par les corps policiers. Ces directives sont à même de fonder de futures lois nationales en la matière. Notons, entre autres, les *Sedona Canada E-Discovery Principles* qui visent à encadrer la collecte de preuves électroniques[414] : par exemple, l'information faisant l'objet de la collecte doit pouvoir être « découverte », en ce sens qu'elle doit pouvoir être « détectable » (*discoverable*) (principe #1), et la collecte, dans toutes les étapes de la découverte de la preuve, doit être « proportionnelle » (principe #2) en tenant compte notamment des conséquences liées à la découverte de la preuve incriminante et à la nature pénale de la procédure en cours[415].

---

[412] Chris SETO, préc., note 396.

[413] Déclaration de la GENDARMERIE ROYALE DU CANADA, « Response to the Report by the Office of the Privacy Commissioner into the RCMP's use of Clearview AI », 10 juin 2021, en ligne : https://www.rcmp-grc.gc.ca/en/news/2021/response-the-report-the-office-the-privacy-commissioner-the-rcmps-use-clearview-ai : « We continue to develop NTOP and are working towards operational status by the fall of 2021. »

[414] En Ontario, en matière civile, ces principes sont imbriqués dans les règlements de la Cour, cf. *Courts of Justice Act,* R.S.O. 1990, c. C.43 et son règlement *Rules of civil procedure*, R.R.O. 1990, Reg. 194 à l'art. 29.1.03(4) : « In preparing the discovery plan, the parties shall consult and have regard to the document titled "The Sedona Canada Principles Addressing Electronic Discovery" developed by and available from The Sedona Conference. O. Reg. 438/08, s. 25. » Pour une application par un tribunal de ces principes, voir *Verge Insurance Brokers* v. *Richard Sherk et al*., 2016 ONSC 4007 (CanLII).

[415] *Sedona Canada E-Discovery Principles*, 3e édition, disponible en ligne : https://thesedonaconference.org/publication/The_Sedona_Canada_Principles

2. Production de preuve par un outil d'IA : preuve d'ADN par génotypage probabiliste

Selon la CDO, l'application logicielle *STRmix*™ serait le premier outil d'IA d'utilisation courante pour produire de la preuve au Canada[416]. Cette application aurait pour fonction d'analyser de la mixture d'ADN et fonctionnerait par génotypage probabiliste (« PG DNA »). Le *Centre of Forensic Sciences* de l'Ontario (depuis août 2016), le *Laboratoire de Sciences Judiciaires et de Médecine Légale du Québec* (depuis février 2018) et la *British Columbia Institute of Technology* (depuis septembre 2018) utiliseraient désormais cette application logicielle spécialisée[417]. L'outil interprète des mixtures complexes composées de plusieurs fragments partiels d'ADN à l'aide de méthodes statistiques et d'algorithmes. Elle permet de comparer deux hypothèses préalablement choisies par l'expert analyste : par exemple, l'expert peut comparer l'hypothèse où l'accusé est à la source de l'ADN avec l'hypothèse où l'accusé n'est pas à la source de l'ADN. L'outil d'IA suggérerait ensuite laquelle de ces hypothèses est la *plus probable* une fois les deux hypothèses mises en relation. La mise en preuve d'un tel résultat peut avoir un effet pervers sur les droits de l'accusé si le juge ou le jury n'est pas bien informé de la *nature* de la suggestion qui est offerte par cet outil. Il s'agit simplement d'une évaluation de la *probabilité* relative d'une hypothèse par rapport à une autre, même si, en réalité, ces hypothèses s'avèrent être fausses toutes les deux[418]. Une telle approche, fondée sur l'hypothèse la *plus probable* parmi deux hypothèses contradictoires est susceptible de venir pervertir la logique derrière la présomption d'innocence et le seuil élevé en matière criminelle qui exige de prouver la culpabilité hors de tout doute raisonnable[419].

Cadre normatif. Il n'existe pas à l'heure actuelle de cadre légal spécifique à même de circonscrire le recours à ce type de preuve. Les règles générales en matière de preuve (*common law*, *Loi sur la preuve au Canada*, *Code criminel*) doivent pallier le vide juridique créé par ces innovations technologiques. Ces règles de *common law* sont appelées par la CDO à être réformées, notamment en inscrivant une présomption d'irrecevabilité des preuves générées par un outil d'IA qui pourrait être repoussée par la Couronne : « Full adherence to placing the burden of proof on the moving party is particularly important to ensure that vulnerable defendants in the criminal justice system are not required to bear a heavy persuasive burden of showing the need for caution and strict scrutiny of novel, AI-based forensic methods. »[420]

Actuellement, comme la production de preuve à l'aide d'un outil d'IA provient d'une pratique réalisée en laboratoire, le recours à cette nouvelle technologie est encadré indirectement par l'entremise des normes de

---

[416] COMMISSION DE DROIT DE L'ONTARIO, Jill R. PRESSER et Kate ROBERTSON (aut.), "AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts", Toronto, Juin 2021 [« CDO2 »]

[417] CDO2, à la p. 7 et à la note de bas de page 6 de la p. 29.

[418] CDO2, p. 11-12 : « The [likelihood ratio] *appears* to answer the question a jury is asked to determine, namely, whether the defendant was the source of some of the DNA at issue. However, LRs do not actually answer this question. Instead, they only "weigh the relative likelihood of two very specific hypotheses." No matter how unlikely two hypotheses are in the real world, when compared and contrasted, one will always be more statistically likely than the other. »; FPT HEADS OF PROSECUTIONS COMMITTEE, "Innocence at Stake: The Need for Continued Vigilance to Prevent Wrongful Convictions in Canada", *Report of the Federal/Provincial/Territorial Heads of Prosecutions Subcommittee on the Prevention of Wrongful Convictions*, 2018, p. 128.

[419] CDO, p. 12-13; Comme on l'explique dans CDO2, à la p.17, la logique derrière le fonctionnement du *STRmix*™ s'apparente à l'approche d'un juge qui a été jugée problématique par la Cour d'appel de l'Ontario, *R.* v. *A.P.*, 2013 ONCA 344 , par. 40 : « The trial judge indicated that in a case such as this, "the truth will be found to be the version of events that is in harmony with what a practical and well-informed person would recognize as most probable in all of the existing circumstances" (emphasis added). » Et au par. 42 : « by setting up the whole case as a choice between two competing versions of events and stating that the version that is "most probable in all of the existing circumstances" will be selected as "true", the trial judge came dangerously close to deciding the ultimate issue at trial on a balance of probabilities. See *R. v. Quidley*, 2008 ONCA 501, 232 C.C.C. (3d) 255. »

[420] CDO2, p. 22-23.

standardisation qui règlementent les activités en laboratoire. En Ontario, les laboratoires doivent se conformer au standard ISO 17025 pour être accrédités. La province s'est assurée de garantir des standards communs d'opérationnalisation en adoptant en 2018 le *Forensic Laboratories Act* : « The accreditation process includes proficiency testing, annual audits, performance reports, surveillance visits, management reviews and a code of conduct. It also requires that reports on testing conducted in the laboratory provide specific information on a form to be prescribed by the Regulations. The Act provides for regular inspections and enforcement mechanisms. »[421] Au Québec, le *Laboratoire de sciences judiciaires et de médecine légale* se conformerait également aux normes ISO 900268, ISO 17025 et à la norme CAN-P-1578[422] et est accrédité par le *Bureau de Normalisation* du Québec[423].

Opacité préservée et fardeau sur l'accusé. Malgré ces garanties de conformité concernant les pratiques en laboratoire, l'utilisation et le fonctionnement du *STRmix™* ne semble pas respecter de garantie minimale en matière de transparence. Le *Centre of Forensic Sciences* de l'Ontario, qui utilise cet outil, admet ne pas avoir accès au code source et il n'a pas l'intention de partager de façon régulière les informations liées à ses examens de validation interne[424]. Par conséquent, le respect de la garantie de transparence de ces outils se retrouve assuré par la politique d'accès aux informations de la compagnie et par ses ententes de confidentialité : « STRmix, will make its underlying information, including source code and foundational validation research, available to defence experts for review upon signing a non-disclosure and confidentiality agreement. However, the restrictions placed on this defence access are so extensive that one is left wondering whether there is much utility in it at all."[425] Ainsi, une requête spécifique devant un juge doit être formulée à chaque fois par l'accusé[426]. Au Québec, dans une affaire où le *STRmix™* avait été utilisé pour produire un des rapports de l'experte en biologie judiciaire, un juge a refusé la requête en divulgation de la preuve faite par un accusé qui cherchait à obtenir des informations liées au processus de validation interne du logiciel *STRmix™* : « *Bien que la preuve entendue démontre que le logiciel STRmix™ est une évolution dans le domaine scientifique*, le requérant n'a démontré aucun élément tendant à établir que la validation de ce logiciel a eu un quelconque impact, si minime soit-il, dans son dossier ou qu'il pourrait avoir une certaine utilité. »[427] Actuellement, le fardeau de la preuve repose donc sur l'accusé qui doit démontrer qu'« il existe une possibilité raisonnable que les renseignements aient une valeur logiquement probante relativement à une question en litige ou à l'habilité à témoigner d'un témoin »[428].

2.1. Preuve d'analyse d'ADN fonctionnant par génotypage probabiliste considérée en tant que « preuve documentaire » ou « preuve matérielle »

[421] FPT HEADS OF PROSECUTIONS COMMITTEE, préc., note 419, p.129.
[422] Nathalie Nicole POIRIER, *L'utilisation de la preuve par l'ADN et ses impacts sur notre société*, Essai fourni à la Faculté de droit En vue de l'obtention du grade de « Maître en droit », Sherbrooke, Faculté de droit, Université de Sherbrooke, p. 20.
[423] *R*. c. *Aithaqi*, 2020 QCCS 870, par. 22.
[424] CDO2, p. 18.
[425] *Id*.
[426] *Id*.
[427] *R*. v. *Aithaqi*, 2020 QCCS 870, par. 107.
[428] *Id*., par. 108.

Les dispositions actuelles concernant la preuve technologique en matière pénale[429] assurent le respect de *la règle de la meilleure preuve*[430] et garantissent *l'intégrité/authentification* des « documents électroniques » qui proviennent d'un « système informatique » (« computer system »)[431]. Ils assurent le principe d'équivalence des supports ce qui permet de présenter validement ces documents électroniques, même une fois imprimés, à titre de preuve documentaire. La preuve issue d'une IA serait-elle assimilable à un système informatique traditionnel ?

L'évolution récente de la *common law* semble permettre également que de l' « information électronique enregistrée automatiquement sans intervention humaine » soit présentée à titre de preuve matérielle (« real evidence »)[432]. Pourrait-on qualifier la preuve produite par le *STRmix*™ de preuve matérielle ?

Plusieurs chercheurs ont jugé problématique que les preuves produites par IA soient présentées sans être automatiquement accompagnées d'un témoignage d'expert afin de vérifier leur admissibilité en preuve. Pour la CDO, « [a]mendments should address what distinctions should be drawn, where necessary and appropriate, between electronic documents stored on traditional information management and computer-processing systems, and new complex AI-based forensic methods. (…) Particular attention should be paid to the circumstances in which parties would be able to adduce electronic records in the absence of an expert human witness to introduce them. »[433] En raison de la nature controversée et inhabituelle des outils d'IA, de leur potentielle impartialité, de leurs potentiels biais[434], de leur opacité, du haut niveau d'expertise requis pour pouvoir évaluer la fiabilité et l'exactitude de ces outils, il nous apparait comme étant nécessaire de s'assurer que la preuve issue d'une IA soit systématiquement produite *dans le cadre* d'un témoignage d'expert afin de déterminer de son admissibilité en preuve. Autrement, cela entrainerait un déséquilibre injuste dans le *rapport de pouvoir* entre l'avocat de la défense et la Couronne lorsque celle-ci fait appel à un tiers qui, lui, a en main les études internes de validation ou le code source[435].

2.2. Preuve d'analyse d'ADN fonctionnant par génotypage probabiliste considérée en tant que témoignage spécial d'expert

Voir-dire sur la science nouvelle. En principe, lorsqu'une expertise se fonde sur une science nouvelle ou qu'elle est utilisée à des fins nouvelles, la partie qui souhaite présenter ce témoignage doit démontrer par prépondérance des probabilités sa « fiabilité » scientifique et juridique. Lors d'un voir-dire - audience parallèle au procès visant à déterminer l'admissibilité d'une preuve, le juge, en respect de son rôle de *gardien*

[429] En droit criminel : *Loi sur la preuve au Canada*, L.R.C. (1985), ch. C-5, art. 31.1 à 31.8. En droit pénal règlementaire : Au Québec, *Code de procédure pénale*, c. C-25.1 et la *Loi concernant le cadre juridique des technologies de l'information*, RLRQ, c. C-1.1; En Ontario, *Evidence Act,* R.S.O. 1990, c. E.23.
[430] La règle de la meilleure preuve trouve son origine en *common law*, elle exige d'une partie qu'elle produise la meilleure preuve que la cause, selon sa nature, permet d'offrir. Par exemple, en matière de preuve écrite, l'original est à privilégier. Voir à ce sujet Giacomo Marchisio, « La règle de la meilleure preuve dans le procès civil », *Revue de Droit de l'Université de Sherbrooke*, 48-2018, p. 1-30.
[431] CDO2, à la p. 35 à la note de bas de page 109.
[432] Concernant l'admissibilité de l' « information électronique enregistrée automatiquement sans intervention humaine » à titre de preuve matérielle, la CDO2 nous renvoie vers *R*. v. *Dennis James Oland*, 2015 NBQB 245 aux paragraphes 99-106.
[433] CDO2, à la p. 23 et à la note de bas de page 110 à la p. 35.
[434] Andrea ROTH, "Machine Testimony", 126 *The Yale Law Journal* 7-1972, p. 2044 : "While such software will surely help combat certain types of bias in forensic interpretation, it will create new types of bias a criminal defendant should have the right to explore."
[435] CDO2, à la p. 23 et à la note de bas de page 111 à la p. 35: « In this review, consideration should be given to the imbalance of knowledge and expertise between vendors of AI-tools and individual defendants in criminal court, and the importance of examining the circumstantial indicators of reliability in evaluating the reliability and accuracy of opaque algorithms. »

de l'intégrité du procès (« gatekeeper »), doit évaluer l'admissibilité d'une expertise portant sur une science nouvelle en fonction des critères établis dans l'arrêt *R*. c. *Mohan* : « (1) la théorie ou la technique peut-elle être vérifiée et l'a-t-elle été? (2) la théorie ou la technique a-t-elle fait l'objet d'un contrôle par des pairs et d'une publication? (3) le taux connu ou potentiel d'erreur ou l'existence de normes, et (4) la théorie ou la technique utilisée est-elle généralement acceptée? » Compte tenu que la preuve ADN par génotypage probabiliste générée par une technologie d'IA touche directement à la « question fondamentale de la culpabilité ou de l'innocence » de l'accusé, chacune des questions devrait alors faire l'objet d'un examen « minutieux » par le tribunal[436].

Même si, à ce jour, il n'y a pas eu d'étude scientifique indépendante, comparative ou évaluée par les pairs permettant d'établir la fiabilité de ces outils en dehors de paramètres très stricts[437], la preuve d'ADN par génotypage probabiliste a toujours été admise en preuve[438] et ce, sans même qu'il y ait la tenue d'un voir-dire de type *Mohan*[439]. Dans les décisions où une telle preuve a été admise, il n'est d'ailleurs pas possible de déterminer, dans les motifs écrits du juge, si ce dernier a porté une attention particulière aux caractéristiques de l'échantillon original d'ADN, notamment en ce qui a trait au nombre de contributeurs dans le mix d'ADN analysé[440]. Le fournisseur admet pourtant une limite en dessous de laquelle l'usage du *STRmix*™ n'est plus valide[441]. Les motifs écrits des juges ayant admis cet outil pour produire de la preuve ne font ni mention d'études de validation externe qui permettrait d'établir la fiabilité scientifique de ces outils, ni de l'état des études de validation interne menées par les laboratoires d'expertise, ni si cet outil respecte les normes internes de ces laboratoires[442].

Pour ces raisons, nous souscrivons aux inquiétudes émises par les commissaires de la CDO qui suggèrent que l'admission actuelle quasi-automatique de ce type de preuve provient de l'« acceptabilité générale » dont bénéficient les innovations technologiques. Cette acceptation généralisée serait d'ailleurs renforcée par les campagnes publicitaires des vendeurs qui mettent l'accent sur le nombre de décisions où le *STRmix*™ a été admis en preuve[443]. Pourtant, au Canada, « l'acceptation généralisée » n'est qu'*un* des *quatre critères* à

---

[436] *R*. c. *Mohan*, (1994) 2 RCS 9; *R*. c. *J.-L.J.*, (2000) 2 RCS 600, par. 33-37.
[437] CDO2, p. 23.
[438] *Id*., à la p. 15 et à la note 64 de la p. 32, où on rapporte 5 décisions dans lesquelles un outil fonctionnaire par génotypage probabiliste a été utilisé : *R.* v. *Ali*, [2019] O.J. No. 5049 (SCJ); *R.* v. *Pereira*, [2019] O.J. No. 3682 (SCJ); *R.* v. *Pereira*, [2019] O.J. No. 2411 (SCJ); *R.* v. *Gautreau*, [2018] O.J. No. 4338 (SCJ); *R.* v. *Dosanjh*, 2019 ONSC 1320.
[439] *R.* v. *Dosanjh*, 2019 ONSC 1320, par. 7 : « The defence concedes that STRmix is not novel science and that probabilistic genotyping evidence, and STRmix in particular, satisfies the criteria for admission into evidence. Both counsel agree that *viva voce* evidence from Mr. Frappier was not required to satisfy my gate keeping function. On my review of the evidence provided, I am satisfied that Mr. Frappier is an expert to provide the opinion evidence as set out in Volumes 1, 2 and 3 of the Crown's DNA application. » CDO2, p. 15 : « In Canada, to date, there appears to have been only one case in which the admissibility of PG DNA evidence was considered, the unreported decision of R. v. Klimowicz. Mr. Klimowicz was unrepresented. He asked no questions of the CFS DNA analyst. There was no defence expert called. Perhaps unsurprisingly, the PG DNA evidence was admitted in Mr. Klimowicz's case. »
[440] CDO2, p. 23 : « the authors have been unable to locate any reported cases in Canada involving PG DNA evidence where the reasons for judgment note the characteristics of the originating DNA sample (including the number of DNA contributors to the mixture that was subject to analysis). The authors have also been unable to find any reported judgments involving PG DNA evidence that advert to the parameters for reliability according to existing validation studies, and whether these were met. Equally, the authors have been unable to locate any reported decisions involving PG DNA evidence that advert to the particular forensic laboratory's validation studies, and whether their own internal standards were met in the particular case. »
[441] *Id*., p. 13 : « This means that PG tools have not been established to be reliable for DNA mixtures where more than three people have contributed DNA, where the minor contributor contributed less than 20% of the DNA, or where there was a low amount of DNA collected. »
[442] *Id*, p. 23.
[443] *Id*, p. 23.

remplir avant de permettre un témoignage qui se fonderait sur une science nouvelle[444]. Cette acceptabilité pourrait également être le fruit de la paresse des avocats de la défense ou de leur ignorance face aux enjeux que soulève l'IA. Cette acceptabilité est certainement renforcée par l'aura de scientificité (*automation bias*) dont bénéficie cette nouvelle technologie. Le juge de la Cour suprême, William Rogers McIntyre, qui s'exprimait alors au nom de la majorité dans une affaire où on a refusé d'admettre en preuve un témoignage soutenu par un détecteur de mensonges, nous mettait déjà en garde contre les biais découlant des innovations technologiques:

> « Il faut se rappeler que toute scientifique que puisse être cette preuve, son utilisation devant le tribunal dépend d'une intervention humaine, celle de l'expert en détecteurs de mensonges. Quels que soient les résultats enregistrés par le détecteur de mensonges, c'est par la bouche de l'expert que leur nature et leur sens sont communiqués au juge des faits. La faillibilité humaine est par conséquent toujours présente, mais on peut dire que maintenant elle est renforcée par la *mystique de la science*. »[445]

Conclusion : pour une approche pragmatique dans la qualification de la preuve. La qualification de la preuve générée par le *STRmix*™ et des preuves générées par des systèmes d'IA en général reste à déterminer au Canada. Néanmoins, certains auteurs canadiens proposent une approche *pragmatique* en soulignant les avantages qui découleraient du fait d'appliquer, par analogie, le cadre normatif s'appliquant aux témoignages d'expert aux preuves ou, plus largement, à toutes prédictions générées par une IA : « The analogy between AI-generated opinions and human expert opinions highlights an existing framework that may be drawn upon to develop robust due process protections applicable to AI-generated evidence. To this end, *Criminal Code* amendments relating to notice requirements, access to disclosure, and requirements regarding the availability of human witness testimony in the courtroom should be carefully considered to address the unique challenges associated with AI-generated evidence such as PG DNA analysis.»[446] Actuellement, les exigences qui s'appliquent lorsqu'une partie veut présenter le témoignage d'un expert comprennent la garantie que l'expert sera présent pour subir un interrogatoire et un contre-interrogatoire lorsque le tribunal l'ordonne, ainsi qu'un préavis raisonnable à l'autre partie lorsqu'elle a l'intention de présenter une preuve d'expert, celui-ci doit inclure une description sommaire des compétences de l'expert et une copie du rapport (art. 657.3 C.cr.). Des garanties similaires pourraient s'appliquer lorsqu'une partie chercherait à mettre en preuve un rapport d'analyse réalisé à l'aide d'un outil d'IA. Cette position était également défendue par les auteurs du Rapport du Citizen Lab pour qui « [a]nalogizing algorithmic policing technologies to the use of expert opinions in the criminal justice system may provide a framework and guiding questions for courts to refer to in assessing the reliability of such technologies. With expert evidence, courts exercise a critical gatekeeper function, ensuring that the admitted evidence is sufficiently reliable, will not usurp the judge or jury, and will not risk undue bias towards the expert. »[447] Cette proposition fait écho à la position défendue par la professeure Andrea Roth aux États-Unis qui soutient que la preuve issue du *STRmix*™ devrait être reçue avec la même retenue que l'est un témoignage d'expert :

> « Applying these principles to machine sources, a jurisdiction might require the proponent of a machine "expert" - a source that generates and conveys information helpful to the jury and beyond the jury's knowledge - to disclose the substance and basis of the machine's conclusion. As one DNA statistics expert told me, "I just want these expert systems to be subject to

[444] *R.* v. *J-LJ*, 2000 SCC 51, par. 34
[445] *R.* c. *Béland*, [1987] 2 RCS 398, par. 20.
[446] CDO2, à la p. 23 et à la note de bas de page 107; Kate ROBERTSON et Jill R. PRESSER, "Algorithmic Technology and Criminal Law in Canada", préc., note 229, p.83.
[447] Citizen Lab, p. 138.

> the same requirements as I am." A jurisdiction might therefore require access to the machine's source code, if a review of the code were deemed necessary to prepare a rebuttal of the machine's claims. »[448]

Les obstacles actuels limitant l'accusé dans l'exercice de son droit de contre-interroger ce « témoin algorithmique » sur les éléments essentiels susceptibles d'établir sa culpabilité seraient à même d'empêcher le juge des faits de « découvrir la vérité et de rendre une décision équitable » contrevenant ainsi à l'art. 7 (défense pleine et entière et procès juste et équitable) et à l'art. 11(d) (présomption d'innocence) de la *Charte*[449]. Appliquer le cadre concernant le témoignage d'expert aux preuves générées par un outil d'IA fonctionnant par génotypage probabiliste permettrait de s'assurer que certaines garanties soient respectées également dans ces circonstances. Cela permettrait de corriger la situation actuelle où l'accès à l'information par la défense est limité par le secret commercial.

[448] Andrea ROTH, préc., note 434, p. 2027.
[449] M. PURCELL et M. ZAIA, préc., note 202, p. 528; *R*. c. *Levogiannis*, [1993] 4 RCS 475; *R*. c. *Seaboyer*, [1991] 2 SCR 577.

LISTE D'ACRONYMES

AIA - Algorithmic Impact Assessment tool

CAI - Commission d'accès à l'information du Québec

CCPE - Centre canadien de protection de l'enfance

C.cr. - Code criminel canadien

CDO - Commission du droit de l'Ontario

Charte - Charte canadienne des droits et libertés

CIPC - Centre d'information de la police canadienne

CLC - Commission des libérations conditionnelles du Canada

CNCEE - Centre national des crimes d'exploitation d'enfants

COMPAS - Correctional Offender Management Profiling for Alternative Sanctions

CPS - Calgary Police Service

CRC - Child Rescue Coalition

CSA - Community Solutions Accelerator

CSDPI - Conseil stratégique des DPI

DPI - Dirigeants principaux de l'information

DRS - Division de la recherche et de la statistique

EPS - Edmonton police service

GPS - Guelph Police Service

GRC - Gendarmerie royale du Canada

LCCJTI - Loi concernant le cadre juridique des technologies de l'information du Québec

LDL - Ligue des droits et Libertés

LSI-OR - Level Service Inventory - Ontario Revised

LSCMLC - Loi sur le système correctionnel et la mise en liberté sous condition

NIJ - National Institute of justice

OPP - Ontario Provincial Police

Ontario’s ICON - Integrated Courts Offenses Network de l’Ontario

PG DNA - ADN par génotypage probabiliste

PMIA - Partenariat Mondial sur l’Intelligence Artificielle

RF - Reconnaissance faciale

SCC - Service Correctionnel Canada

SPPAL - Saskatchewan police predictive analytics lab

SPS - Saskatoon Police Service

SPVM - Service de Police de la Ville de Montréal

SQ - Sûreté du Québec

TPS - Toronto Police Service

WRP - Waterloo Regional Police

VPD - Vancouver Police Department