Université de Montréal

L’intelligence artificielle de recrutement : appréhender les risques de discrimination

Par

Elodie Morton

Faculté de Droit

Mémoire présenté à la Faculté des études supérieures en vue de l’obtention du grade de Maitrise en Droit des Technologies de l’Information (LL.M)

Novembre 2021

Résumé

Traitement des mégadonnées, surveillance, prédictions comportementales ou aides à la décision, les avantages techniques et commerciaux attribués à l'intelligence artificielle emportent l'engouement et l'adhésion des acteurs économiques privés. Ayant vocation à reproduire les facultés cognitives de l'être humain, l'intelligence artificielle s'immisce ainsi progressivement dans nos activités, nos usages et plus largement, dans nos vies.

Pourtant, les défauts de la technologie inquiètent. Utilisée à des fins de reconnaissance faciale, de profilage publicitaire ou encore de recrutement, les biais de l'intelligence artificielle représentent des risques de discriminations pour les personnes qui interagissent avec cette technologie. Or dans un secteur aussi sensible que le recrutement, un tel risque représente un enjeu aussi bien pour les candidats, exposés à une violation de leur droit fondamental à l'égalité, que pour les employeurs qui, eux, s'exposeraient à des sanctions juridiques.

En l'absence d'un cadre juridique spécifique à l'intelligence artificielle, la question se pose donc de savoir si notre droit permet l'appréhension de ces formes de discriminations à l'embauche.

Le propos de ce mémoire consistera donc à proposer des réponses à cette interrogation en trois temps : l'étude du cadre légal applicable, la gestion du risque de biais discriminatoire et l'enjeu de l'accès à la justice des candidats lésés.

Mots-clés : intelligence artificielle, discrimination, recrutement, décision algorithmique, gestion de risques, droit de l'intelligence artificielle, droit des technologies de l'information.

Abstract

Big data processing, surveillance, behavioral predictions or decision aids, the technical and commercial advantages attributed to artificial intelligence have won the enthusiasm and support of private economic players. Designed to reproduce the cognitive abilities of human beings, artificial intelligence is gradually interfering in our activities, our practices and more widely in our lives.

Nevertheless, the flaws of the technology are concerning. Used for facial recognition, advertising profiling or recruitment purposes, the artificial intelligence biases are a risk of discrimination against people who interact with this technology. In a sector as sensitive as recruitment, such a risk constitutes a challenge both for candidates, exposed to a violation of their fundamental right to equality, and for employers who would be exposed to legal sanctions.

Without a specific legal framework for artificial intelligence, the question therefore arises as to whether our law allows for the apprehension of these forms of discrimination in hiring.

The purpose of this thesis will consist in proposing answers to this question in three stages : the study of the applicable legal framework, the challenges of managing the risk of discriminatory bias, and access to justice for aggrieved candidates.

Keywords : artificial intelligence, discrimination, recruitment, algorithmic decision, risk management, artificial intelligence law, information technologies law.

Table des matières

Liste des figures

Liste des sigles et abréviations

Termes relatifs à la législation et à la réglementation

al. : alinéa(s)

art. : article(s)

c. : Chapitre (textes de loi)

c. : contre (décisions)

CCQ : Code civil du Québec

Charte Canadienne : Charte canadienne des droits et libertés

Charte Québécoise : Charte des droits et libertés de la personne

et suiv. : et suivants (pour les articles)

EFVP : Évaluation des facteurs relatifs à la vie privée

LPRPDE : Loi de protection des renseignements personnels et les documents électroniques

LPRPSP : Loi de protection des renseignements personnels dans le secteur privé

L.R.C : Lois révisées du Canada

RLRQ : Recueil des lois et des règlements du Québec

Termes relatifs à la jurisprudence

CSC : Cour suprême du Canada

DTE : Droit du Travail Express

RCS : Recueil des arrêts de la Cour Suprême du Canada

SCR : Canada Supreme Court Reports

TA : Tribunal d'Arbitrage

Termes relatifs aux revues de droit et recueils de doctrine

Can. Bar Rev. : Canadian Bar Review

JCP G : La Semaine Juridique Générale (France)

RGD : Revue générale du droit

RJT : Revue Juridique Thémis

SSRN : Social Science Research Network

Termes relatifs aux institutions de régulation québécoises et fédérales

CAI : Commission d'Accès à l'Information du Québec

CDPDJ : Commission des droits de la personne et de la jeunesse

CNESST : Commission des Normes de l'Équité de la Santé et de la Sécurité au Travail

Commissariat à la vie privée/Commissariat fédéral : Commissariat à la protection de la Vie Privée du Canada

Commission canadienne : Commission canadienne des droits de la personne

Commission québécoise : Commission des droits de la personne et des droits de la jeunesse

Secrétariat du Conseil du Trésor : Secrétariat du Conseil du Trésor du Canada

Tribunal Canadien : Tribunal Canadien des droits de la personne

Tribunal en matière d'emploi : Tribunal de l'équité en matière d'emploi

Termes relatifs aux textes et institutions internationales

Convention sur la Discrimination Raciale : Convention Internationale sur l'Élimination de toutes les formes de Discrimination Raciale

Convention sur la Discrimination à l'égard des Femmes : Convention sur l'Élimination de toutes les formes de Discrimination à l'Égard des Femmes

Déclaration universelle : Déclaration Universelle des Droits de l'Homme

ONU : Organisation des Nations Unies

Pacte sur les droits civils : Pacte International Relatif aux Droits Civils et Politiques

RGPD : Règlement Général sur la Protection des Données

Termes relatifs aux règlements et institutions européens

CNIL : Commission Nationale de l'Informatique et des Libertés (en France)

GT29 : Groupe de Travail « Article 29 »

L'UE : L'Union Européenne

Autres termes

GAFAM : Google Amazon Facebook Apple et Microsoft

IA : intelligence(s) artificielle(s)

IA de recrutement : intelligence(s) artificielle(s) de recrutement

NIST : National Institute of Standards and Technology

À Nounou, TMC et mon Papillon. Mes trois piliers

Remerciements

En premier lieu je tiens à remercier chaleureusement mon directeur, le Pr Benyekhlef. Un mentor de grande qualité qui, avec une bienveillance constante, a su me guider vers l'excellence tout au long de mon projet dont la route fut parsemée d'embuches.

En second lieu, je remercie également Me Camille Aubin ainsi que François Senécal pour leur temps et nos échanges inspirants.

Je remercie mes parents, ainsi que Edwidge L., Adèle T. et Jonathan S. pour leurs relectures, nos échanges enrichissants et leur soutien constant.

Ma gratitude se dirige ensuite tout particulièrement vers quatre personnes : Ève-Marie L., Rachel C., Carole B. et Mylène B.. Pas seulement parce qu'elles ont été des collègues et directrices extraordinaires ; pas seulement parce qu'elles ont été des soutiens et des guides à des moments critiques de mon parcours ; mais surtout, pour la bienveillance et l'esprit de sororité sans commune mesure dont elles m'ont enveloppée. Je les remercie du plus profond de mon cœur.

Enfin, m'accrocher et franchir la ligne d'arrivée n'aurait pas été possible sans l'affection incommensurable que m'ont témoignée Emmy, Dimi et Ma Team Sauvages. Ce mémoire est le résultat d'un entourage de qualité qui m'a soutenue jusqu'au point final de mon mémoire.

Merci à toutes et tous.

Introduction

« Rien de vaste n'entre dans la vie des mortels sans malédiction »[1].

Chaque innovation technique produit sa part d'ombre et de lumière dans une société. Qu'il s'agisse du nucléaire, de l'automobile, d'internet ou des réseaux sociaux, chaque avancée technologique présente autant de bénéfices que d'effets pervers. Ces dernières années, c'est au tour des technologies de l'information d'en faire la démonstration. Amplifiant les mutations sociales à une rapidité sans commune mesure, ces technologies bousculent en effet nos sociétés. Si ces technologies permettent un plus grand accès à l'information, l'organisation citoyenne ou encore l'optimisation de tâches redondantes, l'actualité de ces trois dernières années conduit à une prise de consciente croissante des dangers des intelligences artificielles (ci-après « l'IA »).

Tout au long de l'année 2018, la plateforme de Mark Zuckerberg s'est tristement illustrée par une succession de failles de sécurité, affolant autant le grand public que les législateurs des démocraties occidentales. Au début de l'année 2018, l'affaire « Cambridge Analytica » éclate, éclaboussant au passage la campagne du 45e président des États-Unis d'Amérique[2]. En 2016, l'entreprise britannique a illégalement extrait et analysé les données de plus de 50 millions d'utilisateurs de Facebook, et ce, à des fins de profilage et de communication ciblée dans le cadre de la campagne présidentielle de Donald Trump. Les données volées alimentant le profilage provenaient notamment de tests de personnalité fantaisistes circulant sur Facebook.

Au-delà de l'ébranlement soudain de la vie démocratique d'un État, cette manipulation de masse a révélé aux démocraties occidentales la grande capacité de nuisance d'algorithmes[3] d'IA omniprésents dans leur quotidien. Objets connectés, réseaux sociaux, téléphones

---

[1] Sophocles

[2] Kevin GRANVILLE, « Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens », *The New York Times*, sect. Technology (19 mars 2018), en ligne : <https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html>.

[3] «Séquence de règles opératoires exécutées sur des données et qui permettent l'obtention d'un résultat. », « Algorithme », dans Grand dictionnaire terminologique, coll. Une intelligence artificielle bien réelle : les termes de l'IA, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8367804>.

intelligents, ordinateurs constituent autant de moyens de surveiller, analyser et influencer les comportements et les choix des individus. Omniprésents dans nos quotidiens, nos appareils électroniques connectés représentent des gisements du nouvel or noir : le *Big data*[4]. Ces masses de données (ou mégadonnées) constituent la matière première des algorithmes et de l'IA.

Mais qu'est-ce donc que l'IA ? Plus qu'un domaine d'étude[5], l'IA désigne en premier lieu le « système conçu pour simuler le fonctionnement de l'intelligence humaine afin d'exécuter des fonctions relevant normalement de celle-ci »[6]. Incapables de déroger aux fins pour lesquelles ils sont programmés, les systèmes d'IA fonctionnent par mimétisme. C'est ainsi par le biais des mégadonnées que le comportement humain est observable, puis reproductible pour ces technologies. Loin des mythes d'IA fortes[7] des fictions hollywoodiennes, l'état de l'art permet plutôt de parler d'IA faibles[8], et conduit même parfois à préférer le terme « intelligence augmentée »[9]. Néanmoins, la rapidité d'exécution des analyses des IA ouvre le champ des

[4] « Ensemble d'une très grande quantité de données, structurées ou non, se présentant sous différents formats et en provenance de sources multiples, qui sont collectées, stockées, traitées et analysées dans de courts délais, et qui sont impossibles à gérer avec des outils classiques de gestion de bases de données ou de gestion de l'information.», « Mégadonnées (Big Data) », dans Grand dictionnaire terminologique, coll. Une intelligence artificielle bien réelle : les termes de l'IA, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26507313>.

[5] « Domaine d'étude ayant pour objet la reproduction artificielle des facultés cognitives de l'intelligence humaine dans le but de créer des systèmes ou des machines capables d'exécuter des fonctions relevant normalement de celle-ci. L'intelligence artificielle touche à de nombreux domaines, comme les sciences cognitives et les mathématiques, et à diverses applications, notamment en reconnaissance des formes, en résolution de problèmes, en robotique, dans les jeux vidéo ainsi que dans les systèmes experts. », « Intelligence artificielle », dans Grand dictionnaire terminologique, coll. Une intelligence artificielle bien réelle : les termes de l'IA, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8385376>.

[6] « Système conçu pour simuler le fonctionnement de l'intelligence humaine afin d'exécuter des fonctions relevant normalement de celle-ci. », « Système d'intelligence artificielle », dans Grand dictionnaire terminologique, coll. Une intelligence artificielle bien réelle : les termes de l'IA, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=23798323> (consulté le 28 novembre 2020).

[7] « Système d'intelligence artificielle conçu pour imiter le fonctionnement de l'intelligence humaine dans son ensemble, et ayant la capacité de se questionner, d'analyser et de comprendre ses raisonnements.», « Intelligence artificielle forte », dans Grand dictionnaire terminologique, coll. Une intelligence artificielle bien réelle : les termes de l'IA, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26543873>.

[8] « Système d'intelligence artificielle conçu pour imiter une portion spécifique du fonctionnement de l'intelligence humaine, lui permettant de reproduire certains comportements humains afin d'accomplir une ou des tâches particulières. », « Intelligence artificielle faible », dans Grand dictionnaire terminologique, coll. Une intelligence artificielle bien réelle : les termes de l'IA, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26543874>.

[9] Le terme « intelligence augmentée » vient de la reconnaissance d'un écart palpable entre d'une part, les rêves d'IA fortes capables de substituer ou d'égaler l'intelligence humaine, et d'autre part, l'état de l'art qui semble indiquer

possibles à un point tel, qu'aujourd'hui, les entreprises les plus lucratives sont celles qui ont organisé leur modèle d'affaires sur les données et leur exploitation : les GAFAM[10].

La délégation aux machines de tâches répétitives, chronophages ou complexes est désormais possible. Au nombre des applications bénéfiques de l'IA, et sans prétendre à l'exhaustivité, nous pourrions citer une meilleure détection des tumeurs cancéreuses[11], la facilitation des recherches sur le virus covid-19[12], la mobilité routière[13], la modernisation de l'agriculture[14], une plus grande

---

que la capacité de l'IA se limiterait à la reproduction de certaines tâches ou capacité humaines. Luc JULIA, Ondine KHAYAT et Jean-Louis GASSÉE, *L'intelligence artificielle n'existe pas*, Paris, First éditions, 2019.

« […] it is necessary to introduce human cognitive capabilities or human-like cognitive models into AI systems to develop a new form of AI, that is, hybrid-augmented intelligence. This form of AI or machine intelligence is a feasible and important developing model. Hybrid-augmented intelligence can be divided into two basic models: one is human-in-the-loop augmented intelligence with human-computer collaboration, and the other is cognitive computing based augmented intelligence, in which a cognitive model is embedded in the machine learning system. » dans Nan-ning ZHENG, Zi-yi LIU, Peng-ju REN, Yong-qiang MA, Shi-tao CHEN, Si-yu YU, Jian-ru XUE, Ba-dong CHEN et Fei-yue WANG, « Hybrid-augmented intelligence: collaboration and cognition », (2017) 18-2 *Frontiers Inf Technol Electronic Eng* 153-179, DOI : 10.1631/FITEE.1700053; Maxime AMBLARD, « Idée reçue : Les algorithmes prennent-ils des décisions ? », *Interstices Info* 2018, en ligne : <https://interstices.info/idee-recue-les-algorithmes-prennent-ils-des-decisions/> (consulté le 27 avril 2020); Paul LOUBIÈRE, « Pourquoi l'IA et la voiture autonome sont des mythes », *Challenges* (26 septembre 2019), en ligne : <https://www.challenges.fr/sommet-start-up/sommet-des-startup-de-challenges-pourquoi-l-ia-et-la-voiture-autonome-sont-des-mythes_676606> (consulté le 28 novembre 2020).

[10] GAFAM est l'acronyme désignant Google, Amazon, Facebook, Apple et Microsoft.

[11] Philippe MERCURE, « L'intelligence artificielle contre le cancer », *La Presse*, sect. Sciences (1 mars 2020), en ligne : <https://www.lapresse.ca/actualites/sciences/2020-03-01/l-intelligence-artificielle-contre-le-cancer> (consulté le 28 novembre 2020); Hugo JALINIÈRE, « Dépistage : l'IA du MIT qui prédit l'apparition du cancer du sein », *Sciences et Avenir* (30 mai 2019), en ligne : <https://www.sciencesetavenir.fr/sante/cancer/une-intelligence-artificielle-pour-predire-le-cancer-du-sein_134056> (consulté le 28 novembre 2020).

[12] ICI.RADIO-CANADA.CA, « Une IA détecterait la toux de personnes asymptomatiques atteintes de COVID-19 », *Radio-Canada.ca*, sect. Zone Techno (2 novembre 2020), en ligne : <https://ici.radio-canada.ca/nouvelle/1746457/mit-covid-19-intelligence-artificielle-toux-ia> (consulté le 28 novembre 2020); Philippe MERCURE, « La génomique et l'intelligence artificielle contre la COVID-19 », *La Presse*, sect. Sciences (17 juin 2020), en ligne : <https://www.lapresse.ca/actualites/sciences/2020-06-17/la-genomique-et-l-intelligence-artificielle-contre-la-covid-19> (consulté le 28 novembre 2020).

[13] Arnaud DEVILLARD, « Les voitures autonomes et le casse-tête de la fluidité du trafic », *Sciences et Avenir* (14 novembre 2020), en ligne : <https://www.sciencesetavenir.fr/high-tech/transports/les-voitures-autonomes-et-le-casse-tete-de-la-fluidite-du-trafic_148648> (consulté le 28 novembre 2020); Arnaud DEVILLARD, « Voiture autonome : quand le passager et l'intelligence artificielle collaborent », *Sciences et Avenir* (7 novembre 2019), en ligne : <https://www.sciencesetavenir.fr/high-tech/intelligence-artificielle/conducteur-et-ia-coequipiers-en-vehicule-autonome_138822> (consulté le 28 novembre 2020).

[14] Jean-François VENNE, « L'intelligence artificielle en bref », *Le Devoir* (28 novembre 2020), en ligne : <https://www.ledevoir.com/societe/science/590344/en-bref> (consulté le 28 novembre 2020).

efficacité de la lutte contre la fraude fiscale[15] ou encore l'optimisation des processus de recrutement dans les entreprises[16]. Dans ce dernier cas, les IA facilitent tant la diffusion d'offres d'emploi, la recherche de talents, que le tri des candidatures, voire l'organisation et la réalisation des entretiens d'embauche. Les recruteurs économisent donc un temps considérable grâce à l'IA.

Cependant, comme évoqué en propos liminaire, à l'instar de toute technologie l'IA présente également des externalités négatives. À ce titre, il faut citer la question de l'atteinte aux droits fondamentaux, notamment lorsque l'IA produit des effets sur la situation individuelle d'une personne. Précisons en effet que si l'IA peut reproduire certaines capacités humaines, c'est la prédiction algorithmique qui retiendra notre attention.

Par le profilage algorithmique et ses prédictions subséquentes, l'IA autorise le soutien et la délégation de processus décisionnels. Bien qu'on puisse concéder les gains de temps remarquables qu'amènent les IA d'aide à la décision, il faut rappeler ici que l'IA constitue une création humaine, et donc aussi imparfaite que ses concepteurs. Plus encore, la question des biais de l'IA devient centrale lorsque celle-ci entre en confrontation avec les droits et libertés fondamentaux. Notamment, l'affaire Cambridge Analytica a mis en exergue une incompatibilité de certains usages de l'IA avec la liberté de choix des électeurs américains et britanniques.

De même, reflétant les biais issus des discriminations systémiques des sociétés occidentales, les IA de reconnaissance faciale utilisées dans la justice et les forces de l'ordre se sont déjà révélées

---

[15] « Fraude fiscale : la manne de l'intelligence artificielle », *Le Point* (17 février 2020), en ligne : <https://www.lepoint.fr/economie/fraude-fiscale-la-manne-de-l-intelligence-artificielle-17-02-2020-2363033_28.php> (consulté le 28 novembre 2020); Annick BERGER, « Fraude fiscale : explosion des recouvrements grâce à l'intelligence artificielle », *Capital.fr*, sect. Economie-et-politique (17 février 2020), en ligne : <https://www.capital.fr/economie-politique/fraude-fiscale-explosion-des-recouvrements-grace-a-lintelligence-artificielle-1362432> (consulté le 28 novembre 2020).

[16] Rémy DEMICHELIS, « Le recrutement dans l'oeil de l'intelligence artificielle », *Les Echos*, sect. Tech-Médias (19 février 2019), en ligne : <https://www.lesechos.fr/tech-medias/intelligence-artificielle/le-recrutement-dans-loeil-de-lintelligence-artificielle-991588> (consulté le 30 avril 2020); Quentin PÉRINEL, « Vera, le robot qui recrute dans les grands groupes », *FIGARO* (4 avril 2018), en ligne : <http://www.lefigaro.fr/vie-bureau/2018/04/04/09008-20180404ARTFIG00111-vera-le-robot-qui-recrute-dans-les-grands-groupes.php> (consulté le 7 août 2019); Tomas CHAMORRO-PREMUZIC, « Digital Staffing: The Future of Recruitment-by-Algorithm », *Harvard Business Review* (26 octobre 2012), en ligne : <https://hbr.org/2012/10/digital-staffing-the-future-of> (consulté le 3 novembre 2020).

discriminatoires[17]. En effet, la présence de biais relatifs aux caractères sociodémographiques des individus peut s'expliquer par le manque de diversité dans l'industrie de la Tech[18]. Fondée sur les corrélations, l'IA risque ainsi de reproduire des discriminations systémiques, autant qu'elle peut créer de nouvelles injustices.

Les biais des algorithmes d'IA représentent donc un risque tant pour les individus susceptibles de subir une discrimination, que pour les entreprises qui assumeraient les conséquences de discriminations algorithmiques dont elles n'avaient pas nécessairement connaissance.

L'actualité de l'année 2020 témoigne suffisamment d'une intolérance sociale croissante aux discriminations et traitements inéquitables entre individus. Aussi, plus qu'un risque de déficit d'image, les organisations souhaitant jouir des bénéfices d'IA prédictives se trouvent face à un dilemme : adopter un outil risqué, mais utile, ou bien y renoncer et perdre la chance d'atteindre une performance avantageuse dans leur secteur d'activité. Bien heureusement, le paradigme de la gestion de risques permet bien souvent aux organisations de concilier les risques et les bénéfices de certains outils. En effet, la gestion de risques recouvre :

> « [l'] ensemble des activités qui consistent à recenser les risques auxquels l'entité est exposée, puis à définir et à mettre en place les mesures préventives appropriées en vue de supprimer ou d'atténuer les conséquences d'un risque couru. »[19]

---

[17] Jeff LARSON, Surya MATTU, Lauren KIRCHNER et Julia ANGWIN, « How We Analyzed the COMPAS Recidivism Algorithm », *ProPublica* (23 mai 2016), en ligne : <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=sYBNO6t1202JOb6ILFkA_eTWzPmpol3N> (consulté le 14 juillet 2020); Julia ANGWIN, Jeff LARSON, Surya MATTU, Lauren KIRCHNER, et PROPUBLICA, « Machine Bias », *ProPublica*, éd. ProPublica (2016 2016), en ligne : <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=sYBNO6t1202JOb6ILFkA_eTWzPmpol3N> (consulté le 14 juillet 2020); « Amazon exhorté à ne plus fournir son outil de reconnaissance faciale à la police », *La Presse* (22 mai 2018), en ligne : <https://www.lapresse.ca/techno/actualites/201805/22/01-5182761-amazon-exhorte-a-ne-plus-fournir-son-outil-de-reconnaissance-faciale-a-la-police.php> (consulté le 19 décembre 2018); NIST, « NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software », *NIST* (19 décembre 2019), en ligne : <https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software> (consulté le 13 juillet 2020).

[18] *Discriminating Systems: Gender, Race, and Power in AI*, coll. Publications, New York, AI Now Institute, 2018, en ligne : <https://ainowinstitute.org/discriminatingsystems.html> (consulté le 29 novembre 2019).

[19] « Gestion du risque », dans Grand dictionnaire terminologique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8353448> (consulté le 28 novembre 2020).

Les entreprises peuvent donc intégrer le risque de discrimination algorithmique à leur stratégie de gestion du risque. En la matière, il convient de relever que le droit constitue la pierre angulaire de la gestion de risques de discriminations par IA. Cependant, nonobstant les projets de réformes des législations de protection de la vie privée et des renseignements personnels qui marquent la fin de l'année 2020[20], il faut relever l'absence de législation dédiée à l'IA.

Aussi, la question se pose de savoir dans quelle mesure le droit positif offre des instruments de gestion du risque de discrimination aux organisations utilisatrices et productrices d'IA.

Compte tenu des nombreuses applications de l'IA et de la variété d'aides à la décision existantes, nous articulerons nos réflexions autour des biais discriminatoires des IA utilisées dans le processus de recrutement externe[21] des entreprises. Si les discriminations à l'embauche ont rarement bonne presse pour les organismes qui les pratiquent, elles amènent également la mise en cause de la responsabilité légale de l'organisme. Or, lorsque ce dernier se repose sur des algorithmes d'IA, le relatif vide juridique autour de ces technologies pourrait susciter des hésitations ou des craintes, alors même que ces outils présentent des avantages notables pour les employeurs[22]. Il y a donc plusieurs intérêts à l'étude de la gestion de risque de discrimination par l'IA de recrutement.

---

[20] *Loi sur la protection de la vie privée des consommateurs*, projet de loi n°C-11 (Dépôt-1ère lecture - 17 novembre 2020), 2ème sess., 43e légis. (Can) [*Projet de loi C-11*], en ligne : <https://parl.ca/DocumentViewer/fr/43-2/projet-loi/C-11/premiere-lecture#ID2E0IB0BA> (consulté le 28 novembre 2020); *Loi modernisant des dispositions législatives en matière de protection des renseignements personnels*, RLRQ, 64 (Sanction - 22 septembre 2021) [*Loi n°64*], en ligne : <http://www.assnat.qc.ca/fr/travaux-parlementaires/projets-loi/projet-loi-64-42-1.html> (consulté le 14 novembre 2021).

[21] « Recrutement externe », dans Grand dictionnaire terminologique, coll. Gestion, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8392912> (consulté le 24 juin 2020).

[22] Le vice-président de l'Association nationale des directeurs des ressources humaines français, Benoît Serre, indiquait au journal Les Echos que « L'IA fait gagner du temps, elle peut aussi nous aider à estimer le besoin de formation des candidats […] », dans Rémy DEMICHELIS, « Le recrutement dans l'œil de l'intelligence artificielle », *Les*

En premier lieu, l'accès à l'emploi, notoirement concurrentiel, constitue un contexte propice à la mise en exergue des bénéfices des IA, notamment parce qu'elles facilitent le tri des candidatures.

Ensuite, la recherche d'emploi est fortement touchée par la numérisation et les IA : la diffusion d'offres d'emploi, les candidatures en ligne, la gestion des candidatures reçues, le classement des candidatures reçues et même les entretiens constituent autant d'étapes d'un processus de recrutement potentiellement appuyé, voire délégué, à une IA. Or, s'agissant du tri de candidature, cette étape chronophage est cruciale lors d'un processus de recrutement. En cas de biais affectant l'IA de recrutement, les conséquences paraissent évidentes : les personnes candidates s'exposent à des discriminations à l'embauche.

Enfin, le secteur du recrutement s'est déjà heurté aux imperfections des IA ; l'exemple récent le plus retentissant fut celui de l'IA d'Amazon[23] dont les biais sexistes ont été dénoncés.

Le choix d'une application de l'IA aussi sensible que celle du recrutement nous offre donc l'occasion de questionner le cadre juridique existant, souhaité et utile à la gestion du risque juridique de discrimination dans ce contexte. De plus, une telle perspective autorise le croisement des perspectives fédérale, québécoise et européenne au cours de nos développements.

La gestion du risque de discrimination à l'embauche que soulève l'IA de recrutement constituera donc notre prisme d'étude des particularités d'une discrimination causée par une IA.

Les normes applicables à une discrimination survenue au cours d'un recrutement augmenté (Chapitre 1), les instruments juridiques de mitigation des principaux biais de l'IA de recrutement (Chapitre 2), ainsi que les enjeux d'un accès à la justice des candidats lésés (Chapitre 3) constituent les trois axes autour desquels nos réflexions s'articuleront.

Echos, sect. Tech-Médias (19 février 2019), en ligne : <https://www.lesechos.fr/tech-medias/intelligence-artificielle/le-recrutement-dans-loeil-de-lintelligence-artificielle-991588> (consulté le 30 avril 2020).

[23] Amaury BUCCO, « Le logiciel de recrutement d'Amazon n'aimait pas les femmes », *FIGARO* (11 octobre 2018), en ligne : <http://www.lefigaro.fr/social/2018/10/11/20011-20181011ARTFIG00096-le-logiciel-de-recrutement-d-amazon-n-aimait-pas-les-femmes.php> (consulté le 7 août 2019).

CHAPITRE 1. LE CADRE JURIDIQUE DES DISCRIMINATIONS PAR INTELLIGENCE ARTIFICIELLE DE RECRUTEMENT

Dans l'étude du risque de discrimination à l'embauche par des outils d'intelligence artificielle de recrutement (ci-après « IA de recrutement ») il s'agira, dans le cadre de ce premier chapitre, d'en explorer le cadre juridique. Pour cela trois questionnements serviront de guides : qu'est qu'une discrimination dans l'ordre juridique qui nous intéresse ? Est-il légal de prendre une décision de recrutement (ici de rejet de candidature) par algorithme ? Et enfin, comment sont traitées les discriminations en matière d'emploi ?

Pour répondre à ces interrogations, nous rappellerons, dans un premier temps, les différents textes portant sur les discriminations (A). Puis, dans un second temps, nous mettrons en exergue le cadre légal naissant relatif aux décisions par algorithmes (B).

A. Le cadre juridique des discriminations

L'appréhension juridique des discriminations s'opère à travers deux volets que sont leur définition et leur sanction. Dans le premier volet, il sera question de rappeler les protections constitutionnelles et internationales du droit à l'égalité (1), avant d'aborder l'existence de sanctions des discriminations à l'embauche, tant québécoises que fédérales (2). Enfin, le troisième volet mettra en évidence l'appréhension de la discrimination en matière d'emploi (3).

1. La protection du droit à l'égalité : une définition extensive

La définition du droit à l'égalité et ses implications en matière de discriminations se retrouvent au sein de deux sources : les sources constitutionnelles et les sources internationales.

Au Canada, les Chartes consacrent des articles spécifiques au droit à l'égalité. Il convient cependant de remarquer que si ces textes nationaux permettent une protection étendue, celle-

ci est également fortement influencée par le droit international. À ce titre, nous étudierons donc cinq conventions.

a) Les chartes et textes constitutionnels.

L'égalité se définit comme étant le traitement non différencié des personnes aussi bien en droit, qu'en fait[24]. Dans sa formulation négative, l'égalité incarne donc le droit à ne pas subir de discrimination[25]. En la matière, les Québécois bénéficient d'une protection constitutionnelle du droit à l'égalité tant au niveau national, qu'au niveau provincial. Le droit à l'égalité (ou à la non-discrimination) constitue donc un droit fondamental à part entière.

Au niveau national, la *Charte canadienne des droits et libertés* (ci-après « Charte Canadienne ») figurant dans la *Loi constitutionnelle de 1982*, énumère et définit les droits et libertés fondamentaux garantis. On y retrouve l'article 15 prohibant toute forme de « discriminations fondées sur la race, l'origine nationale ou ethnique, de leur couleur, de leur religion, le sexe, l'âge ou de leurs déficiences mentales ou physiques[26] ».

La Cour suprême a par ailleurs défini la discrimination comme constituant :

> « une distinction, intentionnelle ou non, mais fondée sur des motifs relatifs à des caractéristiques personnelles (…), qui a pour effet d'imposer à cet individu (…) des fardeaux, des obligations ou des désavantages (…) ou d'empêcher ou de restreindre l'accès aux possibilités, aux bénéfices et aux avantages offerts à d'autres membres de la société (…) »[27].

[24] Michel FILION, « égalité », dans Dictionnaire encyclopédique du Droit québécois, Gaudet Éditeur ltée, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26548045> (consulté le 23 avril 2020).

[25] Hubert REID et Simon REID, *Dictionnaire de droit québécois et canadien*, 2016, p. 137.

[26] *Charte Canadienne des Droits et Libertés*, (1982) Loi constitutionnelle de 1982, partie I, c. 11, en ligne : <https://laws-lois.justice.gc.ca/fra/Const/page-15.html> (consulté le 23 avril 2020).

[27] *Andrews c. Law Society of British Columbia*, [1989] 1 RCS 143 (CSC), en ligne : <http://canlii.ca/t/1ft8r> (consulté le 30 juin 2020); Jean-François TRUDEL, *Mémoire à la Commission d'Accès à l'Information sur le document de consultation « Intelligence artificielle »*, Cat. 2.412.133, CDPDJ, 2020, p. 9, en ligne : <https://www.cdpdj.qc.ca/Publications/memoire_consultation_CAI_IA.pdf>.

Au Québec, la Charte des droits et libertés de la personne (ci-après « Charte Québécoise ») réserve un chapitre entier au droit à l'égalité aux articles 10 et suivants. Y sont détaillées les diverses formes de traitements discriminatoires contraires à l'égalité, et c'est à son article 16 que la Charte Québécoise consacre la prohibition de la discrimination à l'embauche[28]. Cette loi quasi-constitutionnelle[29] s'applique aussi bien à l'État provincial québécois et ses démembrements, qu'au secteur privé[30]. En effet, sous réserve de la répartition des compétences entre les législateurs québécois et fédéral, la Charte Québécoise s'applique ainsi à l'ensemble des personnes morales du Québec, qu'elles soient rattachées aux secteurs public ou privé. Ainsi, le droit à l'égalité garanti par la Charte provinciale s'applique autant au gouvernement provincial, qu'aux personnes morales privées[31]. En outre, la coexistence de ces deux chartes renforce les droits qu'elles protègent[32].

Ainsi, par ces textes constitutionnels la discrimination sous toutes ses formes, y compris à l'embauche, est par principe proscrite. Par ailleurs, des protections supplémentaires contre les ruptures d'égalité existent également à l'échelle internationale.

---

[28] *Charte des Droits et Libertés de la Personne*, (1975), RLRQ c C-12, en ligne : <http://legisquebec.gouv.qc.ca/fr/showdoc/cs/C-12> (consulté le 23 avril 2020). « Nul ne peut exercer de discrimination dans l'embauche, l'apprentissage, la durée de la période de probation, la formation professionnelle, la promotion, la mutation, le déplacement, la mise à pied, la suspension, le renvoi ou les conditions de travail d'une personne ainsi que dans l'établissement de catégories ou de classifications d'emploi »

[29] *Singh v. Minister of Employment and Immigration*, [1985] 1 SCR 177, par. 85 (SCC), en ligne : <http://canlii.ca/t/1fv22> (consulté le 13 septembre 2020). « Thus, the Canadian Bill of Rights retains all its force and effect, together with the various provincial charters of rights. Because these constitutional or quasi-constitutional instruments are drafted differently, they are susceptible of producing cumulative effects for the better protection of rights and freedoms ».

[30] Andre MOREL, « La Charte Quebecoise: Un Document Unique dans l'Histoire Legislative Canadienne », (1987) 21-1 *R.J.T. n.s.* 1-24, sect. « III. Le régime de protection des droits et libertés », en ligne : <https://heinonline.org/HOL/P?h=hein.journals/revjurns21&i=24> (consulté le 13 septembre 2020).

[31] *Loi constitutionnelle de 1867*, 1867, Last Modified: 2015-07-30, part. IV. Distribution des pouvoirs législatifs [*L.R.C*], en ligne : <https://laws-lois.justice.gc.ca/fra/Const/page-4.html#h-17> (consulté le 13 septembre 2020). Ainsi, en dehors de secteurs réputés stratégiques comme l'aéronautique, les télécommunications, les banques ou encore le secteur militaire, les entreprises privées entrent dans le champ d'application de la Charte Québécoise.

[32] *Id. 6* ; Michel COUTU et Pierre BOSSET, *Etude 6 : La dynamique juridique de la Charte*, 6, coll. Études, vol.2, CDPDJ, 2003, p. 264, en ligne : <http://www.crimt.org/Publications/CDPDJQ_APRES_25_ANS_2003_II.pdf> (consulté le 12 septembre 2020).

b) Les conventions internationales

Au nombre des traités et conventions internationales que le Canada reconnaît, le principe d'égalité constitue le pilier de cinq d'entre eux : la Déclaration Universelle des Droits de l'Homme[33] (ci-après « Déclaration universelle »), la Convention Internationale sur l'Élimination de toutes les formes de Discrimination Raciale[34] (ci-après « Convention sur la Discrimination Raciale »), le Pacte International Relatif aux Droits Civils et Politiques[35] (ci-après « Pacte sur les droits civils »), la Déclaration des Nations unies sur les Droits des Peuples Autochtones[36], et la Convention sur l'Élimination de toutes les formes de Discrimination à l'Égard des Femmes[37] (ci-après « Convention sur la Discrimination à l'égard des Femmes »). Bien qu'ils n'aient pas de valeur contraignante, ces traités illustrent aussi bien l'approche extensive du principe d'égalité, que son évolution. Quelques développements à leur propos nourriront notre étude, avant de nous attarder sur la protection spécifiquement accordée au droit à l'égalité dans un contexte professionnel.

En premier lieu, la Déclaration universelle et le Pacte sur les droits civils posent le principe d'égalité dans son caractère universel, transcendant toutes les différences. La Déclaration universelle représente l'instrument incontournable en matière d'égalité. En effet, du traumatisme de la Seconde Guerre mondiale, nait une volonté de pacification durable des relations entre États, et plus largement des relations entre humains. Cet instrument constitue

---

[33] *Déclaration Universelle des Droits de l'Homme*, (1948), résolution 217 A (III), art. 1, 2 et 7 [*DUDH*], en ligne : <https://www.un.org/fr/universal-declaration-human-rights/index.html> (consulté le 2 mai 2020).

[34] *Convention internationale sur l'élimination de toutes les formes de discrimination raciale*, (1966) Recueil des Traités, résolution 2106 (XX), art. 1 à 5 (ratifié le 14/10/1970), en ligne : <https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-2&chapter=4&clang=_fr> (consulté le 2 mai 2020).

[35] *Pacte International Relatif aux Droits Civils et Politiques*, (1966) Recueil des Traités, art. 2 et 3, en ligne : <https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-4&chapter=4&clang=_fr> (consulté le 2 mai 2020).

[36] *Déclaration des Nations Unies sur les droits des peuples autochtones*, (2007), Résolution 61/295 [*Déclaration sur les droits des peuples autochtones*].

[37] *Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes*, (1979) Recueil des Traités, résolution 34/180, en ligne : <https://www.ohchr.org/FR/ProfessionalInterest/Pages/CEDAW.aspx> (consulté le 2 mai 2020).

l'acte clé de l'Organisation des Nations Unies (ci-après « l'ONU »). À ce titre, non seulement les États s'engagent à une intervention sur le plan juridique[38], mais également à développer des politiques qui créent et renforcent les conditions d'un exercice plein et entier des droits fondamentaux protégés dans les traités portés par l'ONU. Les traités et déclarations subséquents à la Déclaration universelle s'analysent comme des précisions du spectre des droits fondamentaux, dont la clé de voute serait le droit à l'égalité. On y retrouve donc une égalité à large spectre, déclinée dans les sphères productrices d'interactions humaines[39].

Pour autant, la persistance de certaines discriminations à l'égard de certains groupes de personnes, aboutit à trois textes procurant une protection particulière aux femmes, aux Autochtones et aux groupes racialisés : la Convention sur la Discrimination à l'égard des Femmes, la Déclaration sur les Droits des Peuples Autochtones et la Convention sur la Discrimination Raciale. À travers ces trois instruments, deux observations émergent.

D'abord, la protection accordée par la Déclaration universelle et le Pacte sur les droits civils repose sur les États. Ces textes, aussi protecteurs soient-ils, d'une part, sont pour certains dénués de caractère contraignant. D'autre part, ils dépendent de l'investissement des États quant à l'acquittement de leurs obligations, et de la « bonne foi »[40] dont ils font preuve dans cet aspect de leur office. À ce propos, les trois instruments mentionnés précédemment commencent par

[38] Les interventions sur le plan juridique impliquent l'obligation des États d'abolir tous textes, ou pratique institutionnelle produisant de la discrimination, ainsi que des sanctions des discriminations par le biais des juridictions. Voir à ce propos les articles 1 à 3 de la Déclaration universelle.

[39] *Pacte International Relatif aux Droits Civils et Politiques*, préc., note 35, art. 1 et 2. À ce propos le Pacte sur les droits civils précise les conséquences du principe d'égalité sur les plans démocratiques, politiques, culturels et socio-économiques dans son article premier et deuxième.

[40] Dans la déclaration relative aux droits des peuples Autochtones on peut lire que « Les dispositions énoncées dans la présente déclaration seront interprétées conformément aux principes de justice, de démocratie, de respect des droits de l'homme, d'égalité, de non-discrimination, de bonne gouvernance et de bonne foi ». De même, le Pacte international relatif aux droits civils énumère les engagements contraignants mais uniquement pour les États signataires. *Déclaration sur les droits des peuples autochtones*, préc., note 36, art. 46 3); *Pacte International Relatif aux Droits Civils et Politiques*, préc., note 35, art. 2.

réaffirmer que les droits garantis par la Déclaration universelle s'appliquent aux trois groupes pour lesquels les discriminations persistent et ne reculent pas[41].

De plus, l'ONU définit spécifiquement les discriminations visant les groupes racialisés et les femmes dans les traités qui leur sont consacrés. Malheureusement la déclaration dédiée aux Autochtones ne bénéficie pas d'un effort de définition similaire. En dépit de cette omission regrettable, les cinq instruments susmentionnés reconnaissent néanmoins, et renforcent, les obligations des États à agir activement, et rapidement, pour remédier aux injustices et traitements inégalitaires subis par ces populations.

Aussi, la garantie du respect de ces obligations passe notamment par des sanctions des ruptures d'égalité auxquelles les États s'engagent[42]. Au regard des cinq instruments soutenus[43] par le gouvernement fédéral, ainsi que des deux Chartes rappelées précédemment, le principe d'égalité ferait l'objet d'une protection satisfaisante, ne serait-ce que par son périmètre de définition. La sanction des situations discriminatoires étant renvoyée aux droits nationaux, il convient de s'intéresser aux lois canadiennes. Il est pertinent de relever que la protection des droits fondamentaux définie dans ces deux Chartes est complétée par des lois spécifiques. Au Québec, la Commission québécoise des droits et libertés et le Tribunal des droits de la personne constituent les organes administratifs de référence dans l'application de la Charte québécoise, tandis qu'au niveau national, tous les tribunaux appliquent la Charte canadienne. La Commission canadienne des droits de la personne (ci-après « Commission canadienne ») et le Tribunal

[41] *Pacte International Relatif aux Droits Civils et Politiques*, préc., note 35; *Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes*, préc., note 37, part. Préambule; *Convention internationale sur l'élimination de toutes les formes de discrimination raciale*, préc., note 34, art. 1er 4). La Convention sur la Discrimination à l'égard des Femmes déplore explicitement le nombre d'instruments et politiques dédiés à la lutte contre les discriminations, sans avancées significatives. De même, la Convention sur la discrimination raciale rappelle la nécessité de maintenir des politiques jusqu'à ce que l'objectif d'égalité véritable soit atteint.

[42] *Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes*, préc., note 37, art. 2; *Déclaration sur les droits des peuples autochtones*, préc., note 36, art. 26 al.3 et 40; *Convention internationale sur l'élimination de toutes les formes de discrimination raciale*, préc., note 34, art. 4.

[43] Si la Convention sur la Discrimination Raciale, la Convention sur la Discrimination à l'égard des Femmes et le Pacte sur les droits civils font l'objet de ratifications, il faudrait remarquer que la Déclaration universelle et la Déclaration sur les Droits des Peuples Autochtones ne donnent pas lieu à ratification.

Canadien des droits de la personne (ci-après « Tribunal Canadien ») sont en sus compétents s'agissant de la Loi fédérale sur les droits de la personne.

2. Le traitement des cas de rupture d'égalité : un dialogue institutionnel entre les commissions et tribunaux spécialisés

Dans l'analyse du cadre légal de la discrimination à l'embauche, le système de sanction en place nous apportera une vue d'ensemble sur l'appréhension du droit à l'égalité au Canada et au Québec.

Au Canada, les lois fédérale et québécoise sur les droits de la personne créent chacune deux organes dédiés à la protection et la promotion de l'égalité. En l'occurrence, ce dialogue institutionnel entre une autorité administrative indépendante et un tribunal spécialisé se constate tant au niveau québécois qu'au niveau fédéral. Aussi, nous nous consacrons ici à l'approche canadienne de la protection du droit à l'égalité, par le biais des commissions administratives et des tribunaux administratifs qui assurent la sanction des discriminations en vertu des lois qui y sont consacrées.

En effet, l'existence des commissions aux plans national et québécois permet en premier lieu la promotion et la régulation des droits fondamentaux. En second lieu, à travers ces deux autorités administratives, c'est un premier point de régulation et de contrôle du respect du droit à l'égalité qui opère. Pour autant, il convient de remarquer un deuxième point de contrôle qui permet la sanction judiciaire : les tribunaux spécialisés en matière de droits fondamentaux. Nous présenterons donc l'arrimage entre commissions et tribunaux des droits de la personne au niveau national d'abord, au niveau québécois ensuite.

a) Les institutions canadiennes

Au niveau fédéral, la loi constituant la Commission canadienne des droits de la personne (ci-après « Commission canadienne ») définit la discrimination comme constituant « un motif de distinction illicite »[44]. Les définitions et les motifs de discriminations, ou de harcèlements, sont sensiblement les mêmes dans la loi québécoise et la loi canadienne. À l'instar de la Commission québécoise au Québec, la Commission canadienne est l'organe national crucial de sensibilisation, de promotion et de surveillance des droits garantis par une loi canadienne opposable au gouvernement national, et au Parlement[45].

Instituée par *la Loi canadienne sur les droits de la personne*[46], la Commission canadienne est notamment dotée des pouvoirs de réception des plaintes de discrimination et de saisine du Tribunal Canadien[47]. De même qu'au niveau québécois, la Commission canadienne est habilitée à diligenter des enquêtes afin d'éclairer son jugement quant au bien-fondé, ou à la recevabilité, de la plainte qui lui est soumise[48].

À l'instar du Tribunal des droits de la personne, c'est à l'issue d'une première instruction de la Commission canadienne que le Tribunal canadien des droits de la personne (ci-après le « Tribunal Canadien ») est saisi. Les juges exercent leur pouvoir d'enquête en désignant des enquêteurs dotés des pouvoirs analogues à ceux des juges des cours supérieures[49]. À l'issue de cette phase d'enquête, le Tribunal Canadien peut décider la condamnation de l'auteur des discriminations à différentes peines ; notamment, l'adoption de mesures ou programmes (qui ont vocation à corriger les comportements discriminatoires), la réparation des préjudices causés par

[44] *Loi Canadienne sur les droits de la personne*, (1985), ch. H-6, L.R.C., art. 3 et suivants, en ligne : <https://laws-lois.justice.gc.ca/fra/lois/H-6/page-1.html> (consulté le 24 mai 2020).

[45] *Charte Canadienne des Droits et Libertés*, (1982) Loi constitutionnelle de 1982, partie I, c. 11, art. 32, en ligne : <https://laws-lois.justice.gc.ca/fra/Const/page-15.html> (consulté le 23 avril 2020).

[46] *Loi Canadienne sur les droits de la personne*, préc., note 44, art. 26 et 27.

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

[48] *Id.*, art. 40 (5) et suiv.

[49] *Id.*, art. 48.3 (3) et suiv.

la discrimination établie, ou encore en cas d'actes discriminatoires intentionnels, la condamnation à des dommages et intérêts punitifs[50].

b) Les institutions québécoises

La Commission des droits de la personne et des droits de la jeunesse (ci-après « Commission québécoise ») constitue l'organe dévolu à la promotion de la Charte québécoise auprès des citoyens[51]. Cette institution a également vocation à enquêter, puis à saisir le Tribunal des droits de la personne des plaintes des justiciables victimes de discriminations. À cet égard, la Commission québécoise définit trois types de discriminations : les discriminations directes et indirectes, et la discrimination systémique[52]. Cette typologie permet de distinguer des actes, propos, ou comportements intentionnels, de ceux qui causent des discriminations de façon non intentionnelle, voire qui résultent de la manifestation des oppressions systémiques[53]. Après une enquête non contradictoire[54], lorsque la Commission québécoise identifie un cas de discrimination ou de harcèlement discriminatoire, celle-ci est habilitée à saisir le Tribunal des droits de la personne.

Au-delà de son rôle procédural en matière de réception des plaintes, la Commission québécoise est également habilitée à produire des rapports, des recommandations en matière d'inégalités systémiques. De même, elle est compétente pour contrôler et promouvoir les bonnes pratiques respectueuses du droit à l'égalité.

---

[50] *Id.*, art. 53 (2) et suiv.

[51] *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 57 et suiv.; CDPDJ, « Origine et mission de la Commission des droits de la personne et des droits de la jeunesse », *La Commission*, en ligne : <http://www.cdpdj.qc.ca/fr/commission/Pages/default.aspx> (consulté le 22 mai 2020).

[52] CDPDJ, « Discrimination : pratique interdite », *Pratiques interdites*, en ligne : <http://www.cdpdj.qc.ca/fr/droits-de-la-personne/pratiques/Pages/discrimination.aspx> (consulté le 22 mai 2020); « Lexique | CDPDJ », *Lexique*, en ligne : <http://www.cdpdj.qc.ca/fr/pages/lexique.aspx#lexiqueD> (consulté le 22 mai 2020).

[53] Les discriminations systémiques, par essence, peuvent être inconscientes et sont intégrées aux interactions sociales.

[54] *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 17.

Le Tribunal des droits de la personne est un tribunal québécois spécialisé, dont la compétence est dévolue au contentieu relatif aux discriminations, ou aux cas de harcèlements dont les motifs sont discriminatoires[55].

> « Le Tribunal a compétence pour entendre et disposer de toute demande portée en vertu de l'un des articles 80, 81 et 82 (…), ou en vertu de l'un des articles 88, 90 et 91 relativement à un programme d'accès à l'égalité »[56].

À cet effet, la Commission québécoise lui transmet les dossiers susceptibles de donner lieu à sanction. Le Tribunal québécois est compétent pour trancher aussi bien des atteintes aux droits fondamentaux entre personnes privées[57], que des atteintes survenues entre des personnes privées et la province du Québec[58]. Dotés de pouvoirs d'enquête similaires à ceux d'un juge de Cour supérieure (excepté celui d'emprisonner les individus)[59], les membres du Tribunal disposent de pouvoirs significatifs en matière de discrimination et de harcèlement discriminatoire. En effet, les juges peuvent recourir aussi bien à des auditions de personnes, à des assignations à produire des documents, qu'à des citations à comparaître ou à des injonctions par ordonnance[60].

Enfin, une fois l'enquête close, le tribunal est habilité à prononcer des sanctions proportionnées à la gravité des faits discriminatoires reprochés. Par ce pouvoir d'ordonnance, les juges peuvent enjoindre toute mesure nécessaire à la cessation de l'acte discriminatoire, prononcer la condamnation au paiement de dommages et intérêts au titre de la réparation des

---

[55] Il s'agit des motifs de discrimination interdits aux articles 10 et suivants de la Charte Québécoise. *Id.*, art. 10 et suiv.

[56] *Id.*, art. 111.

[57] Ici le termes personnes privées renvoie aux personnes physiques et morales. Les relations visées sont donc celles entre personnes physiques, entre personnes physiques et morales, ou entre personnes morales.

[58] *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 54.

[59] *Loi sur les commissions d'enquête*, (1964), RLRQ c C-37, art. 1 et suiv., en ligne : <http://legisquebec.gouv.qc.ca/fr/showDoc/cs/C-37?&digest=> (consulté le 23 mai 2020); *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 113.

[60] *Loi sur les commissions d'enquête*, préc., note 59, art. 7 à 13; *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 123.

dommages causés, ou encore la condamnation à des dommages et intérêts punitifs lorsque le tribunal a constaté des actes intentionnels[61].

Ainsi, à la lumière des institutions et juridictions présentées, il convient de conclure que le fédéral et le Québec se sont dotés d'instruments conformes à la protection des droits fondamentaux garantis par les Chartes et conventions internationales.

La dualité organique permet effectivement une complémentarité entre une approche préventive et une approche répressive des discriminations : tel qu'il a été mentionné précédemment, les deux commissions incarnent davantage l'approche préventive, alors que les tribunaux assurent la répression. À ce propos, il faudrait relever qu'avant d'aboutir à une saisine du tribunal compétent, chacune des commissions passe par la recommandation de mesures qui ont vocation à corriger ou à compenser les discriminations constatées. Les commissions privilégient un dialogue avec les personnes et organisations contrevenant aux chartes. La sanction des atteintes au droit à l'égalité passe donc en premier lieu par une approche pédagogique, négociée avec les acteurs concernés. Cette intervention circonstanciée fait également écho à la protection internationale qui organise l'articulation d'obligations positives, avec la répression des atteintes au droit à l'égalité. Nous avons pu observer que les cadres et procédures juridiques en matière d'égalité offrent donc une protection relativement complète face aux discriminations. Par conséquent, *a priori*, les discriminations par IA de recrutement entreraient dans le champ d'application de ces procédures.

Néanmoins, il est à noter que si les textes nationaux précités ont créé des tribunaux spécifiques au traitement du contentieux de la discrimination, le contentieux en matière de discrimination à l'embauche fait l'objet de dispositions spécifiques. Aussi, il convient de s'intéresser à la protection face aux discriminations dans le cadre professionnel.

[61] *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 49 et 130, 131, 136.

3. La discrimination en matière d'emploi : un traitement spécifique

S'agissant des fondements juridiques de la discrimination dans le cadre professionnel, ceux-ci existent aussi bien dans les conventions internationales que dans les textes nationaux cités précédemment. Tous visent expressément la question discriminatoire dans le cadre de la vie professionnelle, et notamment la discrimination à l'embauche[62]. Deux lois particulières prennent en compte cette problématique et donnent compétence aux commissions et tribunaux précités.

a) La Commission canadienne des droits de la personne

Au titre de l'article 22 de la *Loi sur l'équité en matière d'emploi*, la Commission canadienne est compétente pour contrôler l'application de ladite loi.

> « 22 (1) La Commission est responsable de la détermination de l'observation par les employeurs des articles 5, 9 à 15 et 17. »[63]

La Commission peut donc déléguer les contrôles à ses agents, dits agents d'application, ou effectuer une enquête conformément aux pouvoirs généraux qui lui sont conférés par la Loi canadienne sur les droits de la personne.

De plus, en cas de non-respect de ses obligations par l'employeur, conformément aux « orientations générales » de la loi, la Commission recourt à son pouvoir d'ordonnance et d'injonction « en dernier lieu » au profit d'une approche fondée sur la négociation et l'accompagnement des employeurs dans l'établissement de politiques conformes[64]. On retrouve ici cette gradation de l'intervention. En cas de difficultés ou de manque de coopération de bonne foi de l'employeur, l'agent d'application en réfère à la Commission qui, elle, est habilitée à

---

[62] Pour les trois groupes faisant l'objet d'une protection spécifique, on retrouve dans les divers instruments des articles visant les discriminations dans la formation et la vie professionnelles. *Déclaration sur les droits des peuples autochtones*, préc., note 36, art. 17 al.3; *Convention internationale sur l'élimination de toutes les formes de discrimination raciale*, préc., note 34, art. 5 e); *Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes*, préc., note 37, art. 10 et 11.

[63] *Loi sur l'équité en matière d'emploi*, (1995), L.C. 1995, c. 44, Last Modified: 2017-12-12, en ligne : <https://laws-lois.justice.gc.ca/fra/lois/e-5.401/page-1.html> (consulté le 30 mai 2020).

[64] *Id.*, art. 22.

communiquer ses injonctions à l'employeur. Si en dépit de cet ordre, la Commission constate la persistance des manquements de l'employeur, elle peut saisir le Tribunal de l'équité en matière d'emploi (ci-après « Tribunal en matière d'emploi »). De même, si l'employeur souhaite une révision de l'injonction ou de l'ordonnance de la Commission, il peut lui aussi saisir ce même tribunal.

b) Le Tribunal de l'équité en matière d'emploi

À l'instar du dialogue entre la Commission canadienne et le Tribunal en matière d'emploi, on retrouve au Québec la même répartition des rôles entre la Commission québécoise et le Tribunal des droits de la personne : les deux sont habilités au traitement des discriminations en matière d'emploi. Dans les deux ressorts, l'administration est soumise à leur juridiction de la même manière. La différence juridictionnelle notable entre le Tribunal Canadien et le Tribunal Québécois pourrait résider dans le fait que le premier prévoit une formation spécifique en matière de discrimination en contexte professionnel[65], alors que le second l'intègre à sa formation habituelle[66]. Aussi, il convient de s'attarder davantage sur la formation spécifique du Tribunal Canadien : le Tribunal sur l'équité en matière d'emploi.

Après saisine, le président du Tribunal Canadien désigne un membre qui constituera le Tribunal en matière d'emploi. Il s'agit d'une formation à juge unique par principe, la formation collégiale étant un choix laissé à l'appréciation du président du Tribunal en matière d'emploi en fonction de l'importance de l'affaire[67]. Par conséquent, ce tribunal ne constitue pas une juridiction à part entière, mais une formation du Tribunal Canadien. Tout comme les cas de discriminations hors contexte professionnel, la réception des plaintes et la première phase d'investigation se déroulent au niveau de la Commission canadienne. Si au cours de son

[65] « Tribunal de l'équité en matière d'emploi », *Tribunal Canadien de Droits de la Personne*, en ligne : <https://chrt-tcdp.gc.ca/about/employment-equity-review-tribunal-fr.html> (consulté le 23 mai 2020).

[66] *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 111 et 111.1; « Tribunal des droits de la personne », en ligne : <http://www.tribunaux.qc.ca/TDP/index-tdp.html> (consulté le 22 mai 2020).

[67] *Loi sur l'équité en matière d'emploi*, préc., note 63, art. 28 (2).

intervention la Commission canadienne rend une ordonnance, ou juge de l'opportunité d'une saisine du tribunal[68], alors ce dernier commence son office.

En effet, après avoir mené ou délégué son instruction du dossier, le tribunal rend une décision confirmant ou infirmant en tout ou partie l'ordre de la Commission canadienne. Au titre de son instruction, le tribunal en matière d'emploi dispose de pouvoirs comparables à ceux du Tribunal Canadien en formation ordinaire[69].

Enfin, sa décision implique également la capacité à prendre et à modifier toute mesure qu'il jugerait pertinente. S'agissant de la force exécutoire des ordonnances, à défaut d'homologation de la Cour fédérale, celles-ci s'exécutent au même titre que les ordonnances du Tribunal Canadien[70]. Les lois fédérales permettent donc un traitement adapté et particulier aux cas de discriminations en contexte professionnel[71]. De même, les institutions québécoises traitent de manière spécifique les atteintes au droit à l'égalité dans la sphère professionnelle.

c) Une protection double face aux discriminations en matière d'emploi

Ainsi, au regard de la particularité du traitement accordé aux discriminations en contexte professionnel, les justiciables confrontés à une IA de recrutement discriminatoire pourraient se tourner vers le tribunal en matière d'emploi au niveau national. De même ils pourraient s'orienter vers le Tribunal Québécois.

En conclusion, les Canadiens et les Québécois bénéficient d'une protection juridique sophistiquée en matière de discriminations en matière d'emploi. En effet, les engagements

[68] Il faudrait remarquer également que l'employeur qui conteste l'ordre de la commission peut également saisir le tribunal, pour une infirmation ou une révision de l'ordre rendu par la Commission canadienne.

[69] *Loi sur l'équité en matière d'emploi*, préc., note 63, art. 29.

[70] *Id.*, art. 31.

[71] *Id.*, art. 4 b) et suiv. À ce titre il conviendrait de rappeler qu'en vertu de ses articles 4 (b) et suivants, *la Loi sur l'équité en matière d'emploi* est opposable aux employeurs privés et aux secteurs publics des administrations fédérales. La compétence de la Commission canadienne et du Tribunal en matière d'emploi concerne donc aussi bien les situations discriminatoires lorsque l'employeur du secteur privé, que celles qui concernent un employeur du secteur public.

internationaux trouvent application aussi bien par le truchement des textes constituants, que par celui des lois fédérales et québécoises qui nourrissent le cadre légal du droit du travail. Cette particularité du contentieux des discriminations en matière d'emploi reflète une reconnaissance de l'existence et de la complexité des ruptures d'égalité en ce domaine.

Plus encore, la protection du droit à l'égalité est ici également double. Qu'est-ce à dire ? Il s'agit ici de remarquer que les lois fédérales et québécoises ont défini des obligations de moyens et de résultat à la charge de l'employeur. En effet, au-delà de la prohibition des comportements et mécanismes sciemment discriminatoires — qui sont à la fois des obligations de ne pas faire et des obligations de résultat —, et à l'instar des conventions internationales, les textes nationaux définissent des obligations positives de protection du droit à l'égalité, qui sont également des obligations de moyens[72]. Là où l'interdiction de comportements discriminatoires permet de protéger contre les ruptures d'égalité intentionnelles, l'obligation positive consiste pour l'employeur à poser des actes de nature à permettre l'équité et le recul des oppressions systémiques à l'œuvre dans son organisation. Il s'agit donc ici d'une obligation de mettre en place des mesures promouvant une égalité réelle, voire de l'équité. Ce paradigme de protection s'inscrit dans la continuité de celui établi pour les discriminations hors contexte professionnel.

Ainsi, lorsqu'on s'interroge sur l'incidence de l'IA de recrutement face à ce corpus de règles juridiques, on peut remarquer que si les différentes formes de discriminations apparaissent dans les instruments juridiques précités, les outils et moyens de discriminations, eux, ne sont pas visés. Par conséquent, il n'est pas fait référence à la technologie qui contribue à la production des discriminations, que ce soit en général, ou dans la sphère professionnelle. L'exclusion n'étant pas

[72] Stéphanie PAVAGEAU, « Les obligations positives dans les jurisprudences des cours européenne et interaméricaine des droits de l'homme », *International Law : Revista colombiana de derecho internacional* 2005.6.201-246; Samantha BESSON, « Les obligations positives de protection des droits fondamentaux », (2003) 1 *Revue de droit suisse* 49-96; *Loi Canadienne sur les droits de la personne*, préc., note 44, art. 16, 17 et 27; *Loi sur l'équité en matière d'emploi*, préc., note 63, art. 5 et 6; *Charte des Droits et Libertés de la Personne*, préc., note 28, art. 72 et 86; *Loi sur l'équité salariale*, (1996), RLRQ c E-12.001, art. 10 et 11, en ligne : <http://legisquebec.gouv.qc.ca/fr/ShowDoc/cs/E-12.001> (consulté le 21 mai 2020).

apparente, on pourrait conclure que les protections juridiques contre les discriminations s'étendent également aux candidats soumis à une IA de recrutement[73].

Suite aux différentes révélations de discriminations[74], on peut d'ailleurs comprendre les inquiétudes que suscitent celles relatives à l'embauche par IA de recrutement. Si ces dernières ne font pas l'objet d'une exclusion légale, il reste à étudier dans quelle mesure la décision algorithmique fait partie d'un mode légal de prise de décision.

B. La légalité des décisions algorithmiques

Dans le but de comprendre les règles juridiques applicables à la décision de l'IA de recrutement, nous nous interrogerons sur la valeur des décisions intervenant par IA au cours d'un processus de recrutement.

Aussi, avant de nous attarder sur l'étude de la réglementation fédérale et québécoise en matière de décision algorithmique (2), nous définirons ce terme en premier lieu (1).

1. Les contours de la notion de décision algorithmique

Plus que se prêter à un exercice de définition, il s'agira également de justifier le choix de ce terme parmi d'autres existant, avant d'en préciser la portée dans le cadre d'un processus de recrutement.

---

[73] « Speciala generalibus derogant » : cet adage juridique signifie que les dispositions spéciales dérogent aux règles générales. Dans le cas des discriminations par IA de recrutement aucun régime particulier n'est prévu par la loi. Par conséquent c'est la règle générale qui s'applique : ici ce sont les règles de prohibition et de sanction des discriminations en matière d'emploi.

[74] Rory CELLAN-JONES, « Amazon scrapped "sexist AI" tool », sect. Technology (10 octobre 2018), en ligne : <https://www.bbc.com/news/technology-45809919>; Sophie CURTIS, « Google Photos labels black people as "gorillas" », *The Telegraph*, sect. Technology (1 juillet 2015), en ligne : <https://www.telegraph.co.uk/technology/google/11710136/Google-Photos-assigns-gorilla-tag-to-photos-of-black-people.html>; Thierry NOISETTE, « Facebook retire 5.000 filtres publicitaires permettant des discriminations », *L'Obs*, sect. Les internets (22 août 2018), en ligne : <https://www.nouvelobs.com/les-internets/20180822.OBS1184/facebook-retire-5-000-filtres-publicitaires-permettant-des-discriminations.html> (consulté le 10 décembre 2018).

a) La définition de la décision algorithmique de recrutement

La décision consiste en un choix, un acte par lequel on tranche un problème[75]. Selon le contexte de son intervention, la décision répond à des catégories ou critères différents. En tout état de cause, la décision est nécessaire en présence d'un problème auquel plusieurs solutions répondent.

Dans le cadre d'une organisation, et notamment en matière de recrutement externe[76], la décision prend du temps parce qu'elle nécessite une analyse au préalable (sélectionner la meilleure candidature). Plus particulièrement, on pourrait identifier plusieurs étapes du processus de recrutement : la recherche de candidats, le tri des candidatures reçues, la retenue des profils les plus pertinents pour l'organisation, une nouvelle sélection aboutissant à l'embauche finale. Le recrutement signifie donc faire une série de choix et de décisions qui aboutissent à l'embauche des personnes qui seraient, dans l'idéal, les plus indiquées pour les postes offerts. Aussi, au nombre des tâches chronophages qui peuvent être déléguées à un algorithme d'IA, il y a la prise de décision. Parler de « décision algorithmique » permet donc de distinguer les décisions humaines des décisions qui sont déléguées à un algorithme. Ceci étant précisé, il convient de s'intéresser à la terminologie existante.

L'Union Européenne (ci-après « l'UE ») à travers son Règlement Général sur la Protection des Données (ci-après « RGPD ») emploie le terme de « décision automatisée » que l'on retrouve en France, notamment avec la Commission Nationale de l'Informatique et des Libertés[77] (ci-après « CNIL »). De plus, à travers les travaux du Service anti-discrimination du Conseil de l'Europe, on retrouve également la terminologie « décision algorithmique ». Celle-ci renvoie à toute décision

[75] OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, « décision », dans Office Québécois de la Langue Française, coll. Grand Dictionnaire Terminologique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8357768> (consulté le 27 avril 2020). Voir également « décision », dans Le Robert, coll. Dico en ligne, en ligne : <https://dictionnaire.lerobert.com/definition/decision> (consulté le 27 avril 2020).

[76] OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, préc., note 21.

[77] CNIL, « Profilage et décision entièrement automatisée », en ligne : <https://www.cnil.fr/fr/profilage-et-decision-entierement-automatisee> (consulté le 27 avril 2020).

prise par un algorithme ; en d'autres termes, il s'agit du résultat de l'algorithme[78]. Ici, la référence explicite à l'algorithme ou au caractère automatique de la décision ne permet pas de distinction nette entre les décisions qui sont automatiques en tout ou en partie.

Au Canada par ailleurs, le gouvernement fédéral semble privilégier le terme « système décisionnel automatisé »[79]. La principale différence entre ces deux expressions réside dans l'attachement au processus conduisant à la décision. En effet, dans la définition fédérale, ce sont les moyens techniques et les méthodes de choix délégués à la technologie qui déterminent les contours de la notion. Si ces deux termes visent les décisions prises ou appuyées par un algorithme d'IA, les propos développés privilégieront la définition fédérale.

Afin d'englober les décisions prises en tout, en partie ou simplement appuyées par un algorithme d'IA, le terme « décision algorithmique » sera utilisé dans une conception extensive rejoignant autant les travaux du Canada fédéral que ceux du Conseil de l'Europe.

b) Le type de processus visés, un champ d'application étendu

Une fois la décision algorithmique définie, quelques éclaircissements demeurent nécessaires. Dans la Directive sur la prise de décision automatisée[80] du Secrétariat du Conseil du Trésor du Canada (ci-après « Secrétariat du Conseil du Trésor »), on retrouve une définition visant plusieurs technologies. Ainsi, sont explicitement visés « les systèmes basés sur des règles, la régression, l'analytique prédictive, l'apprentissage automatique, l'apprentissage en profondeur et les réseaux neuronaux. » Les techniques et technologies concernées font donc l'objet d'une

[78] Frederik Zuiderveen BORGESIUS, *Discrimination, intelligence artificielle et décisions algorithmiques*, Strasbourg, Conseil de l'Europe, 2018, en ligne : <https://www.coe.int/fr/web/european-commission-against-racism-and-intolerance/-/-discrimination-artificial-intelligence-and-algorithmic-decision-making-> (consulté le 30 avril 2020).

[79] SECRÉTARIAT DU CONSEIL DU TRÉSOR DU CANADA, *Directive sur la prise de décision automatisée*, 5 février 2019, en ligne : <http://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=32592>.

[80] *Id.* La directive du Conseil du Trésor ne constitue pas une règle de droit *erga omnes*, mais concerne la gestion interne des administrations fédérales. L'intérêt ici est donc d'éclairer la notion de décision algorithmique et ses conséquences par le biais des travaux que le gouvernement fédéral a élaboré dans le cadre de sa gestion interne. Ainsi, même si cette directive ne consiste pas un acte juridique *a priori*, il sert néanmoins d'indicateur sur les orientations administratives sur ces enjeux.

énumération claire. Tout processus de décision recourant à ces procédés est ainsi compris dans le terme « décision algorithmique ».

Plus encore, il convient de relever que tant dans le vocabulaire européen, que dans le vocabulaire fédéral ou québécois, les processus de décisions visés englobent autant l'IA aide à la décision, que l'IA suffisamment autonome pour assurer le processus décisionnel, en tout ou partie. Concrètement, une IA de recrutement aide à la décision serait, par exemple, celle qui classe les candidatures de la plus pertinente à la moins pertinente[81]. Une IA de recrutement plus autonome interviendrait à plusieurs étapes du processus de recrutement : par exemple, plus que du classement, l'IA de recrutement trie les candidatures pour n'en retenir que 10 que le service des Ressources humaines prendra le temps de gérer sans technologie par la suite[82]. Enfin, il y aurait l'IA de recrutement qui serait totalement autonome, sans supervision humaine, de l'examen des candidatures jusqu'à la décision finale de recrutement. Il convient de remarquer ici que les IA de recrutement les plus répandues, aujourd'hui, répondent à la catégorie d'aide à la décision, soit la catégorie la moins autonome d'IA. En effet, à ce jour, le recours aux IA demeure majoritairement circonscrit à l'aide au tri des candidatures[83], ainsi qu'à la recherche de talents[84].

Par conséquent, on peut conclure que les autorités régulatrices du Canada et de l'UE ont une conception extensive des IA. Si aujourd'hui il ne semble pas y avoir d'IA entièrement autonomes, les inclure dans le champ d'application du cadre juridique en construction parait bienvenu[85].

---

[81] Jean Baptiste SU, « SmartRecruiters Unveils Artificial Intelligence Recruiting Assistant, Hiring Scorecard », *Forbes*, en ligne : <https://www.forbes.com/sites/jeanbaptiste/2018/03/14/smartrecruiters-unveils-artificial-intelligence-recruiting-assistant-hiring-scorecard/> (consulté le 7 août 2019); Q. PÉRINEL, préc., note 16.

[82] HIREVUE, « HireVue's Industry Solutions From Technical Hiring to Campus Recruiting », *HireVue*, en ligne : <https://www.hirevue.com/why-hirevue/solutions> (consulté le 1 juillet 2020); Drew HARWELL, « A face-scanning algorithm increasingly decides whether you deserve the job », *Washington Post*, sect. Technology (6 novembre 2019), en ligne : <https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/> (consulté le 1 juillet 2020).

[83] Notons ici que l'aide au tri de candidatures comprend parfois la tenue d'entretiens d'embauche par IA.

[84] R. DEMICHELIS, préc., note 16.

[85] Chris REYNOLDS, « Réglementation de l'intelligence artificielle: le Canada à la traîne », *La Presse*, sect. Techno (19 mai 2019), en ligne : <https://www.lapresse.ca/affaires/techno/201905/19/01-5226739-reglementation-de-lintelligence-artificielle-le-canada-a-la-traine.php> (consulté le 30 avril 2020); Brad SMITH, « Facial recognition

Par ailleurs, plus que la conception extensive, l'influence du RGPD dans les travaux et réflexions canadiens est manifeste dans la réglementation embryonnaire de l'IA. Un exercice comparatif permettrait une mise en perspective des orientations qui émergent au Canada aussi bien au niveau fédéral qu'au Québec.

2. La protection de la vie privée : une protection contre la décision automatique discriminatoire

Après avoir clarifié le contenu de la notion de décision algorithmique, il convient maintenant de nous attarder sur les règles juridiques qui encadrent les risques de discrimination découlant de ce type de décision. Parler de la régulation des utilisations de l'IA amène inévitablement à s'interroger sur les données traitées par celle-ci. Aussi, les IA de recrutement sont fonctionnelles parce qu'elles ont été entraînées sur des données extraites des candidatures, ou des profils de personnes candidates à des emplois. Il convient de s'interroger sur les données sensibles prises en compte par l'algorithme. En effet, selon la manière dont l'IA traite des données telles que la race, le genre, ou encore la situation de validité des individus candidats, les risques de discriminations fondées sur ces caractères sont plus ou moins importants.

Aussi, l'encadrement de la collecte et du traitement de données sensibles relatives aux personnes candidates à un emploi, pourrait constituer un premier moyen de limiter le risque de discrimination par IA de recrutement. En la matière, les législations relatives à la protection des données personnelles et à la vie privée constituent des supports non négligeables. Inspiré par le RGPD européen, le Canada érige en effet la protection des données personnelles en véritable rempart de protection des droits fondamentaux. Ce paradigme de protection par le droit à la vie privée s'observe à travers la notion extensive de traitement de données, mais également par la réglementation dédiée à la décision automatique.

technology: The need for public regulation and corporate responsibility », *Microsoft on the Issues*, en ligne : <https://blogs.microsoft.com/on-the-issues/2018/07/13/facial-recognition-technology-the-need-for-public-regulation-and-corporate-responsibility/>.

a) La notion de traitement des données étendue au profilage

Le RGPD, entré en application en 2018, constitue une référence qui inspire les législateurs du reste du monde en matière de protection des données personnelles. En effet, le RGPD s'applique ainsi « à la collecte de données pour la création de profils, ainsi qu'à l'application de ces profils aux particuliers »[86], mais également au profilage et aux décisions automatisées. Pour recourir à l'un de ces deux derniers modes de traitement, le gestionnaire des données s'engage à deux égards : le respect des garanties[87] et la légalité du traitement.

En la matière l'article 22 § 1 du RGPD semble poser le principe : le traitement des données personnelles par profilage ou décision automatique est en effet conditionné[88]. Cet article est en outre analysé par l'UE comme formulant « […] une interdiction générale de ce type de traitement afin de tenir compte des risques potentiels pour les droits et les libertés des personnes »[89]. Le traitement des données personnelles intègre donc le profilage. En effet, le RGPD définit le profilage comme constituant :

> « toute forme de traitement automatisé de données à caractère personnel consistant à utiliser ces données à caractère personnel pour évaluer certains aspects personnels relatifs à une personne physique, notamment pour analyser ou prédire des éléments concernant (…) cette personne physique »[90].

Trois caractères apparaissent dans cette définition : « [un] traitement automatisé (…) effectué sur des données à caractère personnel [dont l'objectif est] d'évaluer les aspects personnels d'une

[86] GROUPE DE TRAVAIL « ARTICLE 29 », *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, 17/FR WP251rev.01, Bruxelles, Commission Européenne, 2018, p. 6, en ligne : <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053> (consulté le 6 juillet 2018).

[87] Les garanties auxquelles il est fait référence sont celles de transparence, de licéité, de loyauté, d'information de la personne concernée lorsque les finalités de collecte et de traitement évoluent, la minimisation des données, et la limitation de la conservation de celles-ci.

[88] *Règlement Général sur la Protection des Données personnelles*, 88 (2016), 2016/679, art. « 22. Décision individuelle automatisée, y compris le profilage » [*RGPD*], en ligne : <https://eur-lex.europa.eu/eli/reg/2016/679/oj> (consulté le 20 février 2020).« 1. La personne concernée a le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire. »

[89] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, p. 9.

[90] *RGPD*, préc., note 88, art. 4 4).

personne physique »[91]. Par conséquent, le profilage constituerait une forme de « classification » fondée sur des informations personnelles des individus. C'est sa finalité, ou son utilisation qui détermine s'il s'agit — ou non — de profilage. La pratique du profilage permet donc, grâce aux données collectées, d'établir des corrélations, ou de faire ressortir des inférences.

À titre d'exemple, le ciblage publicitaire[92] de plateformes telles que Facebook[93] constitue une forme de profilage des utilisateurs. L'intérêt d'appliquer le RGPD à ce type de pratique réside dans la possibilité d'engager la responsabilité de la plateforme en cas d'utilisation discriminatoire. En effet, le profilage reposant sur la collecte de renseignements personnels, les risques de détournement de la finalité — ici, c'est le ciblage d'audience — peuvent conduire à des atteintes au droit à l'égalité. Notamment, dans le cas de Facebook, ce risque s'est réalisé en matière de discriminations à l'embauche fondée sur le genre. En effet, en septembre 2018, une plainte a été déposée contre Facebook et plusieurs employeurs devant la Commission américaine pour l'égalité des opportunités d'emploi. Cette plainte reposait alors sur la constatation d'une inégalité d'accès à des offres d'emploi entre les hommes et les femmes : les employeurs ont été accusés d'utiliser les filtres de Facebook pour réserver leurs offres d'emploi à des hommes[94]. Ce contentieux reflète les liens entre le profilage des plateformes et les risques de discriminations à l'embauche. Par son recours au profilage, Facebook permet d'identifier (à tort ou à raison) le genre de ses utilisateurs. C'est grâce à ses algorithmes que les clients de la plateforme (les entreprises) ont accès à un éventail de filtres leur permettant de cibler leur public. Or, lorsqu'il s'agit de publier des offres d'emploi, donner la possibilité aux employeurs de filtrer l'audience en fonction du genre constitue une forme de discrimination prohibée.

---

[91] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, par. « A. Profilage ».

[92] OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, « Ciblage », dans Grand dictionnaire terminologique, coll. Commerce, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8400654> (consulté le 31 mai 2020).

[93] « Toucher de nouvelles audiences », *Pages d'aide de Facebook Business*, en ligne : <https://fr-fr.facebook.com/business/help/717368264947302> (consulté le 31 mai 2020).

[94] « Ciblage publicitaire. Facebook est accusé de favoriser la discrimination à l'égard des femmes », *Courrier international* (19 septembre 2018), en ligne : <https://www.courrierinternational.com/revue-de-presse/ciblage-publicitaire-facebook-est-accuse-de-favoriser-la-discrimination-legard-des> (consulté le 10 décembre 2018); T. NOISETTE, préc., note 74.

La nécessité d'appliquer les garanties du RGPD au profilage permet donc d'étendre les protections du règlement aux personnes qui y sont soumises et, par extension, aux différentes utilisations qui pourraient découler dudit profilage. En l'occurrence si Facebook ne collectait pas d'informations sur le genre de ses utilisateurs, la plateforme ne serait pas en mesure de proposer ce type de filtre à ses entreprises clientes. Sans ce profilage détaillé de ses utilisateurs, Facebook n'aurait peut-être pas contribué à nuire aux opportunités d'embauche des femmes concernées[95].

Aussi, si le RGPD autorise les décisions entièrement automatisées et le profilage, il les encadre strictement. Ce principe apparaît notamment à l'article 22, et est réaffirmé dans les travaux du Groupe de Travail « Article 29 » (ci-après « GT29 ») sur le profilage et la prise de décision automatique[96]. Trois exceptions viennent encadrer le recours au profilage et à la prise de décision automatique : le consentement de la personne concernée, lorsque c'est « nécessaire à l'exécution du contrat », en présence d'une obligation légale[97] ou encore un « intérêt légitime »[98]. Cette dernière exception renvoie aux intérêts de celui qui collecte et traite les données par rapport à ses activités.

Toutefois, le RGPD rend obligatoire l'équilibre entre l'intérêt légitime du gestionnaire des données et les droits fondamentaux de la personne profilée. La protection des droits fondamentaux prévaut donc sur l'intérêt légitime du gestionnaire de données. La recherche d'équilibre s'effectue en fonction du degré de précision du profilage, et de ses conséquences sur l'individu concerné. On retrouve dans cette approche une forme d'obligation de moyen dont l'intensité varie en fonction du risque pour les droits de l'individu concerné. Dans le cas des offres d'emploi par le biais de Facebook, il aurait fallu rechercher l'équilibre entre le besoin des entreprises employeuses de cibler une audience pertinente, et la nécessité de préserver les

[95] T. NOISETTE, préc., note 74. Notons ici que le recrutement n'est pas le seul domaine qui a fait l'objet d'une utilisation discriminatoire des filtres de Facebook. Les reproches et plaintes répétées ont d'ailleurs conduit la multinationale à retirer plusieurs filtres de sa plateforme.

[96] *RGPD*, préc., note 88, art. 22.

[97] Au titre de l'obligation légale on retrouve notamment celles qui concernent le blanchiment d'argent, la lutte contre la fraude ou l'évasion fiscale, la sauvegarde d'intérêts vitaux, les missions de service public.

[98] *RGPD*, préc., note 88, Considérant 47.

utilisateurs de Facebook de toutes formes de discriminations (en l'occurrence fondée sur le genre). La recherche de cet équilibre aurait pu conduire par exemple à l'impossibilité de filtrer l'audience des offres d'emploi en fonction du genre, de la race, ou des origines ethnoculturelles des utilisateurs.

S'agissant de la perspective fédérale, dans la Directive du Secrétariat du Conseil au Trésor du Canada, sont visés les systèmes décisionnels automatisés, soit :

> « toute technologie qui soit informe ou remplace le jugement des décideurs humains. Ces systèmes (…) utilisent des techniques telles que les systèmes basés sur des règles, la régression, l'analytique prédictive, l'apprentissage automatique, l'apprentissage en profondeur et les réseaux neuronaux ».

S'il n'est pas directement fait mention de la technique de profilage, on pourrait interpréter cette définition large comme incluant le profilage. En effet, l'analyse comportementale constitue la base du profilage[99]. Utilisé aussi bien comme support d'une décision entièrement automatisée, que pour éclairer la décision, le profilage analyse des données, des habitudes, des usages d'individus pour en tirer des corrélations et des analyses prédictives qui permettent d'établir un profil des individus concernés.

Aussi, à travers la référence à l'information ou le remplacement du jugement humain, la Directive du Secrétariat du Conseil au Trésor du Canada semble intégrer également le profilage dans les systèmes de décisions automatisés qu'elle vise. On pourrait regretter que l'occasion d'y référer explicitement n'ait pas été saisie ; pour autant, l'étude des pistes d'améliorations des différentes lois canadiennes étant plus que jamais d'actualité[100], il conviendrait probablement de nuancer les critiques quant au manque de précision d'une directive alors que les législations sont en cours d'évolution. À ce propos, on pourrait relever que les orientations et recommandations

---

[99] COMMISSION D'ENRICHISSEMENT DE LA LANGUE FRANÇAISE, « profilage », dans Grand dictionnaire terminologique, coll. Psychologie sociale, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26538781> (consulté le 18 juin 2020).

[100] *Loi n°64*, préc., note 20; *Projet de loi C-11*, préc., note 20. Au moment où nous écrivons ces lignes, les recommandations du Commissariat à la Vie privée sont déjà publiées, de même que le projet de loi québécois de modernisation de la LPRPDE, dit « projet de loi 64 », a été récemment adopté.

des autorités indépendantes convergent vers de forts rapprochements avec le RGPD[101]. Plus spécifiquement, les propositions 1 et 3 du Commissariat à la protection de la Vie Privée du Canada[102] (ci-après « Commissariat à la vie privée ») réfèrent explicitement aux protections des articles 21 et 22 RGPD applicables au profilage. Si le Commissariat ne propose pas de définition du profilage, il recommande clairement de s'inspirer des dispositions européennes dans la conception de la protection fédérale.

Aussi, les propositions susvisées, si elles étaient suivies, entraîneraient deux conséquences : non seulement la définition de l'IA engloberait le cas du profilage, mais la législation fédérale ouvrirait également un droit d'opposition à ces modes de traitements[103]. Dans le cas des offres d'emplois publiées sur Facebook, un droit d'opposition signifierait que les utilisateurs de la plateforme pourraient choisir de désactiver le filtrage des publications selon leur genre. Ainsi, ceux-ci seraient déjà en capacité de se préserver des utilisations discriminatoires des filtres appliqués aux offres d'emploi.

---

[101] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, *Consultation sur les propositions du Commissariat visant à assurer une réglementation adéquate de l'intelligence artificielle*, coll. Consultations, Commissariat à la protection de la vie privée du Canada, 2018, en ligne : <https://www.priv.gc.ca/fr/a-propos-du-commissariat/ce-que-nous-faisons/consultations/consultations-terminees/consultation-ai/pos_ai_202001/> (consulté le 15 mars 2020).

[102] *Id. « Proposition 1 : Incorporer dans la loi une définition de l'IA qui servirait à distinguer les règles juridiques qui ne s'appliqueraient qu'à elle, tandis que les autres règles s'appliqueraient à tout type de traitement, y compris l'IA.* La [Loi sur la protection des renseignements personnels et les documents électroniques (LPRPDE)] est neutre sur le plan technologique et constitue une loi d'application générale. À ce titre, elle ne comprend pas de définition relative à l'IA, à la prise de décision automatisée ou au traitement automatisé. Toutefois, tel que nous le suggérons dans certaines autres propositions de ce document, il pourrait être nécessaire d'établir des règles particulières pour certaines utilisations de l'IA, ce qui justifierait de la définir dans la loi pour préciser quand ces règles s'appliqueraient. […] Le Règlement général sur la protection des données (RGPD) de l'Union européenne traite explicitement de l'IA en parlant de « décision individuelle automatisée » et de « profilage » à l'article 22 […]. » (nos italiques)

« Proposition 3 : La loi devrait prévoir le droit de s'opposer à la prise de décision automatisée et de ne pas être soumis à des décisions fondées uniquement sur un traitement automatisé, sous réserve de certaines exceptions. Si nous voulons protéger efficacement la vie privée en tant que droit de la personne dans un contexte numérique où des systèmes d'IA sont actifs, l'un des droits qui doit être envisagé est la capacité de s'opposer aux décisions prises par les ordinateurs et de demander une intervention humaine. […] L'article 21 confère également à la personne concernée le droit de s'opposer au traitement de ses données à caractère personnel à des fins de prospection directe (marketing), et tout profilage ou traitement connexe doit cesser dès la réception de la signification de l'opposition. Il n'existe pas d'exemptions ni de motifs permettant de refuser l'objection d'une personne à l'égard de la prospection directe. […] En cette matière, nous sommes en faveur de l'incorporation dans la LPRPDE d'un droit d'opposition limité, semblable à celui que l'on trouve dans le RGPD. » (nos italiques)

[103] Nous étudierons plus en détail le projet de loi fédéral C-11 et la *loi n°64* au chapitre 3

En outre, dans sa *loi n° 64*, le Québec définit le profilage comme constituant :

> « la collecte et de l'utilisation de renseignements personnels afin d'évaluer certaines caractéristiques d'une personne physique, notamment à des fins d'analyse (…) du comportement de cette personne »[104].

On peut remarquer ici un alignement du projet de loi sur la recommandation du mémoire de la Commission québécoise à la CAI (Commission d'Accès à l'Information du Québec)[105]. Par ailleurs, il convient de relever que dans ce mémoire, la Commission québécoise recommande l'interdiction de traiter certains renseignements personnels dans le cadre du profilage[106].

Enfin, si un droit à l'opposition n'existe aujourd'hui ni au Canada ni au Québec, la récente *Loi n° 64* permet de confirmer une inspiration européenne. Bien que le cadre légal soit émergent s'agissant du profilage, les autorités québécoise et fédérale de protection des données personnelles s'inscrivent dans la même démarche d'encadrement strict du profilage. En effet, dans la L*oi n° 64*, plus qu'une définition du profilage proche de celle du RGPD, on y retrouve

[104] *Loi n°64*, préc., note 20, art. 14 et 37. Le projet de loi 64 prévoit en effet l'insertion d'articles dédiés au profilage dans les législations de protection des renseignements personnels des secteurs public et privé. Le projet prévoit l'insertion d'un article 8.1 dans LPRPSP dont le second alinéa dispose « Le profilage s'entend de la collecte et de l'utilisation de renseignements personnels afin d'évaluer certaines caractéristiques d'une personne physique, notamment à des fins d'analyse du rendement au travail, de la situation économique, de la santé, des préférences personnelles, des intérêts ou du comportement de cette personne ». On retrouve la même formulation dans la loi relative au secteur public avec un projet d'article 65.0.1

[105] J.-F. TRUDEL, préc., note 27, p. 14. « La Commission est d'accord avec les principes énoncés au point 2 et visant l'encadrement des activités de profilage, d'analyse et de prédiction. Le document de consultation propose de s'inspirer de la définition de profilage prévue à l'article 4 du Règlement général sur la protection des données84, qui se lit comme suit : Aux fins du présent règlement, on entend par […] «profilage», toute forme de traitement automatisé de données à caractère personnel consistant à utiliser ces données à caractère personnel pour évaluer certains aspects personnels relatifs à une personne physique, notamment pour analyser ou prédire des éléments concernant le rendement au travail, la situation économique, la santé, les préférences personnelles, les intérêts, la fiabilité, le comportement, la localisation ou les déplacements de cette personne physique »

[106] *Id.* « Il est aussi proposé, au point 2.1, d'interdire l'utilisation de certains types de renseignements personnels afin d'effectuer du profilage. À cet effet, la Commission recommande que les 14 motifs prohibés de discrimination prévus à l'article 10 de la Charte fassent partie d'une telle liste ».

également un droit d'information et de refus des traceurs[107], pierres angulaires du profilage[108]. On retrouve un paradigme similaire s'agissant des données inférées.

b) Une extension de la protection aux inférences résultant du traitement

Plus que l'analyse comportementale subséquente au traçage, le traitement de données offre la capacité de révéler des informations, grâce aux recoupements. Il s'agit d'inférences[109]. Celles-ci peuvent apparaître au cours de traitement de données qu'il s'agisse ou non de profilage. Ces informations inférées, *a priori*, ne constituent pas des renseignements personnels puisque c'est par des recoupements qu'elles peuvent dévoiler des renseignements. Or ceux-ci peuvent constituer des renseignements sensibles. Les inférences constituent donc des données déduites des renseignements collectés. La question est de savoir comment appréhender ces données qu'on pourrait qualifier de secondaires.

Dans l'approche européenne, la lecture conjointe de ses articles 6 et 9 permet au RGPD d'inclure ces données inférées dans son champ d'application. Ainsi, par sa conception extensive, le RGPD protège les données produites par les inférences au même titre que des renseignements personnels par essence[110].

---

[107] Les traceurs, plus connus sous l'appellation « cookies », sont des programmes rattachés à des sites internet ou à des applications. Ces programmes permettent la collecte de données de différentes natures aussi bien sur l'utilisateur, que sur sa navigation. Ce sont les données collectées par le biais de ces cookies qui constituent la base du profilage.

[108] *Loi n°64*, préc., note 20, art. 8.1, 2°.

[109] OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, « inférence », dans Grand dictionnaire terminologique, coll. Informatique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=2074389> (consulté le 20 juin 2020).

[110] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, p. 16. « Le profilage peut engendrer des données d'une catégorie particulière par inférence à partir de données qui n'appartiennent pas à une catégorie particulière en soi, mais qui le deviennent lorsqu'elles sont combinées avec d'autres données. Par exemple, il peut être possible de déduire l'état de santé d'une personne à partir des historiques de ses achats d'aliments combinés à des données sur la qualité et la teneur énergétique des aliments.

Il est alors possible de découvrir des corrélations qui donnent des indications au sujet de la santé, des convictions politiques, des croyances religieuses ou de l'orientation sexuelle des individus […] ».

La protection de la vie privée d'une personne concernée par une décision automatique répond donc aux expositions qu'elle risque face à un traitement automatisé de cet ordre.

Au Canada, le Commissaire à la protection de la vie privée du Canada recommande également d'étendre la protection de la vie privée — et des renseignements personnels — aux données inférées[111].

L'intérêt de la protection contre l'utilisation de données inférées est d'offrir une garantie supplémentaire face à l'utilisation d'informations sensibles telles que la race, le genre, la religion, l'origine ethnique. En effet, par le biais des traceurs, les algorithmes de profilage permettent de déduire des informations sensibles. Ainsi, une plateforme telle que celle de Facebook a déjà été mise à l'index à propos des préjudices découlant de ses algorithmes de profilage[112].

En conclusion, par leur encadrement du profilage et des données qui l'appuient, les législations relatives à la vie privée et à la protection des données personnelles apportent un premier niveau de protection contre les discriminations. On pourrait qualifier ce premier niveau d'encadrement de régulation à la source. Il s'agit de réguler l'accès aux données sensibles afin qu'elles ne génèrent pas des profilages — et par extension, des décisions algorithmiques — discriminatoires.

Dans un contexte de recrutement, les responsables du traitement seront donc amenés à faire preuve de vigilance quant aux IA utilisées afin de faciliter le processus de tri des candidatures.

---

[111] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, *Un cadre réglementaire pour l'IA : recommandations pour la réforme de la LPRPDE*, Gatineau, Commissariat à la protection de la vie privée du Canada, 2020, en ligne : <https://priv.gc.ca/fr/a-propos-du-commissariat/ce-que-nous-faisons/consultations/consultations-terminees/consultation-ai/reg-fw_202011/> (consulté le 17 novembre 2020). « Une autre mesure importante de cette approche fondée sur les droits de la personne consisterait à modifier la définition des renseignements personnels dans la LPRPDE afin d'inclure les inférences au sujet d'une personne. […] Faire en sorte que les renseignements [inférés] soient clairement visés par la loi est essentiel à la protection des droits de la personne, car il est souvent possible de faire des inférences sur un individu à son insu et de les utiliser pour prendre des décisions à son sujet. »

[112] K. GRANVILLE, préc., note 2. C'est notamment le cas avec l'affaire Cambridge Analytica qui a influencé les électeurs américains en 2016. Grâce aux informations collectées par le biais de tests de personnalité fantaisie publiés sur Facebook, la firme a permis le profilage de milliers d'électeurs dont les opinions politiques ont pu être identifiées.

c) La modulation des garanties dans l'encadrement de la prise de décision automatique

La prise de décision automatique « est la capacité de prendre des décisions par des moyens technologiques sans intervention humaine[113] ». À ce titre, les IA utilisées dans les processus de recrutement produisent des décisions pour automatiser certaines étapes du processus[114]. Pour autant, on peut distinguer deux niveaux d'automatisation dans les IA de tri de candidatures. En effet, là où certaines classent les CV et profils reçus, d'autres vont jusqu'à l'organisation d'entrevues avec les personnes candidates. Dans cette dernière hypothèse, c'est sur la base de l'analyse de l'IA qu'une décision officielle est communiquée aux personnes candidates[115].

Par conséquent, dans les IA de tri de candidatures, on peut différencier les décisions de rejet de candidatures entièrement automatisées, de celles qui sont semi-automatisées. Dès lors, il convient de s'intéresser aux implications légales de ces deux recours à l'IA. Nous reprendrons cette distinction entre décision semi-automatique et décision automatique afin de mettre en exergue les similitudes entre régimes canadiens[116] et européen.

Au niveau européen, le RGPD établit une distinction explicite entre les principes s'appliquant à une décision semi-automatique, et ceux qui s'appliquent en cas de décision entièrement automatique. Selon le niveau d'autonomie de la technologie utilisée, les garanties sont donc plus ou moins importantes[117]. On peut lire en effet dans les Lignes directrices du Groupe de Travail « Article 29 » (GT29) que « Des garanties et restrictions supplémentaires s'appliquent dans le cas d'une prise de décision exclusivement automatisée, y compris le

[113] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, par. « B. Prise de décision automatisée ».

[114] Rappelons à cet effet, qu'aujourd'hui les IA de recrutement ne concernent que l'automatisation de certaines étapes, mais ne permettent une délégation totale du processus de recrutement à la technologie. En d'autres termes, les IA permettent d'affiner la décision de recrutement finale (par l'élimination progressive des candidatures), mais ne procèdent pas à une sélection du début à la fin de recrutement.

[115] D. HARWELL, préc., note 82; Q. PÉRINEL, préc., note 16; J. B. SU, préc., note 81.

[116] En cohérence avec l'ensemble des propos développés, seules les perspectives québécoise et fédérale seront étudiées.

[117] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, p. 9-16.

profilage, visée à l'article 22, paragraphe 1 ». Autrement dit, d'une part le profilage entre dans la notion de décision automatisée, et d'autre part, en présence de telles automatisations de la prise de décision, le GT29 tire de l'article 22 du RGPD des exigences et des limites protectrices additionnelles. Plus encore, le GT29 opère une distinction explicite entre les décisions exclusivement automatiques, et les décisions semi-automatisées.

Les garanties offertes par le RGPD interviennent à plusieurs niveaux. Les protections accrues auxquelles nous faisons référence se manifestent à travers les exigences qui s'imposent au responsable du traitement. La protection générale encadrant la décision semi-automatique s'organise en effet autour de trois exigences : le respect des 6 principes de protection des données, l'existence d'une base légale du traitement et le respect des droits de la personne concernée[118].

Premièrement, s'agissant des principes de protection des données, l'article 5 RGPD permet de citer dans en premier lieu[119], les principes de licéité, de loyauté et de transparence, la limitation des finalités[120], la minimisation des données, l'exactitude de celles-ci et la limitation de leur conservation dans le temps. Dans le cadre d'un processus de recrutement, les quatre premiers principes se traduiraient respectivement par :

- La légalité d'un processus de sélection lors du recrutement,
- L'information des candidats quant à la technologie utilisée[121] afin qu'ils puissent s'y opposer ou y consentir[122],

[118] *Id.*

[119] Nous verrons en second lieu la base légale du traitement infra.

[120] Ici, il s'agit de traiter des finalités qui apparaitraient après que le consentement ait été obtenu. Les nouvelles finalités pourraient être acceptées à plusieurs conditions : « le lien entre les finalités pour lesquelles les données ont été collectées et les finalités du traitement ultérieur; le contexte [de collecte] et les attentes raisonnables (…) quant à leur utilisation ultérieure ; la nature des données; l'impact du traitement ultérieur sur les personnes concernées; et les garanties appliquées par le responsable du traitement afin d'assurer un traitement loyal et d'éviter tout impact indu sur les personnes concernées. » *RGPD*, préc., note 88, art. 5.

[121] *Id.*, art. 13 et 14.

[122] *Id.*, art. 21.

- L'information quant aux critères de sélection et, pour finir, la collecte et le traitement des données dans le cadre exclusif du recrutement[123].

Quant aux derniers principes, il s'agirait d'une part, d'un accès aux données[124] notamment pour les corriger et les actualiser[125], et d'autre part, de la mise en œuvre de processus de suppression des données[126].

En guise d'illustration de ces 6 principes, on pourrait prendre l'exemple d'un système de gestion des données répandu dans les entreprises : l'espace candidat. Sur les sites internet de plusieurs entreprises, les personnes candidates à des emplois peuvent se créer un profil virtuel de sorte qu'elles entrent leurs informations professionnelles dans la base de données de l'entreprise. Ainsi, une personne candidate crée son profil, actualise ses données professionnelles, parcourt l'ensemble des offres d'emplois émises par l'entreprise ; elle peut également suivre le traitement de sa candidature, notamment avec un suivi des étapes, des décisions de refus potentielles, ou encore des alertes pour des offres correspondant à son profil. Tout cela est géré de manière autonome par l'individu. Notons par ailleurs que dans un « espace candidat » les profils comportent des rubriques similaires aux CV (formations, diplômes, expériences professionnelles ou bénévoles antérieures, activités de loisirs…). Avec ce procédé, l'entreprise laisse la personne candidate contrôler[127] les données qui serviront à évaluer son profil au cours du processus de recrutement.

En second lieu, outre les 6 principes évoqués précédemment, la base légale de traitement constitue l'élément permettant une modulation du principe d'interdiction des traitements automatisés et des décisions automatiques : « Le consentement explicite est l'une des exceptions à l'interdiction de la prise de décision et du profilage automatisés définis à l'article 22,

[123] *Id.*, art. 18.

[124] *Id.*, art. 15.

[125] *Id.*, art. 16.

[126] *Id.*, art. 17.

[127] Le contrôle des données partagées par l'individu est partiel : certaines informations comme celles du CV sont souvent obligatoires.

paragraphe 1. »[128] En sus du consentement, l'article 6 du RGPD identifie cinq situations de nécessité qui permettent de faire exception à ce principe d'interdiction : l'exécution d'un contrat, le respect d'une obligation légale, la sauvegarde d'intérêts vitaux, l'exécution d'une mission de service public, ou l'exercice de l'autorité publique, l'intérêt légitime du responsable du traitement ou d'un tiers[129].

Enfin, le respect des droits de la personne concernée constitue la troisième exigence. Cette dernière exigence consiste à s'assurer que le traitement automatisé ou la décision semi-automatique demeure compatible et conforme aux droits de la personne qui y sera soumise.

Le cumul de ces trois catégories d'exigences — le respect des principes, l'existence d'une base légale et le respect des droits de la personne visée — constitue les garanties générales du RGPD en présence d'un traitement automatisé ou d'une décision semi-automatique. Ainsi, face à un processus semi-automatique de tri de candidatures, l'intérêt légitime du responsable de traitement pourrait justifier le recours à des algorithmes d'IA de classement. Aussi, le candidat doit partager un certain nombre de données avec l'employeur potentiel pour appuyer la décision semi-automatique de rejet de sa candidature. Ce dernier est donc soumis aux exigences susmentionnées.

Au Canada, en sus des 6 principes précités[130], le Commissariat à la vie privée recommande la création d'une obligation de tracer les données et les algorithmes, notamment en présence de décision automatique, et ce, tout au long du cycle de vie du système d'IA[131]. La proposition

---

[128] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, p. 14 par.2.

[129] *Id.*, p. 14-16.

[130] Soit les principes de licéité, de loyauté et de transparence, de limitation des finalités, de minimisation des données, d'exactitude des données et de limitation de leur conservation dans le temps.

[131] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101, Proposition 9.« *Exiger des organisations qu'elles assurent la traçabilité des données et des algorithmes, notamment en ce qui concerne les ensembles de données, les processus et les décisions prises pendant le cycle de vie du système d'IA.* Une exigence de traçabilité des algorithmes faciliterait l'application de plusieurs principes, notamment la responsabilité, l'exactitude, la transparence, la minimisation des données ainsi que l'accès et la correction. […] Compte tenu de ces points de vue d'experts et de l'importance de pouvoir retracer, analyser et valider les résultats du système d'IA pour que les personnes puissent se prévaloir des droits d'accès et de correction existants et pour améliorer la protection des droits de la personne en vertu de la LPRPDE modifiée, nous recommandons l'inclusion d'une exigence en matière de traçabilité des algorithmes des systèmes d'IA. » (nos italiques)

poursuit plusieurs objectifs. Dans un premier temps, le principe de traçabilité constituerait un rempart de protection d'autres principes : être en mesure de tracer les données et les algorithmes assurerait en effet le contrôle des principes de minimisation, de responsabilité, d'exactitude et de transparence. De facto, ces derniers principes seraient mieux garantis et contrôlés avec une traçabilité des données.

De plus, il s'agit de documenter tout le cycle de vie des données liées à l'algorithme, y compris l'origine des données et leur étiquetage[132]. Par conséquent, appliqué au processus de recrutement passant par les plateformes de Facebook ou Google, le principe de traçabilité impliquerait de savoir en premier lieu, quelles sont les données utilisées pour profiler/analyser les candidatures des utilisateurs, et en second lieu, quelles sont les étiquettes des données sensibles. De sorte que, si les plateformes utilisent des données comme la race, le genre, l'origine ethnique, l'orientation sexuelle, pour nourrir les IA de recrutement, le principe de transparence permettrait de le savoir. Plus encore, une traçabilité des données adéquate permettrait aussi de savoir qu'elles sont les étiquettes apposées en fonction du genre, de la race, des origines des candidats. Ce principe serait donc un instrument utile pour identifier les biais discriminatoires des IA de recrutement.

> « Les processus de préparation des données et d'"étiquetage" des données devraient être traçables. Autrement dit, il devrait être possible de montrer une piste de vérification de tout ce qui est arrivé aux données au fil du temps, au cas où il y aurait une vérification ou une enquête ultérieure. »[133]

In fine, le principe de traçabilité des données permettrait un meilleur contrôle des décisions semi-automatiques et autres formes de traitements automatisés dans les processus de recrutement. Remonter le fil des données utilisées — et du traitement appliqué — ouvre la voie à un plus grand contrôle. En d'autres termes, lorsque la décision de l'IA de recrutement écarte une candidature, le principe de traçabilité permettrait de vérifier la façon dont ont été traitées les données de la

[132] Par « origine des données », il est fait référence à la source des données, puis à leur classement dans la base de données aussi appelé étiquetage. Par exemple, dans une base de données de reconnaissance faciale fonctionnant sur les animaux, d'une part, la source des données X pourrait être un catalogue animalier, et d'autre part l'étiquetage constitue l'opération par laquelle les différentes photos sont classées en leur attribuant le marqueur « chien », « chat », ou encore « ours ».

[133] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101, Proposition 9.

personne candidate, selon quels filtres, quelles étiquettes. Plus encore, une telle capacité offrirait probablement une opportunité d'améliorer la qualité des données de sortie des algorithmes[134].

Aussi, connaître le traitement précis et l'interprétation des données permet de déceler des erreurs et biais potentiels de l'IA. Par conséquent, en cas de contestation du rejet d'une candidature, le principe de traçabilité met à la disposition de la personne candidate, ou de la Commission compétente[135], l'ensemble des informations nécessaires au contrôle de l'équité de la décision.

> « Assurer "l'exactitude et l'intégrité des renseignements sur une personne tout au long de la chaîne de possession en exigeant des organisations qu'elles informent toute autre organisation à qui ils ont été communiqués de leur modification ou de leur suppression" ».[136]

Bien que le GT29 fasse émerger une distinction entre les garanties dévolues aux décisions semi-automatisées, et celles relatives aux décisions entièrement automatisées, il convient de remarquer que l'on y retrouve fondamentalement les mêmes droits et exigences[137]. La particularité des exigences tiendrait davantage à leur intensité et à la responsabilité subséquente des responsables de traitement. Notons néanmoins que l'article 22 du RGPD semble insister sur la nécessité irrémédiable d'un contrôle humain. Par cette supervision humaine, le RGPD assurerait que la personne visée par une décision entièrement automatisée soit préalablement informée en des termes intelligibles et clairs, puis qu'elle dispose d'une explication de la décision et des motifs y ayant conduit. De même, il conviendrait de remarquer que l'article 35 du RGPD

[134] Voir infra *Chapitre 2 A*. *La base de données de l'IA de recrutement un vecteur de transmission des biais humains à la machine : le cas Amazon*.

[135] Commission canadienne, ou Commission québécoise.

[136] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101, Proposition 9.

[137] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, part. « IV. Dispositions spécifiques concernant la prise de décision exclusivement automatisée définie à l'article 22 ».

exige une évaluation de l'impact des décisions entièrement automatisées lorsque celles-ci présentent un risque fort d'atteinte à la vie privée ou aux droits fondamentaux[138] :

> « 1. Lorsqu'un type de traitement, en particulier par le recours de nouvelles technologies, et compte tenu de la nature, de la porte, du contexte et des finalités du traitement, est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes physiques, le responsable du traitement effectue, avant le traitement, une analyse de l'impact des opérations de traitement envisagées sur la protection des données à caractère personnel. […]»[139]

Enfin, la personne faisant l'objet d'une décision automatique doit conserver un droit de la contester.

Si contrairement à l'UE, le Canada ne s'est pas encore doté d'une législation claire en matière d'IA, la Directive sur la prise de décision automatisée du Secrétariat du Conseil du Trésor[140] constitue le texte de référence au niveau fédéral sur la décision automatique dans la sphère administrative. Aussi, bien qu'il n'existe pas de distinction nette entre les exigences relatives aux décisions semi-automatiques et celles relatives aux décisions automatiques, les prolégomènes légaux fédéraux et québécois permettent néanmoins l'identification d'un paradigme de gradation des exigences, similaire à celui retrouvé dans le RGPD.

Dans cette directive, c'est l'évaluation de l'impact algorithmique qui ressort comme condition sine qua non de la décision automatique[141]. Il s'agit en effet d'une grille d'évaluation permettant d'accompagner le respect des principes éthiques au stade de la conception, du développement et de l'utilisation d'une IA permettant la prise de décision automatique[142]. Par cette directive, l'administration fédérale instaure donc une gradation des exigences requises : c'est en fonction du niveau de risque d'atteinte aux droits fondamentaux que représente une IA, que les exigences

---

[138] F. Z. BORGESIUS, préc., note 78.

[139] *RGPD*, préc., note 88, art. 35 (1).

[140] *Directive sur la prise de décision automatisée*, préc., note 79.

[141] *L'Outil d'évaluation de l'impact algorithmique*, en ligne : <https://www.youtube.com/watch?v=X7QQyDpBS8Q> (consulté le 7 mars 2020).

[142] *Id.*; SECRÉTARIAT DU CONSEIL DU TRÉSOR DU CANADA, « Évaluation de l'incidence algorithmique (EIA) », *aem* (31 mai 2019), en ligne : <https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/technologiques-modernes-nouveaux/utilisation-responsable-ai/evaluation-incidence-algorithmique.html> (consulté le 7 mars 2020).

seront définies. Par conséquent, la logique consistant à prévenir les risques induits par les biais algorithmiques prime au niveau fédéral. Dans le secteur privé et, en l'absence de législation dédiée, les recommandations du Commissariat à la vie privée et la *Loi n° 64* s'inscrivent dans la même logique que la directive précitée : l'obligation d'Évaluation des Facteurs relatifs à la Vie Privée (ci-après « EFVP ») émerge comme une condition incontournable de recours aux décisions automatiques.

En outre, les recommandations des autorités régulatrices en matière de vie privée indiquent la stratégie de régulation souhaitée aux niveaux fédéral et québécois[143]. On observe à ce titre une volonté du Commissariat à la vie privée d'ériger la vie privée au rang des droits fondamentaux. L'objectif est d'en faire la clé de voute de la protection des autres droits fondamentaux, notamment le droit à l'égalité[144].

De même, en présence d'une décision automatique, le Commissariat à la vie privée rejoint la stratégie européenne en recommandant la garantie d'un droit d'opposition à la décision automatique[145]. Plus encore, l'autorité administrative indépendante fédérale propose qu'un droit

---

[143] Daniel THERRIEN, « Modernizing federal privacy laws to better protect Canadians », *Commissariat à la Protection de la Vie Privée du Canada* (1 juin 2020), en ligne : <https://priv.gc.ca/en/opc-news/speeches/2020/sp-d_20200309/> (consulté le 18 juin 2020); Simon DU PERRON, « Projet de loi 64 : une réforme à l'Européenne du droit à la protection des renseignements personnels », *Laboratoire de cyberjustice* (17 juin 2020), en ligne : <https://www.cyberjustice.ca/2020/06/17/projet-de-loi-64-une-reforme-a-leuropeenne-du-droit-a-la-protection-des-renseignements-personnels/> (consulté le 18 juin 2020); Commissariat à la protection de la vie privée du CANADA, « Nos attentes : Guide du Commissariat au sujet du processus d'évaluation des facteurs relatifs à la vie privée » (13 octobre 2011), en ligne : <https://priv.gc.ca/fr/sujets-lies-a-la-protection-de-la-vie-privee/evaluations-des-facteurs-relatifs-a-la-vie-privee/gd_exp_202003/> (consulté le 18 juin 2020).

[144] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101, Proposition 2. « *Adopter une approche fondée sur les droits dans la loi, selon laquelle les principes de protection des données sont mis en œuvre comme moyen de protéger un droit plus général à la vie privée – reconnu comme un droit fondamental de la personne et comme fondement de l'exercice d'autres droits de la personne.* […] Le but de la loi devrait être de protéger la vie privée au sens le plus large du terme, compris comme un droit de la personne en soi et comme fondement de l'exercice des autres droits de la personne. […] Afin d'assurer la protection des droits, nous sommes d'avis que la LPRPDE devrait être fondée sur les droits de façon à reconnaître la protection de la vie privée dans toute son ampleur et sa portée, et fournir une orientation sur la manière dont les autres dispositions de la Loi devraient être interprétées. » (nos italiques)

[145] *Id.* Proposition 3. *« La loi devrait prévoir le droit de s'opposer à la prise de décision automatisée et de ne pas être soumis à des décisions fondées uniquement sur un traitement automatisé, sous réserve de certaines exceptions***.** Si nous voulons protéger efficacement la vie privée en tant que droit de la personne dans un contexte numérique où des systèmes d'IA sont actifs, l'un des droits qui doit être envisagé est la capacité de s'opposer aux décisions prises par les ordinateurs et de demander une intervention humaine. Les lois de plusieurs pays prévoient le droit de ne pas être soumis à la prise de décision automatisée, ou un droit analogue de contester le traitement automatisé des

à l'explication soit explicitement reconnu aux personnes concernées par la décision automatique[146]. Ici, il convient de remarquer que si l'existence d'un droit européen à l'explication fait débat, le Commissariat à la vie privée prend le parti d'éviter cet écueil en consacrant une recommandation claire à ce sujet. Ramenée à l'IA de recrutement, que signifierait la garantie de droits d'opposition à la décision automatisée et de son explication ? D'abord, cela implique que la personne candidate peut s'opposer à ce que son profil soit examiné par le biais d'une IA. Ensuite, cela induit également que si la personne candidate a accepté le processus de recrutement par IA, et qu'elle est écartée, demander l'explication de la décision algorithmique lui sera possible. Par conséquent, ces garanties permettent aux candidats évincés d'obtenir une justification de la décision de l'algorithme utilisé. Il s'agit donc là d'un moyen de détecter et de contrôler l'existence de discriminations.

De plus, le Commissaire assortit le droit à l'explication de deux autres exigences. En sus de l'exécution de l'Évaluation des facteurs relatifs à la vie privée (EFVP), le Commissaire recommande son dépôt obligatoire. Cette dernière exigence complémentaire au droit à l'explicabilité a vocation à donner un caractère contraignant aux recommandations de la proposition 4[147] :

> « En outre, nous serions potentiellement en faveur d'un renforcement des exigences de transparence prévues par la loi de façon à mandater :
>
> La réalisation et la publication d'évaluations des facteurs relatifs à la vie privée (EFVP), y compris des évaluations relatives aux répercussions du traitement par IA sur la vie privée et les droits de la personne. […]. Des dépôts publics d'algorithmes […] assortis de sanctions pour non-divulgation et non-conformité. »

---

données personnelles, ainsi qu'un droit de ne pas être soumis à des décisions fondées uniquement sur l'automatisation. » (nos italiques)

[146] *Id.* Proposition 4.« *Donner aux personnes le droit à une explication et à une plus grande transparence lorsqu'elles interagissent avec un traitement automatisé ou font l'objet d'un tel traitement***.** […] Nous croyons que le principe de transparence de la LPRPDE devrait comprendre un droit d'explication qui fournirait aux personnes qui interagissent avec les systèmes d'IA le raisonnement qui sous-tend tout traitement automatisé de leurs données et les conséquences de ce raisonnement pour leurs droits et intérêts. Cela contribuerait également à satisfaire aux obligations actuelles de la LPRPDE qui consistent à donner aux personnes le droit d'accéder aux renseignements les concernant détenus par les organisations et de les corriger. » (nos italiques)

[147] *Id.*

Les acteurs privés seraient ainsi contraints de déposer leurs EFVP sur leurs algorithmes de décisions automatiques, sous peine de sanction pour « non-divulgation et non-conformité ». Cette exposition à la sanction est particulièrement intéressante, dans la mesure où, s'inspirant de la directive du Trésor susmentionnée, la proposition du Commissariat à la vie privée semble exiger l'EFVP dans les mêmes conditions que le Secrétariat au Trésor. Autrement dit, l'EFVP devra être réalisée non seulement aux stades de la conception et du développement, mais également à chaque mise à jour ou modification de l'algorithme.

En conclusion, les perspectives d'encadrement des décisions automatiques, au Québec et au niveau fédéral, se rapprochent étroitement de celles de l'Europe. Ces stratégies similaires s'observent également lorsque l'on s'attarde spécifiquement à l'IA de recrutement.

En l'absence de règles dédiées à l'IA, et par voie de conséquence à l'IA de recrutement, il convient de se tourner vers les travaux émergents.

Dans son livre blanc sur l'IA, la Commission Européenne expose sa stratégie de réglementation de l'IA[148]. L'approche retenue par l'UE repose sur la gestion des risques. Il s'agit donc de construire une législation dont la pierre angulaire est le risque que représente l'IA pour les droits fondamentaux ou pour l'intégrité physique des personnes. En effet, la Commission européenne définit une échelle des risques à travers une distinction entre les IA à haut risque et les IA présentant des risques modérés. Deux critères permettraient l'identification des IA à haut risque : le secteur concerné et l'utilisation faite dans le secteur à risque[149]. Selon la Commission :

> « […] compte tenu de son importance pour les particuliers et de l'acquis de l'UE sur l'égalité en matière d'emploi, l'utilisation d'applications d'IA dans les procédures de recrutement et dans des situations ayant une incidence sur les droits des travailleurs serait toujours considérée comme étant "à haut risque" ».

[148] Catelijne MULLER, *Intelligence artificielle - Une approche européenne axée sur l'excellence et la confiance*, Livre Blanc, INT/894-EESC-2020, 2020, en ligne : <https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/livre-blanc-sur-lintelligence-artificielle> (consulté le 23 juin 2020).

[149] *Id.*, p. 20.

En d'autres termes, en vertu de son potentiel discriminatoire, l'IA de recrutement présenterait de hauts risques par essence.

Ainsi, la tendance européenne consiste à aligner le niveau d'exigences sur le niveau de risque des IA : celles à haut risque, telles que les IA de recrutement, appellent un niveau de protection élevé, là où des IA de faible risque impliquent des garanties moins lourdes. L'analogie serait permise avec la logique de la gestion des risques que l'on retrouve en matière de protection de la vie privée. Force est de constater que le Canada s'inscrit dans le même courant. De même que pour l'UE, au Canada les orientations stratégiques constituent les seules ressources sur lesquelles on peut se fonder en matière d'IA.

En premier lieu, on peut noter une reconnaissance explicite du risque accru de discrimination en présence d'IA. En effet, aussi bien à travers les publications que par des prises de position, le Commissariat à la vie privée[150] et les organismes de lutte contre les discriminations[151] ont exprimé leurs inquiétudes, et leurs recommandations, pour appeler le législateur fédéral à se saisir de cet enjeu. La discrimination par IA de recrutement, si elle ne fait pas l'objet de recommandations spécifiques, est néanmoins intégrée dans l'appréhension des risques de l'IA au regard des droits fondamentaux. C'est effectivement dans plusieurs publications du Commissariat à la vie privée que l'on retrouve ses préoccupations et recommandations[152].

En second lieu, face à l'exigence de recueil du consentement, le Commissariat à la vie privée reconnaît, malgré tout, la possibilité d'exceptions au consentement sous certaines conditions « subordonnées à des pouvoirs d'application renforcés [qu'autoriseraient le

[150] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101; MINISTÈRE DE LA JUSTICE, « Modernisation de la Loi sur la protection des renseignements personnels du Canada », *Site Internet du ministère de la Justice Canada* (4 juin 2020), en ligne : <https://www.justice.gc.ca/fra/sjc-csj/lprp-pa/modern.html#s2> (consulté le 18 juin 2020); D. THERRIEN, préc., note 143; COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, « Pour une législation efficace sur la protection des renseignements personnels et l'accès à l'information dans une société guidée par les données » (6 novembre 2019), en ligne : <https://www.priv.gc.ca/fr/a-propos-du-commissariat/ce-que-nous-faisons/collaboration-avec-les-provinces-et-les-territoires/resolutions-conjointes-avec-les-provinces-et-territoires/res_191001/> (consulté le 30 juin 2020).

[151] M. WHITTAKER et al., préc., note 18; J.-F. TRUDEL, préc., note 27.

[152] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101; COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 150; D. THERRIEN, préc., note 143.

Commissariat lui-même] »[153]. À ce propos, contrairement au RGPD[154], le Commissariat à la vie privée ne semble pas considérer aussi précisément les différentes situations dans lesquelles le consentement des individus ne serait pas libre. En effet, si les organismes utilisateurs d'IA doivent recueillir le consentement des personnes concernées, il convient de s'interroger sur les contextes dans lesquels le consentement serait vicié. Or, en matière de recrutement, on pourrait questionner la liberté réelle des personnes candidates à refuser une IA de recrutement utilisée par leur potentiel employeur[155].

D'une part, pour la personne candidate consentir à l'IA de recrutement signifierait qu'elle accepte sa soumission à un procédé de tri potentiellement discriminatoire. Et bien que le tri de candidatures humain puisse également faire l'objet de discriminations, la particularité du recours à l'IA réside dans une plus grande difficulté à identifier la discrimination, ou ses causes.[156] En effet, là où une personne employeuse peut mettre en place des procédés assurant la neutralité du recrutement, l'IA en revanche repose souvent sur des algorithmes dont le fonctionnement

---

[153] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101, Propositions 7 et 11.« *Proposition 7 : Inclure dans la loi d'autres motifs de traitement et des solutions pour protéger la vie privée lorsqu'il n'est pas possible d'obtenir un consentement valable*. […] Une nouvelle exception au consentement de cette nature devrait nécessairement être subordonnée à des pouvoirs d'application renforcés qui autoriseraient l'organisme de réglementation en matière de la protection de la vie privée, le cas échéant, à déterminer si l'utilisation des renseignements personnels est effectivement faite à des fins sociales plus générales et si elle respecte les conditions juridiques prescrites. » (nos italiques)

« *Proposition 11 : Donner au Commissariat le pouvoir d'émettre des ordonnances exécutoires et d'imposer des sanctions financières aux organisations qui ne se conforment pas à la loi.* […] Dans d'autres administrations au Canada et à l'étranger, […] L'éventail des pouvoirs de rendre des ordonnances comprend la capacité d'exiger d'une organisation qu'elle cesse de recueillir, d'utiliser ou de communiquer des renseignements personnels, de détruire les renseignements personnels recueillis en contravention de la loi et, plus généralement, d'ordonner l'application des mesures correctives adéquates pour assurer la protection des renseignements personnels, notamment. […] De véritables pouvoirs de rendre des ordonnances et d'appliquer des sanctions financières permettraient aux Canadiens d'obtenir des règlements plus rapidement et les rassureraient quant à leur capacité de participer en toute confiance au marché numérique. » (nos italiques)

[154] À propos de l'invalidité du consentement dans une relation hiérarchique an acte ou en devenir voir GROUPE DE TRAVAIL « ARTICLE 29 », *Lignes directrices sur le consentement au sens du règlement 2016679.pdf*, 17/FR WP259 rév.01, Bruxelles, Commission Européenne, 2018, p. 7-11, en ligne : <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051> (consulté le 6 juillet 2018). À propos de l'invalidité du consentement dans une relation hiérarchique en acte ou en devenir.

[155] Andrée LAJOIE, *Pouvoir disciplinaire et tests de dépistage de drogues en milieu de travail: illégalité ou pluralisme*, coll. Collection Relations industrielles ; 27, Cowansville, Québec, Éditions Y. Blais, 1995.

[156] Les biais des IA de recrutement ainsi que l'enjeux de leur identification seront abordés dans le prochain chapitre.

demeure inintelligible au commun des mortels. En d'autres termes, consentir à l'IA de recrutement expose la personne candidate à un risque accru de discrimination à l'embauche. Le consentement aurait donc pour conséquence une renonciation aux droits et libertés fondamentaux lui assurant tant le respect de son droit à l'égalité, et plus largement à son droit à la dignité[157].

D'autre part, dans l'hypothèse où une telle renonciation serait conforme aux chartes et lois relatives aux droits de la personne, la validité du consentement à une telle renonciation parait peu probable. En effet, les personnes candidates à un emploi, bien qu'informées des risques de l'IA, sont-elles en position de refuser l'utilisation d'une IA dans l'évaluation de leur candidature ? Ce questionnement rappelle celui suscité par les tests de dépistage de drogues à l'embauche des années 1980[158]. Dans son analyse de cet enjeu, Andrée Lajoie rappelait notamment l'obligation de bonne foi reposant sur l'employeur dans l'exercice de sa capacité contractuelle[159]. L'employeur serait ainsi tenu d'exercer son pouvoir contractuel dans le respect du principe de bonne foi, *a fortiori* lorsqu'il bénéficie d'un déséquilibre contractuel :

> « La règle me semble d'autant plus impérative qu'elle s'applique à un moment où la liberté contractuelle de l'employé, surtout dans une conjoncture économique caractérisée par le chômage comme celle que nous vivons présentement, n'est qu'une fiction juridique formelle : qui, aujourd'hui, est vraiment libre de négocier les conditions que lui impose une offre d'emploi ? »[160].

Près de 30 ans plus tard, à l'heure des candidatures numériques et du chômage accru par la pandémie de covid-19, cette question fait grandement écho à la situation des personnes

---

[157] On retrouve sous la plume d'A. Lajoie un rappel des arguments appuyant l'impossible renonciation à droits fondamentaux. Citant les décisions *Dickason* et *Turpin,* certains droits fondamentaux ne sauraient être écartés car ils conditionnent l'exercice du droit à l'intégrité : « c'est par la protection qu'ils constituent pour la dignité et la valeur de la personne humaine qu'elles caractérisent les droits auxquels on ne peut renoncer ». A. LAJOIE, préc., note 155, p. 38-39; *University of Alberta c. Alberta (Human Rights Comm.)*, [1992] 2 R.C.S 1103-1198, par. 112 (CSC); *Sharon Turpin and Latif Siddiqui c. Her Majesty the Queen and the Attorney General of Canada et al. Case*, [1989] 1 R.C.S 1296-1336, 1316 (CSC).

[158] A. LAJOIE, préc., note 155.

[159] *Id.*, p. 41. L'absence de contrat de travail ferait obstacle à l'exercice du pouvoir disciplinaire et normatif de l'employeur, néanmoins celui-ci exerce des pouvoirs contractuels issus de la capacité de sa société à poser des actes juridiques nécessaires à son objet et ses finalités. Dans une relation de recrutement, c'est donc en vertu de ce pouvoir contractuel que l'employeur pourrait contrevenir aux droits fondamentaux de la personne candidate à l'emploi.

[160] *Id.*, p. 46.

demandeuses d'emploi soumises aux IA de recrutement imposées : sont-elles libres de refuser ces IA de tri de candidatures ? La lucidité amenant une réponse négative, l'intérêt d'étudier les protections face au risque d'IA de recrutement discriminatoires demeure entier.

Aussi, il convient d'observer qu'au titre des garanties dédiées aux droits fondamentaux — et donc exigées en présence d'IA de recrutement — le fédéral et le Québec s'orientent vers la définition d'obligations renforcées en présence de traitement automatisé, notamment en matière de décision algorithmique.

De surcroît, l'anonymisation ou la responsabilité démontrable constituent autant de garanties qui ressortent autant des travaux de la province québécoise que du fédéral. Plus spécifiquement, l'obligation de réaliser et de publier une Évaluation d'Impact algorithmique ou une EFVP[161] semble constituer la clé de voute de l'encadrement des IA. Au niveau fédéral, on peut d'ailleurs relever ce prérequis dans les propositions du Commissariat à la vie privée. Compte tenu de la gravité potentielle des discriminations à l'embauche engendrées par une IA, EFVP pourrait en effet représenter un moyen incontournable de prévenir, et de corriger les biais d'IA de recrutement.

En conclusion, ce rôle central accordé à l'évaluation des algorithmes parait donc indispensable dans une approche fondée sur les risques. Une telle obligation d'évaluation des systèmes de décision automatique se déclinerait ainsi à travers une véritable logique de gestion du risque de discrimination des IA de recrutement. Nous étudierons par ailleurs dans un deuxième chapitre les outils juridiques qui offrent, d'ores et déjà, des occasions de développer une politique de gestion des risques juridiques des IA de recrutement.

[161] Évaluation des facteurs relatifs à la vie privée.

CHAPITRE 2 — LA GESTION DES RISQUES DE DISCRIMINATION DU FAIT DE L'IA DE RECRUTEMENT, LE RENFORCEMENT NÉCESSAIRE DES INSTRUMENTS JURIDIQUES

Dans le précédent chapitre, nous avons exposé le régime juridique des discriminations à l'embauche par IA de recrutement. En l'absence de normes dédiées à ce type d'IA, les repères légaux se résument au régime des discriminations en matière d'emploi, ainsi qu'aux législations sur la protection de la vie privée et des données personnelles. Nous verrons donc dans ce deuxième chapitre, les différents mécanismes juridiques susceptibles de soutenir une gestion du risque de discrimination que présente l'IA de recrutement. En effet, le premier enjeu face à l'IA de recrutement repose sur son imperfection. Ayant vocation à guider la décision, l'IA de recrutement produit une analyse, un jugement, sur les informations qui lui sont soumises, et ce, à partir des données sur lesquelles l'algorithme a été entraîné.

Par conséquent, les biais[162] d'une IA dans le cadre d'un recrutement impliquent une influence menant parfois à des analyses erronées sur les candidatures. Transposé au vocabulaire juridique, le biais algorithmique de l'IA de recrutement a pour effet une discrimination à l'embauche. Ces technologies étant fondées sur des corrélations[163], l'origine des failles de l'IA peut résider aussi bien dans les données que dans l'algorithme en lui-même.

[162] « Erreur qui affecte toutes les observations de façon identique, mais pas nécessairement de façon égale, et qui produit des résultats plus hauts ou plus bas que les valeurs réelles avec cohérence », dans OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, « biais », dans Grand dictionnaire terminologique, coll. Statistique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=2077248> (consulté le 1 octobre 2020).

[163] « Liaison entre des données quantitatives ou qualitatives, mise en évidence par une analyse statistique. […] Les études de corrélation permettent notamment de déterminer s'il existe ou non une association entre deux séries d'observations et d'évaluer l'importance et le sens de cette association (coefficient de corrélation). Par exemple, la corrélation entre le nombre d'appels téléphoniques et les heures auxquelles ils sont reçus. » OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, « corrélation », dans Grand dictionnaire terminologique, coll. Statistique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8420133> (consulté le 1 octobre 2020).

En outre, nonobstant l'absence d'un cadre spécifique aux IA, plusieurs obligations et notions juridiques existantes peuvent guider les pratiques des organisations ayant recours à l'IA de recrutement, voire leur permettre d'élaborer leur stratégie de gestion du risque de discrimination. Ainsi, même en l'absence de législation spécifique à l'IA, nous verrons que des instruments juridiques existent pour répondre aux risques que soulève cette technologie à mesure que les processus de recrutement augmentés se développent.

Nous étudierons donc en premier lieu les risques intrinsèques aux bases de données des IA de recrutement (A). Dans un second temps, s'appuyant sur deux des sources du « risque-algorithmes »[164], il conviendra d'évaluer les moyens juridiques pertinents dans la gestion de ce risque dans les processus de recrutement augmenté (B).

A. La base de données de l'IA de recrutement, un vecteur de transmission des biais humains à la machine : le cas Amazon

Alors que les humains ne peuvent entièrement se départir de leurs propres biais, ils sont pourtant créateurs de technologies qui se veulent objectives. En effet, en matière de recrutement les préjugés, les *a priori* ainsi que les discriminations constituent un enjeu connu, au point que les IA de recrutement ont nourri l'espoir d'atteindre l'idéal de recrutements dénués de biais cognitifs[165]. Pour autant, les biais de l'IA proviennent souvent d'une contamination humaine.

[164] Sihem AMER-YAHIA, Amélie FAVREAU et Juliette SÉNÉCHAL, « D'où vient le risque ? Des données et des algorithmes », *Le Monde*, sect. Binaire (5 février 2020), en ligne : <https://www.lemonde.fr/blog/binaire/2020/02/05/les-plateformes-numeriques-un-foyer-pour-les-risques-donnees-et-algorithmes/> (consulté le 13 juillet 2020). Le risque algorithmique, ou « risque-algorithmes », est une notion qui « se réfère aux dérives de discrimination » par le biais des algorithmes des plateformes. La notion permet l'identification de trois sources de dérives : l'enfermement algorithmique (l'exposition exclusives aux contenus qui confortent les opinions personnelles), le traçage des données des utilisateurs (impliquant le profilage et la surveillance accrue des activités internet des individus), et pour finir, la discrimination à proprement parler (le traitement inégalitaire vis-à-vis des services et ressources sociales comme la justice, l'emploi ou les forces de l'ordre). Dans le cas présent c'est le traçage et la discrimination qui nous intéressent dans l'identification des biais des IA de recrutement.

[165] *Id.,* par. « Des risques aux biais sur les données et dans les algorithmes ». On retrouve en effet une liste de différents types de biais cognitifs affectant le jugement des humais : le biais de conformité, le biais de confirmation, le biais de corrélation illusoire, le biais d'endogénéité. Voir également Patrice BERTAIL, David BOUNIE, Patrick WAELBROECK et Stephan CLÉMENÇON, *Algorithmes : biais, discrimination et équité*, Télécom ParisTech & Fondation

Cette contagion de l'humain à la machine s'observe particulièrement à travers la base de données d'apprentissage. Nous verrons donc en premier lieu le processus par lequel l'IA de recrutement apprend à discriminer (1), puis en second lieu, il s'agira de mettre en exergue les biais résultant de l'apprentissage autonome et de l'étiquetage des données (2).

1. Les biais communiqués à l'IA de recrutement durant son entraînement

La base de données d'apprentissage constitue un échantillon de données. Cet échantillon sert à entraîner l'algorithme. Afin de saisir les particularités du processus d'apprentissage machine, nous verrons dans un premier temps le rôle de la base de données dans l'apprentissage machine, avant d'étudier les contaminations d'une IA de recrutement à travers le cas d'Amazon.

a) Le rôle de la base de données dans l'apprentissage des algorithmes d'IA

Le mode d'apprentissage de l'IA de recrutement suit un procédé commun à d'autres IA. Afin de cerner ses attributs, comparons-le à un mode d'apprentissage humain.

Dans l'apprentissage de la lecture à un humain, la méthode du « B A BA » consiste à enseigner les syllabes, afin qu'ensuite, l'humain puisse décortiquer des mots par syllabes, et les lire. Avec une telle méthode, même des mots nouveaux demeurent lisibles. L'apprentissage de la lecture a donc pour objectif la capacité à lire n'importe quel mot, qu'il soit découvert ou connu par le lecteur. Dans l'apprentissage machine l'objectif est analogue. En matière de reconnaissance d'images par exemple, les bases de données seront constituées de milliers de photos numériques d'animaux. Celles-ci seront classées en fonction des animaux présents sur les images, par exemple : chien, chat, oiseau, ours. L'algorithme devra donc examiner les photos et étiqueter les données soumises selon les quatre catégories d'animaux apprises, et ce, au même titre qu'un enfant

ABEONA, 2019, par. « 3.1 Les biais cognitifs », en ligne : <https://www.telecom-paris.fr/wp-content-EvDsK19/uploads/2019/02/Algorithmes-Biais-discrimination-equite.pdf> (consulté le 12 octobre 2020).

décomposerait les mots en fonction des syllabes apprises. In fine, l'apprentissage machine permet à l'algorithme d'étiqueter correctement des données sur lesquelles il n'a pas été entraîné (par exemple en identifiant des races de chiens absentes de sa base de données). Plus la marge d'erreur de l'algorithme sera réduite, et plus on l'estimera fiable. De ce fait, la base de données d'apprentissage est à l'algorithme, ce qu'un manuel scolaire serait à l'élève : la qualité de son contenu et l'orientation de celui-ci auront une conséquence directe sur ce qui est assimilé.

Pour autant, là où pour identifier des chiens, un enfant n'aura pas besoin d'être exposé à une diversité de représentations de chiens, l'algorithme lui n'établira pas de corrélations instinctives. Il faudra donc lui apprendre une grande variété de représentations de chiens avant qu'il puisse procéder à des identifications de manière autonome. La capacité d'un algorithme d'IA à produire des résultats pertinents dépend donc radicalement de ce qui lui a été appris.

S'agissant des formes d'IA de recrutement qui nous intéressent, on attend d'elles deux choses : le tri des candidatures, ou la recherche des profils pertinents (aussi appelés « talents »). Dans le cas du tri de candidatures, on aura appris à l'IA de recrutement la reconnaissance des mots clés, des notions, et autres informations pertinentes extraites d'un CV ou d'une candidature[166].

Une fois compris le rôle central des données d'apprentissage se pose une question : comment l'algorithme peut-il produire de la discrimination ? Pour ce faire, revenons à l'exemple de la reconnaissance d'images. Si la base de données n'est constituée que de labradors et d'ours bruns, alors l'algorithme n'aura pas appris à reconnaître des chiens et des ours ; il aura appris à reconnaître uniquement des labradors et des ours bruns. Face à des photos de berger allemand ou de dalmatien, l'algorithme ne parviendra pas à les identifier à des chiens. De même, les photos d'ours polaires ne seront pas non plus étiquetées correctement. Dans l'entraînement d'IA de recrutement, la même logique se vérifie. Si l'IA est entraînée avec des données qui ne reflètent pas la diversité des profils professionnels existants, alors l'IA effectuera son tri ou sa recherche

[166] Ici le terme candidature renvoie à toute information communiquée dans le cadre d'une réponse à une offre d'emploi qui diffère du CV classique. Par exemple, certaines offres d'emploi publiée sur le réseau professionnel LinkedIn permettent aux utilisateurs d'effectuer une candidature en partageant leur profil LinkedIn. Il n'y a donc pas de CV ou de lettre de présentation.

de talents, uniquement auprès des personnes dont le profil correspond à ce qu'on lui a appris. Lorsque les données d'entraînement reflètent surtout les profils d'hommes blancs, issus de grandes écoles, de 20 à 35 ans, il y a de fortes chances que l'IA de recrutement leur accorde une priorité face à des profils qui sortent de l'uniformité apprise. Ce phénomène s'est d'ailleurs produit avec l'IA de recrutement d'Amazon. Nous utiliserons ce cas concret de discrimination par IA de recrutement afin d'illustrer la transmission des biais humains à la machine.

b) L'échec de l'IA de recrutement d'Amazon

Dans un article du 9 octobre 2018, Reuters révélait les failles de l'IA de recrutement d'Amazon[167]. Soucieux de gagner en productivité dans son processus d'embauche, le géant d'internet a développé sa propre IA de recrutement. Il s'agissait d'une IA de tri de candidatures. Un an après sa mise en fonction, l'algorithme a été mis à l'index comme présentant des biais sexistes : à compétences égales, les candidatures de femmes étaient moins bien classées que celles des candidatures d'hommes. Prenant acte des résultats biaisés de sa technologie, Amazon a tenté de corriger le biais, avant de se résoudre à l'abandon de sa technologie. Plusieurs observations peuvent être formulées.

En tout premier lieu, la source de ces classements discriminatoires est vraisemblablement identifiée au niveau de la base de données d'entraînement. Développé sur la base de l'historique de recrutement d'Amazon et de profils du personnel de la firme, l'algorithme a corrélé le genre à la pertinence du profil professionnel. Le milieu des Tech, à plus de 80 % masculin[168], fait partie des secteurs les moins avancés en matière de parité et d'inclusion dans la composition des équipes salariées. Loin d'échapper à cette inégalité systémique, Amazon a donc entraîné son IA de recrutement sur la base de données reflétant un salariat masculin en écrasante majorité. En d'autres termes, la composition inégalitaire des équipes d'Amazon a servi de base

[167] Jeffrey DASTIN, « Amazon scraps secret AI recruiting tool that showed bias against women », *Reuters* (09 octobre 2018), en ligne : <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> (consulté le 7 août 2019).

[168] M. WHITTAKER et al., préc., note 18.

d'apprentissage à l'algorithme. L'algorithme a donc appris à classer les profils féminins comme étant éloignés du profil idéal recherché. Plus précisément, l'IA de recrutement semble avoir intégré un champ lexical genré, de sorte que les profils comportant des marques de féminité soient mal classés par l'algorithme :

> « In effect, Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter. »[169]

Il est intéressant de remarquer que même en programmant l'algorithme à la détection des compétences clés, l'IA de recrutement n'a appris qu'un vocabulaire représentatif des CV masculins.

> « They taught each to recognize some 50,000 terms that showed up on past candidates' resumes. (…) Instead, the technology favored candidates who described themselves using verbs more commonly found on male engineers' resumes (...) »[170]

Dans le cas d'Amazon, c'est un biais de stéréotype[171], transmis à l'IA par la constitution d'un biais des données[172], qui conduit l'IA de recrutement à défavoriser les candidatures féminines. C'est donc le manque de diversité de la base de données d'Amazon qui a véhiculé des biais humains à deux égards.

Premièrement, parce que les références s'appuyaient sur les usages linguistiques plutôt masculins. Il s'agit donc d'un choix d'échantillon biaisé. Deuxièmement, parce que ces usages linguistiques étaient majoritairement représentés dans la base de données d'apprentissage, et ce au détriment d'autres usages linguistiques. L'IA n'a donc pas été entraînée de manière neutre, et n'a pas pu apprendre à reconnaître des compétences ou des profils purement professionnels. Par la carence de données d'apprentissages représentant les profils féminins, des biais de genre ont

---

[169] J. DASTIN, préc., note 167.*Id.*

[170] J. DASTIN, préc., note 167.*Id.*

[171] Le biais de stéréotype constitue « une tendance qui consiste à agir en référence au groupe social auquel nous appartenons » S. AMER-YAHIA, A. FAVREAU et J. SÉNÉCHAL, préc., note 164.

[172] Le biais de données, les auteurs l'identifient dans les valeurs des données d'entraînement qui composent la base de données « Par exemple, c'est le cas pour un algorithme de recrutement entraîné sur une base de données dans laquelle les hommes sont sur-représentés exclura les femmes » *Id.*

été appris et reproduits à plus grande échelle par l'algorithme. Le cas Amazon permet de rappeler une réalité : jusqu'à nouvel ordre, l'entraînement de l'IA constitue une démarche rétrospective et non prédictive. Les solutions d'IA sont fondées sur l'historique et non sur l'imagination de l'avenir. La marge d'adaptation, d'évolution, demeure donc très limitée. En d'autres termes, l'algorithme reproduit ce qu'on lui a appris, et en la matière un adage demeure évocateur : « garbage in, garbage out »[173].

Or, dans la mesure où il n'est pas possible de prévoir toutes les situations dans lesquelles l'IA intervient, tout l'enjeu de l'apprentissage machine repose sur la capacité de l'algorithme à demeurer efficace face à des données nouvelles et de situations variées auxquelles il n'aurait pas été entraîné. C'est sa capacité de généralisation.

> « […] la règle de décision déterminée par un algorithme d'apprentissage opérant sur une base de données étiquetées (i.e. les "données d'apprentissage") ne doit pas seulement pouvoir prédire le passé, mais permettre de prédire efficacement, lorsqu'elle sera déployée, le label Y associé à de nouvelles données d'entrée X, non encore observées. On dira le cas échéant que la règle prédictive a alors de "bonnes capacités de généralisation" ». [174]

Dans le cas d'Amazon, on comprend donc que le biais qui affecte la base de données, et ainsi nuit à sa capacité de généralisation, constitue la carence de profils féminins dans les données d'apprentissage. Il s'agit donc bien d'un biais humain — excluant les profils féminins de la base de données — qui a été transmis à l'IA de recrutement. Par ailleurs, relevons ici que les biais de l'IA de recrutement d'Amazon lui ont été appris *a priori*, c'est-à-dire avant sa mise en fonction. Il ne s'agissait donc pas de biais appris au fur et à mesure que les candidatures ont été soumises à l'algorithme. Or, nous verrons dans les développements suivants que l'entraînement de l'IA ne forme pas la seule source de contamination de l'algorithme.

---

[173] FOLDOC, « Garbage In, Garbage Out », dans The Free On-line Dictionnary of Compunting, en ligne: <http://foldoc.org/garbage+in+garbage+out> (consulté le 13 juillet 2020). Cet adage signifie que la qualité des résultats (ou données de sortie), dépend de la qualité des données d'apprentissage (ou données d'entrée).

[174] P. BERTAIL, D. BOUNIE, P. WAELBROECK et S. CLÉMENÇON, préc., note 165, p. « 3.2 Les biais algorithmiques ».

2. Les contaminations de la base de données au fil de l'utilisation de l'IA

À travers le cas d'Amazon, nous avons pu mettre en évidence le rôle central de la base de données d'entraînement d'une IA de recrutement. Nous verrons ici que l'apprentissage autonome et l'étiquetage constituent également des moyens par lesquels l'IA apprend des biais humains.

a) L'apprentissage autonome, un risque accru d'apprentissage de biais discriminatoires

Certaines IA poursuivent leur apprentissage de manière autonome afin de couvrir une plus grande diversité de situations apprises. Dans cette configuration, l'enjeu réside dans la conciliation d'un apprentissage autonome sans que soit affectée la capacité de généralisation acquise. Plus simplement, la problématique qui se pose avec l'apprentissage autonome se résume à une question : comment conserver un taux d'erreur minimal alors que l'IA poursuivra son apprentissage de manière autonome ? La question repose donc sur la robustesse de la capacité de généralisation. La comparaison entre les erreurs d'une IA de Microsoft et celles d'Amazon nous permettra une mise en exergue des contours de cet enjeu.

En effet, si l'IA de recrutement d'Amazon nous confirme les risques inhérents à la constitution de la base de données, nous verrons dans les prochains développements que l'exemple du *chatbot* Tay met en exergue un autre risque : l'apprentissage algorithmique biaisé. Nous constaterons qu'ici également, la base de données est demeurée le vecteur de biais.

Tay était un *chatbot*[175] censé simuler les publications d'une jeune personne utilisant Twitter[176]. Le principe était le suivant : au fil des interactions avec les personnes utilisatrices de

[175] FOLDOC, « Chatbot », dans Free On-line Dictionnary of Computing, en ligne : <https://foldoc.org/chatbot> (consulté le 1 novembre 2020). « (Ou "chatterbot") C'est un robot conçu pour être capable de converser avec les humains. Un chatbot est soit un exercice d'IA, ou une interface comme l'infobot » […]. » On retrouve notamment ces chatbot sur de nombreux sites internet d'entreprise pour assurer un premier niveau de service client, orienter ou répondre aux questions courantes des clients. Dans le cas de Tay il s'agissait d'un chatbot destiné à converser avec les utilisateurs de Twitter.

[176] Elle HUNT, « Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter », *The Guardian*, sect. Technology (24 mars 2016), en ligne : <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>.

l'Oiseau bleu, le *chatbot* poursuivait son apprentissage afin de rester au plus près des habitudes des dits millénniaux. L'apprentissage s'effectue donc par mimétisme, en fonction de ses interactions. Or, sur un tel réseau social il n'était pas possible d'anticiper le type d'interactions dans lesquelles Tay serait impliqué. Et, à la suite d'une pollution massive[177], en moins d'une journée le chatbot publiait des propos antisémites, misogynes et racistes. Microsoft a dû mettre un terme à l'expérience. Quand bien même on comprendrait que Microsoft n'ait pas anticipé une attaque malveillante, la question de la robustesse de la capacité de généralisation pourrait interroger. En effet, il pourrait être étonnant qu'une IA censée reproduire le comportement des jeunes internautes se retrouve à véhiculer des discours fascistes crus.

Or, en l'espèce ce n'est pas cette capacité de généralisation qui pose problème : et pour cause. C'est justement parce que l'attaque visant Tay l'a exposé à des contenus fascistes en nombre disproportionné que l'IA les a intégrés. Le problème avec Tay venait de l'incapacité de l'IA à trier le type de contenu qu'elle aurait dû exclure de son apprentissage. Pour Tay, l'exposition à de nouvelles données sciemment biaisées, et en grand nombre, a eu pour résultat de lui apprendre des biais racistes, antisémites et sexistes. Par conséquent, dans le cas d'un apprentissage autonome et continu, il est peut-être plus complexe de contrôler la qualité des données apprises par l'IA et donc de limiter l'illustration brutale de « garbage in, garbage out ».

Toute proportion gardée, une attaque analogue dirigée contre une IA de recrutement pourrait théoriquement générer l'apparition de biais discriminatoires, pourtant absents de la base de données initiale. La robustesse d'une IA ne dépend donc pas uniquement de la qualité de l'entraînement fait en laboratoire ; le cas de Tay met clairement en évidence l'importance de sécuriser les données de l'apprentissage autonome. La sécurité et la qualité de la base de données constituent donc des enjeux interdépendants dans la préservation contre des discriminations par IA de recrutement.

[177] LIBÉRATION, « Microsoft muselle son robot «Tay», devenu nazi en 24 heures », *Libération*, sect. Futurs (25 mars 2016), en ligne : <https://www.liberation.fr/futurs/2016/03/25/microsoft-muselle-son-robot-tay-devenu-nazi-en-24-heures_1441963> (consulté le 14 juillet 2020).

Enfin, au-delà des réflexions relatives à la protection de la base de données de l'IA, les cas d'Amazon et de Google Photos, soulèvent un autre vecteur de discrimination dans la base de données d'apprentissage : l'étiquetage. L'étiquetage constitue l'étape essentielle du tri de données, l'étiquette apposée sur les données d'apprentissage permet à l'IA de les identifier et influe directement sur les données de sortie (ou résultat) qu'elle produit. Autrement dit, au stade de l'entraînement de l'IA, l'étiquetage représente l'opération par laquelle on lui indique ce qu'est une photo de chat, et celle d'un chien. Appliqué à l'IA de recrutement, l'étiquetage consistera à classer les caractéristiques que l'algorithme devra rechercher et analyser dans les profils de personnes candidates à un poste, ou encore de talents[178]. Chaque élément pertinent sera donc classé dans une ou plusieurs catégories qu'on appelle « étiquettes ». Selon la précision et la qualité de l'étiquetage, nous verrons que celui-ci peut constituer un facteur de discrimination, même face à des données représentatives de la diversité des candidatures possibles. Sur la base des étiquettes apprises à l'IA, il est possible de classer un nombre considérable d'informations, de filtrer les données ou les résultats de l'algorithme en fonction de ce que l'on recherche. C'est par exemple par ce moyen que les moteurs de recherche définissent un ordre d'apparition des résultats. Un moteur de recherche comme Google peut en effet s'appuyer sur les activités d'une personne X afin de définir un ordre d'apparition des résultats qui est adapté à ses habitudes.

De même, les entreprises clientes de Facebook utilisent des filtres d'audience pour toucher un public plus précis à travers leurs publications. Or, les filtres d'audiences de Facebook ne pourraient pas guider les algorithmes de la firme sans un travail constant d'étiquetage que les utilisateurs de la plateforme réalisent quotidiennement et gratuitement[179]. Alors pourquoi ce

[178] Les caractéristiques recherchées et analysées dépendront du type d'IA de recrutement selon qu'elle aide au tri de candidatures, ou qu'elle permette la recherche de talents.

[179] « Le « travail gratuit » correspond au travail accompli par les utilisateurs de réseaux sociaux ou de moteur de recherche. Il revient à réaliser des tâches d'étiquetage gratuitement, sans s'en rendre compte, tel que l'identification de personnes ou de lieux sur des photos » dans Laetitia DIMANCHE et Erwan JONCHÈRES, « Les travailleurs du clic et l'intelligence artificielle », *Laboratoire de cyberjustice* (21 juin 2019), en ligne : <https://www.cyberjustice.ca/2019/06/21/les-travailleurs-du-clic-et-lintelligence-artificielle/> (consulté le 10 août 2020). Rappelons par ailleurs « Au deuxième trimestre 2020, Facebook revendiquait 2,7 milliards d'utilisateurs actifs chaque mois, en hausse de 12% par rapport au deuxième trimestre 2019 », voir à ce propos « Nombre d'utilisateurs de Facebook dans le monde », *Journal du Net* (31 juillet 2020), en ligne : <https://www.journaldunet.com/ebusiness/le-net/1125265-nombre-d-utilisateurs-de-facebook-dans-le-monde/> (consulté le 7 octobre 2020).

procédé consistant à apposer des étiquettes facilite-t-il la transmission des biais de l'humain à la machine ? Le cas de Google Photos apporte quelques réponses.

b) L'étiquetage, un vecteur insidieux de biais discriminatoires

En juillet 2015**,** The Telegraph rapportait la mise à l'index de Google Photos à propos de dénonciations de racisme subi par des personnes afro-américaines[180]. Une fonctionnalité de l'application permet en effet la reconnaissance d'objet, d'animaux et de personnes. Or, il s'avère qu'en présence de personnes dites noires, l'application les identifiait sous l'étiquette « Gorille ». Comment l'IA a-t-elle pu apprendre une corrélation aussi raciste ? La base de données demeure, ici également, le point d'achoppement.

D'abord, rappelons que la technologie d'IA intégrée à Google photos apprend de manière autonome par le biais des contenus présents sur internet. Par conséquent, il suffit que ce biais existe dans les publications analysées pour que l'algorithme ait appris que ces personnes sont régulièrement associées à l'étiquette « gorilles ». Donc si statistiquement on retrouve de fréquentes associations entre les termes gorilles et des représentations de personnes noires, l'algorithme de reconnaissance d'images pourra apprendre cette corrélation et la généraliser.

Dans ce cas, l'apprentissage automatique est biaisé par son incapacité à discerner la fiabilité des étiquetages présents sur la Toile. Il est d'ailleurs important de rappeler que, *a priori*, évaluer la pertinence des corrélations d'internet n'est pas ce qui est demandé à l'algorithme de Google Photos. Le cas de Google Photos renforce donc l'idée que les données d'apprentissage constituent un vecteur critique de transmission des biais.

Ensuite, il y a une autre origine à ces erreurs répétées de Google Photos : un manque de diversité des données. En effet, corriger le biais de l'application passerait par exemple par un entraînement renforcé de l'IA quant à la reconnaissance de personnes noires et de gorilles.

[180] S. CURTIS, préc., note 74.

Néanmoins, plutôt que de passer par ce procédé, Google Photos a opté pour le blocage des étiquettes « gorilles », « chimpanzé », « homme noir », et « femme afro »[181]. Comparée à l'IA de recrutement d'Amazon, la suppression des étiquettes se référant au genre n'aurait pas réglé le biais, de même que pour Google Photos, la disparition de ces étiquettes a pour seul intérêt d'éviter des identifications malheureuses et insultantes. Dans le cas de l'IA d'Amazon, supprimer des étiquettes n'était cependant pas envisageable : l'IA n'avait pas uniquement appris à rejeter les marqueurs de féminité, il a surtout appris à reconnaître ceux de la masculinité comme constituant des critères positifs de classement des profils. En d'autres termes, l'IA d'Amazon ne distingue pas les hommes des femmes, l'IA se contente de comparer un modèle à des profils, et de les classer en fonction de la proximité avec le modèle. Quand bien même ledit modèle serait genré, l'IA de recrutement se contente de le rechercher. Supprimer des marqueurs de genre ne pourrait donc passer que par une réinitialisation de l'apprentissage de l'algorithme. Amazon a d'ailleurs mis un terme à son expérience avec l'IA de recrutement plutôt que passer par un processus aussi chronophage. Dans le cas de Google Photos comme d'Amazon, la suppression des étiquettes ou de l'IA productrices de discriminations nous permet de mettre le doigt sur une problématique connexe : qui procède à l'étiquetage des données ?

À ce propos, deux procédés d'étiquetage biaisés nous intéressent ici : l'étiquetage des travailleurs du clic et celui issu des personnes utilisatrices des grandes plateformes.

S'agissant de l'étiquetage par les personnes travailleuses du clic d'une part, en 2019 le sociologue Antonio Casilli exposait l'existence des « tâcherons du numérique »[182]. Ses travaux ont mis au jour le travail humain sur lequel reposent les algorithmes, derrière les plateformes et l'IA.

---

[181] Jacky SNOW, « Google Photos Still Has a Problem with Gorillas », *MIT Technology Review* 2020, en ligne : <https://www.technologyreview.com/2018/01/11/146257/google-photos-still-has-a-problem-with-gorillas/> (consulté le 13 octobre 2020); Tom SIMONITE, « When It Comes to Gorillas, Google Photos Remains Blind », *Wired* (1 novembre 2018), en ligne : <https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/> (consulté le 13 octobre 2020).

[182] « IA: les fermes à clics remplacent les sweatshops », *Les affaires* (21 mars 2019), en ligne : <https://www.lesaffaires.com/blogues/diane-berard/ia-les-fermes-a-clics-remplacent-les-sweatshops/609039> (consulté le 14 juillet 2020).

Plus particulièrement, les travaux des sociologues mettent en exergue le développement du « micro-travail »[183]. Ces « tâcherons » procèdent donc à l'étiquetage des données constituant les bases de données d'entraînement supervisé des algorithmes, ou contribuent à l'entraînement de ceux-ci.

> « l'IA AlphaGo a réussi à battre l'humain dans l'un des derniers bastions de l'intelligence naturelle, toutefois cette IA est incapable de différencier un chat d'un chien, chose que tout enfant de 5 ans fait naturellement. C'est à cette étape qu'interviennent les ouvriers du clic, dont le travail consiste à enrichir les bases de données des machines, en réalisant un tri ou en affinant l'information qui sera ensuite traitée […]. »[184]

Si ces travaux soulèvent des questions complexes quant aux mutations du travail, ils permettent également d'identifier l'étiquetage et l'entraînement des IA parmi les tâches déléguées à de la main-d'œuvre précaire[185].

> « Le principe de datafication […] est un travail fondé sur un flux constant de données générées et exploitées. Les travailleurs du clic accomplissent donc des tâches telles que l'écoute de conversation, la transcription de texte, le visionnage de vidéo ou répondent à des questionnaires. »[186]

[183] « Cette activité rémunératrice relativement nouvelle consiste à réaliser des tâches très fragmentées (micro-tâches) que des plateformes dédiées confient à des prestataires, payés généralement la pièce. Il peut s'agir d'identifier des objets dans une image, de transcrire des factures, de modérer du contenu sur les médias sociaux, de visionner des vidéos de courte durée, de copier-coller du texte ou de répondre à des sondages en ligne. Le plus souvent, ces tâches répétitives nécessitent une faible qualification pour une rémunération tout aussi faible, de l'ordre de quelques centimes. » Antonio CASILLI, Paola TUBARO, Clément LE LUDEC, Marion COVILLE, Maxime BESENVAL, Touhfat MOUHTARE et Elinor WAHAL, *Le Micro-Travail en France*, DiPLab, 2019, p. 4, en ligne : <http://diplab.eu/wp-content/uploads/2019/05/Le-Micro-Travail-En-France_DiPLab-2019.pdf> (consulté le 19 juillet 2020).

[184] L. DIMANCHE et E. JONCHÈRES, préc., note 179.

[185] A. CASILLI et al., préc., note 183, p. 55-60.

[186] L. DIMANCHE et E. JONCHÈRES, préc., note 179.

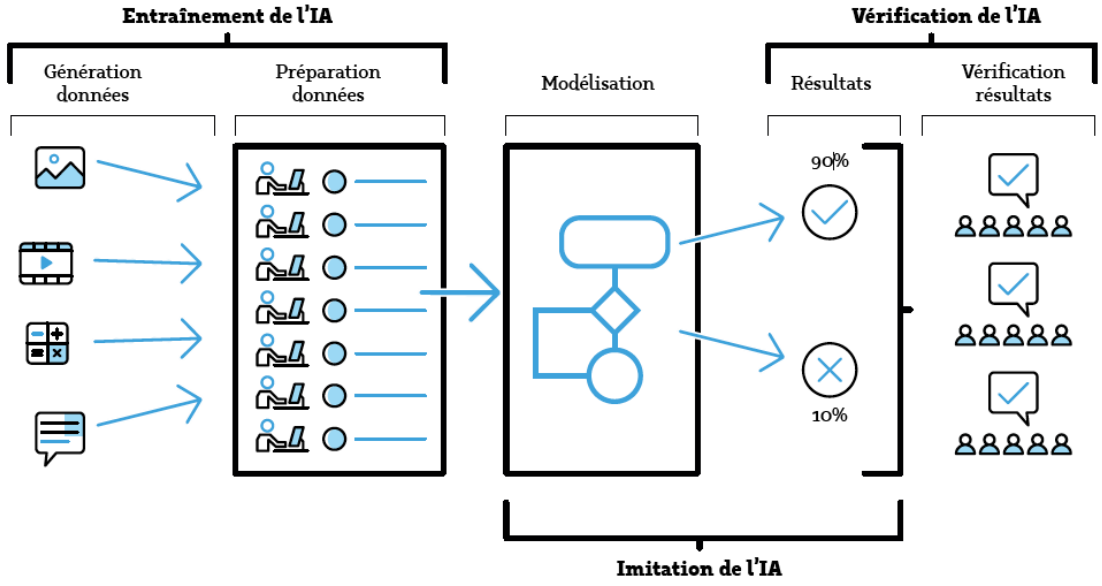

Figure 1. – Trois types d'activités que les microtravailleurs réalisent : entraînement, vérification et imitation des intelligences artificielles[187]

Si par essence le procédé semble anodin, il devient problématique dans la mesure où ces tâcherons ne sont soumis à aucune règle particulière, aucun mode de contrôle des biais individuels. En effet, si de tels processus de constitution de base de données des IA répondent aux besoins quantitatifs, ils conduisent cependant à se reposer sur un travail à la chaîne numérique dont on peine à garantir la qualité. Dans les fermes à clics ou sur les plateformes de microtravail, des personnes issues de pays du monde entier produisent de la donnée d'apprentissage. Cela passe par la véritable création de données, notamment en répondant à des questionnaires et sondages avec toute la subjectivité que suppose un tel travail rémunéré.

En outre, une fois la donnée produite, elle doit être étiquetée. Ici, pareillement, les tâcherons du clic entrent en action. Or, si l'identification d'animaux ne présente pas d'enjeu de discrimination, qu'en est-il de l'identification de qualités humaines ?

À ce propos, on peut notamment prendre l'exemple des solutions d'IA offertes par l'entreprise *HireVue*. L'une de ces solutions consiste à recourir à la reconnaissance faciale au cours

[187] A. CASILLI et al., préc., note 183, p. 60 figure 13.

d'entretiens virtuels avec les personnes candidates. Les entretiens d'embauche traditionnels sont donc réalisés avec une IA de recrutement.

> « HireVue's "AI-driven assessments," which more than 100 employers have used on a million-plus job candidates, use video interviews to analyze hundreds of thousands of data points related to a person's speaking voice, word selection and facial movements. The system then creates a computer-generated estimate of the candidates' skills and behaviors, including their "willingness to learn" and "personal stability." »[188].

Autrement dit, *HireVue* vend une solution aux services de ressources humaines qui leur permet de déléguer à l'IA l'analyse d'une personnalité ou d'aptitudes des personnes candidates. L'IA procède donc à leur profilage[189]. Or, ces dernières années on peut relever l'échec de plusieurs tentatives d'IA en mesure d'assurer les entretiens d'embauche. Par exemple, Vera, une IA de recrutement russe, proposait des solutions similaires[190]. L'IA se voyait déléguer la recherche de talents, la prise de contact avec eux, puis l'entretien. L'analyse reposait alors sur la reconnaissance vocale par téléphone ou par vidéoconférence. Pour autant, tous les grands groupes, au départ séduits, ont mis un terme à leur utilisation de Vera au bout de quelques semaines[191].

Dans le cas de *HireVue*, le profilage prétend aller jusqu'à l'évaluation de la stabilité personnelle des candidats en se fondant sur des algorithmes de prétendue reconnaissance émotionnelle faciale et vocale. De telles analyses par IA supposent des étiquetages par nature intrinsèquement subjectifs et aléatoires, dont les biais et les risques ont déjà été dénoncés[192]. Au-

---

[188] Drew HARWELL, « Rights group files federal complaint against AI-hiring firm HireVue, citing 'unfair and deceptive' practices », *Washington Post*, sect. Technology (6 novembre 2019), en ligne : <https://www.washingtonpost.com/technology/2019/11/06/prominent-rights-group-files-federal-complaint-against-ai-hiring-firm-hirevue-citing-unfair-deceptive-practices/> (consulté le 1 juillet 2020).

[189] Or, comme cela a été abordé au précédent chapitre, le profilage est soumis à des exigences particulières. Voir Supra, Chapitre 1, *2. La protection de la vie privée : une protection contre la décision automatique discriminatoire*.

[190] Q. PÉRINEL, préc., note 16.

[191] Amaury BUCCO, « Vera, le robot recruteur... licencié par ses employeurs », *FIGARO*, sect. Vie de Bureau (30 mai 2018), en ligne : <http://www.lefigaro.fr/vie-bureau/2018/05/30/09008-20180530ARTFIG00003-vera-le-robot-recruteur-licencie-par-ses-employeurs.php> (consulté le 7 août 2019).

[192] Sahil CHINOY, « The Racist History Behind Facial Recognition », *The New York Times*, sect. Opinion (10 juillet 2019), en ligne : <https://www.nytimes.com/2019/07/10/opinion/facial-recognition-race.html> (consulté le 28 août 2019). Dans cette critique virulente des utilisations d'IA de reconnaissance faciale, l'auteur compare, le XIXème siècle et nos jours. En effet ces deux époques ont en commun les tentatives de corrélations fallacieuses entre les traits des

delà du manque de fiabilité des IA de reconnaissance faciale, aujourd'hui incontestable[193], l'utilisation de tels procédés dans le cadre d'un processus de recrutement entraîne d'ores et déjà des plaintes aux États-Unis[194]. En effet, les personnes candidates n'ont aucun moyen de connaître les analyses produites par l'IA de recrutement d'*HireVue*. La solution de profilage vidéo de *HireVue* entraîne d'autant plus de contestations que les personnes candidates n'ont pas accès à l'évaluation rendue, les demandes d'audit indépendant sont refusées, et la violation des principes éthiques définis par l'OCDE est alléguée par les personnes plaignantes.

Parallèlement à l'opacité de l'entreprise, les spécialistes du comportement humain et de ce type de technologies semblent unanimement constater le manque de fiabilité d'un tel profilage appliqué à un processus de recrutement.

En effet, si ces spécialistes reconnaissent que l'IA sait repérer des expressions faciales, des mouvements oculaires ou les variations de la voix, leur interprétation demeure très peu fiable. On pourrait même parler de pseudoscience[195] lorsque *HireVue* prétend que son IA de recrutement est en mesure de déterminer la « stabilité » des individus. Et pour cause : la reconnaissance d'émotions, de traits de caractère par IA de reconnaissance vocale ou faciale soulève autant de questions portant sur les erreurs commises que d'inquiétudes face au risque

individus et leurs personnalités. Aujourd'hui avec la reconnaissance faciale, l'auteur rappelle que cette technologie ne permet pas de reconnaître l'orientation sexuelle, les traits de caractères ou les émotions sur la simple analyse des visages d'individus. De même qu'au XIXème siècle, ces mêmes analyses (manuelles) ne permettaient pas davantage d'identifier les personnes prédisposées à la criminalité. Il s'agit selon l'auteur de procédés connus reposant sur des préjugés racistes séculaires, qui reviennent au gout du jour par le biais de l'IA et de la crédibilité qu'on lui accorde. Les IA de reconnaissance faciale ne permettraient donc qu'un déploiement à grande échelle des délits de faciès.

[193] NIST, préc., note 17. Dans ce rapport, le NIST (National Institute of Standard and Technology) a audité près de 200 IA de reconnaissance faciale développées par plus de 99 ingénieurs différents. L'enquête a abouti au constat des biais racistes et sexistes de ces technologies. Aux États-Unis, les afro-américains et les hommes autochtones sont les premières victimes de ces biais avec des faux positifs bien plus élevés que le taux d'erreurs constaté pour les hommes blancs. De même, les femmes sont moins reconnues par la technologie, *a fortiori* lorsqu'elles sont racialisées (autochtones, afro-américaines en particulier).

[194] D. HARWELL, préc., note 188. Le 06 novembre 2019, une association de défense des droits a saisi la *Federal Trade Commission* des États-Unis d'une plainte contre *HireVue*. L'entreprise propose en effet une IA de recrutement reposant notamment sur la reconnaissance faciale pour évaluer les personnes candidates à un poste. Il est reproché à l'entreprise de recourir à des pratiques dont la fiabilité n'est pas prouvée, ce qui porterait atteinte aux droits des travailleurs américains.

[195] « Pseudoscience », dans Dictionnaire Larousse, en ligne : <https://www.larousse.fr/dictionnaires/francais/pseudoscience/64764> (consulté le 10 novembre 2020).

accru de discrimination. L'étiquetage des données pour de telles IA d'aide au recrutement pose un problème essentiel : la standardisation des réactions humaines dans les bases de données. Qui décide de la signification d'un sourire, d'une absence de sourire, d'un rictus ? Qui décide de ce que constitue une réaction d'embarras, d'irritation, ou de concentration ? Dans la mesure où ces émotions s'expriment différemment en fonction de la personnalité, la culture, les usages, et le contexte dans lesquels évoluent les individus, comment une base de données pourrait-elle répondre aux exigences de neutralité ?

De plus, il ressort des communications de *HireVue* que l'entraînement de l'algorithme s'effectue grâce aux données des salariés de l'entreprise cliente. Les algorithmes ne reposent donc pas sur un entraînement délégué aux tâcherons du numérique. Néanmoins, en choisissant les données des salariés, l'entraînement de l'algorithme serait à nouveau soumis aux risques de biais que l'on a évoqué pour Amazon : l'uniformité des données représenterait un risque accru d'inégalité de traitement des candidatures. Par analogie avec l'adage « Garbage in, garbage out » — et sans préjudice pour la qualité des salariés inspirant le modèle de l'IA —, l'algorithme de recrutement serait donc programmé pour rechercher des profils similaires à ceux existant dans l'entreprise. Si cela peut se comprendre lorsque la recherche est circonscrite aux compétences, elle devient dangereuse pour les droits fondamentaux des personnes candidates lorsque le profilage va jusqu'à une évaluation de la stabilité des individus. Dans cette dernière hypothèse, l'étiquetage présente un danger direct pour le droit à l'égalité des demandeurs d'emploi. En effet, le profilage des candidats dépend exclusivement des étiquettes associées aux comportements repérés par l'IA. Autrement dit, toutes les caractéristiques recherchées chez les personnes candidates dépendront des salariés de l'entreprise recruteuse. En extrapolant les biais de recrutement subséquents, d'aucuns pourraient considérer ces IA comme une forme déguisée de cooptation des candidats par les salariés.

En conclusion, à travers l'exemple d'Amazon et des situations connexes, il y aurait donc trois canaux de transmission des biais humains aux bases de données des IA de recrutement : le

manque de diversité des données d'entraînement, le manque d'encadrement de l'apprentissage autonome continu, et l'étiquetage des données composant la base de données de l'IA.

En outre, après l'analyse comparative de ces différents exemples, voyons maintenant des mécanismes juridiques disponibles pour gérer ces risques de contamination de l'IA de recrutement par les biais humains.

B. La nécessité d'une obligation de moyens renforcée, un encadrement *a priori* des risques inhérents aux bases de données

Dans une perspective de gestion du risque de discrimination, le droit fournit plusieurs mécanismes utiles. On peut notamment citer des obligations juridiques qui nourrissent une stratégie de mitigation des risques dans l'utilisation d'une IA de recrutement. Nous avons identifié le principal vecteur de discrimination : la base de données. La mitigation du risque entourant la base de données de l'IA consiste donc à limiter les transmissions de biais humains à l'IA de recrutement. Nous verrons dans les prochains développements que le renforcement de plusieurs obligations de moyens constitue une stratégie intéressante d'encadrement du risque de contamination de la base de données d'une IA de recrutement.

Donc dans un premier temps nous verrons en quoi l'obligation de moyens renforcée servirait de mécanisme approprié aux enjeux de l'IA de recrutement (1). Dans un second temps, il s'agira ensuite d'étudier plus précisément des exemples d'obligations de moyens renforcées utiles à la mitigation du risque de biais de données (2).

1. L'obligation de moyens renforcée : un mécanisme applicable à l'IA de recrutement

Dans la classification des types d'obligations juridiques, on retrouve une *summa divisio* entre les obligations de moyens et les obligations de résultat.

L'intérêt de la distinction entre ces deux obligations entraîne deux conséquences : en premier lieu, l'identification de la personne qui supporte la charge de la preuve, et en second lieu,

l'intensité de la responsabilité[196] des parties à un contrat[197]. En effet, l'obligation de résultat est une « obligation en vertu de laquelle le débiteur est tenu de parvenir à un résultat précis »[198]. À l'inverse, l'obligation de moyens incarne une « obligation en vertu de laquelle le débiteur est tenu, non pas d'obtenir un résultat précis, mais uniquement de mettre en œuvre tous les moyens pour y parvenir »[199].

Pour autant, Paul-André Crépeau apporte des subtilités à cette distinction. À ce titre, il identifie l'obligation de moyens renforcée, ou obligation de résultat atténuée,[200] à mi-chemin entre l'obligation de moyens et l'obligation de résultat[201]. En effet, l'obligation de moyens entraîne un engagement à agir de son mieux, là où l'obligation de résultat engage à fournir le résultat. Dans le premier cas, l'existence d'un aléa est admise, lorsque dans le second, peu de place est laissée à l'aléa. La conséquence est qu'à l'inverse de l'effet d'une obligation de moyens, lorsque le contrat stipule une obligation de résultat, l'absence du résultat fera peser une présomption de faute sur la personne débitrice. La personne cocontractante n'aura donc pas à démontrer qu'il y a eu faute ou négligence de la part du débiteur, ce sera à ce dernier de démontrer qu'il a été empêché par un cas de fortuit[202]. La question se pose donc de savoir

---

[196] L'engagement de la responsabilité nécessite la réunion de trois éléments de preuve : une faute ou une négligence, un dommage et au moins un préjudice causé par ce dommage. L'intensité de la responsabilité permet de distinguer s'il faut prouver une faute ou une négligence.

[197] H. REID et S. REID, préc., note 25, par. « obligation de résultat ». En effet, dans le cas d'une obligation de résultat, le débiteur ne peut bénéficier d'une exonération de sa responsabilité que s'il parvient à démontrer un cas fortuit. Alors qu'en présence d'une obligation de moyens, le créancier devra démontrer que son débiteur a fait preuve de négligence, ou qu'il n'a pas mis en œuvre tous les moyens d'exécuter sa part du contrat.

[198] *Id.*, par. « Obligation de résultat ».

[199] *Id.*, par. « Obligation de moyens ».

[200] Paul-Andre CRÉPEAU, « Le Contenu Obligationnel d'un Contrat », (1965) 43-1 *Can. Bar Rev.* 1-48, en ligne : <https://heinonline.org/HOL/P?h=hein.journals/canbarev43&i=40> (consulté le 20 octobre 2020) ; Jean-Claude ROYER, « PAUL-ANDRÉ CRÉPEAU, L'intensité de l'obligation juridique ou Des obligations de diligence, de résultat et de garantie, Montréal, Édition Yvon Blais, 1989, 232 p., ISBN 2-89073-726-8. », (1991) 32-1 cd1 240-240 ; Daniel ROUSSY, Chantal BERNIER et Charles MALONE, « Paul-André Crépeau, L'intensité de l'obligation juridique ou des obligations de diligence, de résultat et de garantie », (1991) 22-1 rgd 249-254

[201] Paul-André CRÉPEAU et CENTRE RECHERCHE EN DROIT PRIVÉ ET COMPARÉ DU QUÉBEC, *L'intensité de l'obligation juridique ou des obligations de diligence, de résultat et de garantie*, Cowansville (Québec), Y. Blais, 1989, p. 16-17.

[202] P.-A. CREPEAU, préc., note 200, 40.

comment l'obligation de moyens renforcée peut répondre à la gestion du risque de discrimination par IA de recrutement.

a) La définition de l'obligation de moyens renforcée

L'obligation de moyens renforcée[203] représente l'engagement accru d'une personne à faire, ou à ne pas faire quelque chose par tous les moyens disponibles. La preuve du respect de l'obligation de moyens renforcée entraîne donc la charge pour le débiteur de démontrer sa proactivité dans le respect de son engagement.

Mais alors en quoi cette forme d'obligation présente-t-elle une utilité dans le cadre de la gestion des risques discriminatoires de l'IA de recrutement ? Dans ce contexte technologique, l'obligation de moyens renforcée impliquerait celle de prévenir les risques de transmissions de biais à la machine. En effet, la base de données des IA de recrutement forme donc une compilation d'informations numériques destinées à l'entraînement de l'IA. La protection des informations numériques de l'IA représente donc un enjeu analogue à celui de la sécurité informationnelle.

Aussi, en utilisant l'obligation de moyens renforcée, l'employeur assumerait la survenance de discriminations générées par l'IA. Par analogie avec l'obligation de sécurité informationnelle[204], cela impliquerait qu'en cas de contentieux relatif à une discrimination à l'embauche par IA, l'employeur devra démontrer soit que le rejet de candidature opéré par l'IA ne repose pas sur des biais discriminatoires, soit qu'il a raisonnablement[205] pris toutes les mesures possibles pour prévenir la transmission de ce ou ces biais à l'IA de recrutement utilisée.

---

[203] Les termes « obligation de moyens renforcée » et « obligation de diligence renforcée » seront utilisés comme synonymes. Il en va de même pour les termes « obligation de moyens » et « obligation de diligence »

[204] « Sécurité de l'information », dans Grand dictionnaire terminologique, coll. Sciences de l'information, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8358572> (consulté le 10 novembre 2020); « Ressources informationnelles », dans Grand dictionnaire terminologique, coll. Informatique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8364533> (consulté le 10 novembre 2020).

[205] Le caractère raisonnable tient essentiellement aux coûts que génèrent ces moyens de prévention. Il est généralement admis que les moyens exigés doivent être proportionnés aux ressources de l'entreprise. Par exemple, Les TPE ou les PME ne pourront pas accéder aux mêmes moyens que les GAFAM.

En effet, si dans les lois de protection des renseignements, le législateur ne définit pas une obligation de sécurité informationnelle de moyens ou de résultat, il est possible de déduire que la sécurité informationnelle constitue une obligation de moyens renforcée. Dans la *Loi concernant le cadre juridique des technologies de l'information,*[206] le législateur vise des « mesures de sécurité propres à assurer » la confidentialité des renseignements personnels à l'article 25.

De même, à l'article 10 de la *Loi sur la protection des renseignements personnels dans le secteur privé*[207] il est également question des « mesures de sécurité propres à assurer la protection » desdits renseignements. Dans ces deux dispositions législatives, il s'agit davantage d'objectifs de protection que de l'exigence d'un résultat qui refléterait le risque zéro. Un tel résultat reste, de plus, inatteignable en l'état actuel de l'art. D'ailleurs, l'un des principes de limitation de la responsabilité se résume à l'adage : « à l'impossible, nul n'est tenu ». La protection parfaite des informations constituerait donc un idéal vers lequel tendre, plutôt qu'un résultat impossible par essence.

Suivant le paradigme de l'obligation informationnelle, dans le contexte de l'IA de recrutement la protection de la base de données pourrait ainsi se décliner en deux volets : la sécurité[208] de la base de données d'une part, et la sûreté[209] de celle-ci d'autre part. La distinction signifie que la « sécurité » de la base de données porte sur les risques qu'autrui représente pour la base de données[210], lorsque la sûreté de la base de données renvoie à la neutralisation des risques qu'elle représente pour autrui. En l'occurrence, notre propos porte sur les risques que la

---

[206] *Loi concernant le cadre juridique des technologies de l'information*, R.L.R.Q c C-1.1, en ligne : <http://legisquebec.gouv.qc.ca/fr/ShowDoc/cs/C-1.1> (consulté le 21 février 2020). Ci-après « LCCJTI »

[207] *Loi sur la protection des renseignements personnels dans le secteur privé*, (1993) RLRQ c P-39.1, en ligne : <http://legisquebec.gouv.qc.ca/fr/ShowDoc/cs/P-39.1> (consulté le 18 avril 2019).

[208] « Sécurité », dans Grand dictionnaire terminologique, en ligne : <http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=8366714>. « État de quelqu'un ou de quelque chose qui est à l'abri du danger. »

[209] « Sûreté », dans Grand dictionnaire terminologique, en ligne : <http://www.granddictionnaire.com/ficheOqlf.aspx?Id_Fiche=8366713>. « Caractère de quelque chose qui présente des garanties de protection ou sur quoi on peut compter de manière certaine. »

[210] Il s'agit ici des altérations accidentelles et volontaires de la base de données. A titre d'exemple d'attaques malveillantes, on peut se référer au cas précité du *chatbot* Tay. C'est en effet un exemple de faille de sécurité affectant la base de données.

base de données génère des atteintes au droit à l'égalité des personnes demandeuses d'emploi. Dans une perspective de gestion du risque de discrimination perpétrée par l'IA de recrutement, c'est donc la sûreté de la base de données qui nous intéresse. Dans cette logique de sûreté, l'IA de recrutement dont la base de données n'est pas « sûre » engendrerait des discriminations à l'embauche à l'encontre des personnes candidates qui y sont soumises. Ce paradigme de distinction entre sûreté et sécurité de la base de données renforce l'idée d'une gradation de l'intensité des obligations. En effet, autant en Europe qu'au Canada c'est une stratégie de régulation fondée sur les risques de l'IA qui émerge ; plus spécifiquement, il s'agit pour les autorités existantes d'encadrer le risque d'atteinte aux droits fondamentaux dans le recours aux IA.

Aussi, dans le recrutement, il convient de remarquer que l'obligation de moyens renforcée visant la prévention des biais algorithmiques entérinerait le principe de la responsabilité démontrable proposée par le Commissaire à la vie privée[211].

En effet, la responsabilité démontrable consiste en un mécanisme par lequel l'organisation doit prouver qu'elle se conforme à la loi, par exemple par le biais de l'EFVP (évaluation des facteurs relatifs à la vie privée)[212]. Cela passerait notamment par une obligation de documentation des mesures par lesquelles l'organisme met en œuvre sa conformité. Dans le cas de la prévention des biais algorithmiques dans le recrutement, la responsabilité démontrable suppose une obligation de moyens accrue de mobiliser toutes les ressources raisonnablement disponibles pour limiter au mieux le risque de discriminations à l'embauche. L'obligation de moyens renforcée pourrait donc constituer un mécanisme juridique approprié à la gestion du risque de discrimination par IA de recrutement.

Par ailleurs, les obligations de sécurité et de sûreté semblent également accrues aux yeux de la jurisprudence[213] lorsque des moyens technologiques permettent la réduction des chances de

[211] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101.

[212] Ce concept sera étudié plus en profondeur dans la deuxième partie de ce chapitre. Voir infra *2. La mitigation du biais de données par le renforcement d'obligations existantes.* Nous y reviendrons également dans le chapitre 3.

[213] Voir infra, *b) L'obligation de sûreté de la base de données, une obligation de moyens renforcée par essence.*

survenance de l'aléa (ou du risque). On serait donc en présence d'une obligation de moyens renforcée dès lors que le progrès technique permet une amélioration de la protection des biens ou des personnes. Voyons donc dans quelle mesure la définition de cette exigence dans la jurisprudence offre une perspective prometteuse en matière d'IA de recrutement.

b) L'obligation de sûreté de la base de données, une obligation de moyens renforcée par essence

Poursuivant l'analogie avec l'obligation de sécurité informationnelle, la sûreté de la base de données d'IA de recrutement pourrait constituer également une obligation de moyens renforcée. Le biais de données provoquant souvent des décisions algorithmiques discriminatoires, la soumission de personnes candidates à l'IA de recrutement implique une exposition des candidats au risque de discriminations. Il y a donc bien un enjeu de sûreté de la base de données de l'IA de recrutement vis-à-vis du droit fondamental à l'égalité.

En outre, la définition d'une obligation de moyens renforcée de prévention de la contamination de la base de données permettrait de rééquilibrer la relation contractuelle entre les personnes candidates et les employeurs ayant recours à l'IA.

En effet, aussi bien l'UE que le Canada développent l'idée que le consentement[214] des individus ne constitue pas un critère légal suffisant en matière d'IA. Cette idée s'inscrit dans la continuité québécoise, autant que fédérale, des réflexions jurisprudentielles et doctrinales relatives au consentement dans le cadre des tests de dépistages de drogues[215]. Dans la perspective européenne, c'est le GT 29 précité qui soulève les lacunes du consentement[216]. Dans ses *Lignes*

[214] Les failles du consentement seront étudiées dans le chapitre 3, voir infra *La désuétude du consentement des personnes candidates à une sélection par IA.*

[215] Voir supra, la fin du chapitre 1, au paragraphe *Les exigences particulières en présence d'IA de recrutement* dans *2. La protection de la vie privée : une protection contre la décision automatique discriminatoire.*

[216] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 154. On retrouve en effet à la note 11 : « Dans plusieurs avis, le groupe de travail « Article 29» a exploré les limites du consentement dans des situations où il ne peut être donné librement. C'était notamment le cas de l'avis 15/2011 sur la définition du consentement (WP 187), […] de l'avis 8/2001 sur le traitement des données à caractère personnel dans le contexte professionnel (WP 48) […] » ; *Règlement Général sur la Protection des Données personnelles*, 88 (2016), 2016/679, Considérant 43 : «Pour garantir que le consentement est donné librement, il convient que celui-ci ne constitue pas un fondement juridique valable pour le traitement de

directrices sur le consentement au sens du règlement 2016/679, le GT29 (Groupe de Travail « Article 29 ») rappelle que si l'art. 22 RGPD dispose que la décision automatisée est soumise au consentement des individus concernés, dans certaines circonstances il ne serait pas pertinent de faire reposer la légalité de la décision sur le seul consentement des concernés[217].

De même, dans sa *Consultation sur les propositions du Commissariat visant à assurer une réglementation adéquate de l'intelligence artificielle,* le Commissaire à la vie privée consacre sa proposition 7 au même défi :

> « L'expérience démontre que le modèle de consentement actuel n'est peut-être pas viable dans toutes les situations, y compris pour certaines utilisations de l'IA. Cela s'explique en partie par l'incapacité d'obtenir un consentement valable lorsque les organisations ne peuvent informer les personnes des fins auxquelles leurs renseignements sont collectés, utilisés ou communiqués de façon suffisamment détaillée pour consentir de façon éclairée »[218].

Dans le contexte du recrutement, nous avons relevé au précédent chapitre la fragilité du consentement. Plus encore, nous avons établi que ce dernier n'était pas valide dans le cadre du recrutement. Dans la mesure où les employeurs organisent des processus de recrutement augmentés pour gagner du temps, on pourrait douter légitimement de leur ouverture ou de leur bonne foi lorsqu'une personne candidate refuse que l'IA puisse servir à l'examen de son profil professionnel. Cette vulnérabilité de la personne en demande d'emploi s'observe même très concrètement face à des IA de recrutement qui provoquent un inconfort tel qu'elles déstabilisent[219] la personne candidate. Dans la recherche de stratégie de gestion du risque de

---

données à caractère personnel dans un cas particulier lorsqu'il existe un déséquilibre manifeste entre la personne concernée et le responsable du traitement, en particulier lorsque le responsable du traitement est une autorité publique et qu'il est improbable que le consentement ait été donné librement au vu de toutes les circonstances de cette situation particulière. [...]»

[217] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 154. « L'adjectif « libre » implique un choix et un contrôle réel pour les personnes concernées. En règle générale, le RGPD dispose que si la personne concernée n'est pas véritablement en mesure d'exercer un choix, se sent contrainte de consentir ou subira des conséquences négatives importantes si elle ne donne pas son consentement, le consentement n'est pas valable ».

[218] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, *Consultation sur les propositions du Commissariat visant à assurer une réglementation adéquate de l'intelligence artificielle*, coll. Consultations, Commissariat à la protection de la vie privée du Canada, 2018, Proposition 7

[219] D. HARWELL, préc., note 82. L'article relate entre autres, l'anxiété nouvelle que vivent les personnes demandeuses d'emploi qui sont confrontées à l'IA de recrutement proposées par *HireVue*. « The inscrutable algorithms have forced job seekers to confront a new kind of interview anxiety. […] But when she didn't get the investment banking job, she

discrimination, recueillir le consentement a donc peu d'intérêt. La seule utilité d'un tel consentement serait au bénéficie de l'employeur qui pourrait se prévaloir du consentement des candidats vulnérables qui renonceraient donc à leur droit de refuser un processus d'embauche susceptible de les discriminer. Cependant, nous avons déjà relevé le manque de lucidité d'une telle configuration, l'alternative offerte par l'obligation de sûreté prend donc tout son sens pour contraindre l'employeur, en position de force, à respecter le droit à l'égalité des personnes qui postulent à son offre d'emploi.

Aussi, l'obligation de diligence renforcée appliquée à la sûreté des IA de recrutement entraînerait pour l'employeur la charge d'apporter la preuve qu'il ou elle a pris toutes les mesures raisonnables pour limiter le risque de discriminations algorithmiques. Il s'agirait d'une forme de rééquilibrage de la relation entre les candidats et l'employeur dans la mesure où les demandeurs d'emploi ne sont pas pleinement libres de refuser les processus de recrutement par IA. La charge du risque de discrimination par IA de recrutement constituerait ainsi une sorte de compensation du rapport de pouvoir inégalitaire entre l'employeur et la personne en recherche d'emploi.

Par ailleurs, la jurisprudence semble s'inscrire dans cette logique de renforcement de l'obligation de moyens et de transfert du fardeau de la preuve sur la partie forte de la relation contractuelle. Les deux affaires que nous verrons dans les prochains développements illustrent le renforcement de l'obligation de diligence : dans les deux affaires, l'entreprise est jugée responsable des dommages causés lorsqu'elle ne s'est pas dotée des technologies les plus sécuritaires pour les individus ou pour les biens.

En premier lieu, dans l'affaire *T.J Hooper* c**.** *Northern Barge*[220], la Cour fédérale d'appel du second circuit des États-Unis a jugé que lorsque le progrès technologique fournit les moyens d'assurer

---

couldn't see how the computer had rated her or ask how she could improve, and she agonized over what she had missed. Had she not looked friendly enough? Did she talk too loudly? What did the AI hiring system believe she had gotten wrong? ».

[220] *T.J Hooper* c*. Nothern Barge*, [1932] 60 F. 2d 737 (2d Cir.) [*The T.J Hooper*], en ligne : <https://heinonline.org/HOL/CaseLaw?1=1&native_id=7016937&collection=journals> (consulté le 20 octobre 2020).

une meilleure sécurité, le fait que les normes d'industrie n'encouragent pas l'adoption desdites technologies n'a pas pour effet d'atténuer l'obligation de diligence. Dans cette affaire, une livraison de charbon devait s'effectuer par voie maritime entre la Virginie et New York. Les navires transportant la cargaison ont été remorqués par le Hooper et le Morgan alors qu'une tempête était annoncée. Les vents violents ont eu raison d'un des navires et la cargaison a été perdue. Or, à l'époque, il était désormais possible d'obtenir deux bulletins météo au cours d'une journée. En y ayant accès, il aurait été possible d'éviter de prendre la mer avant la tempête. Seulement la plupart des transporteurs n'étaient pas équipés de radios bien que celles-ci puissent être acquises à un coût considéré raisonnable. La juridiction a donc retenu que les remorqueurs avaient manqué à leur obligation de diligence et de sécurité dès lors que l'aléa météorologique pouvait être mitigé par la technologie nouvellement disponible.

Cette décision présente un intérêt particulier puisqu'elle reconnaît l'existence d'une obligation de sécuriser la marchandise. Cette obligation est renforcée par la disponibilité d'une technologie capable de mitiger le risque existant (c'est-à-dire le risque d'intempéries). Le progrès technologique a donc accru l'obligation de sécurité pesant sur les transporteurs et remorqueurs. La jurisprudence a ici retenu que de tels moyens technologiques, s'ils sont accessibles, induisent une obligation pour le débiteur ou la débitrice de s'en saisir sans attendre une adoption généralisée.

Par conséquent, satisfaire à l'obligation de sécurité implique de recourir aux moyens existants, récents et accessibles de manière proactive. Dans la seconde espèce relative à la protection des données sensibles cette fois, les autorités compétentes abondent dans le même sens.

En second lieu, dans l'affaire *TJX/WMI*[221], informés d'une fuite de données bancaires de plus de 45 millions de personnes en septembre 2007, le Commissariat à la protection de la vie privée du Canada et le Commissariat à l'information et à la protection de la vie privée de l'Alberta ont ouvert une enquête conjointe visant à savoir si TJX/WMI avait manqué à ses obligations

[221] *Rapport d'enquête sur la sécurité, la collecte et la conservation des renseignements personnels TJX Companies Inc./Winners Merchant International L.P.*, 2007 Commissaire à la protection de la vie privée du Canada [*Rapport d'enquête TJX Companies Inc./Winners Merchant International L.P.*], en ligne : <http://canlii.ca/t/1t3tc> (consulté le 20 octobre 2020).

légales de protection des données sensibles. En l'espèce, un intrus avait infiltré le système informatique de l'entreprise, ayant ainsi obtenu accès aux renseignements personnels de la clientèle : leurs coordonnées (noms, prénoms, adresses, numéros de téléphone), cartes de crédit, permis de conduire et autres numéros d'identification gouvernementaux.

Or, au moment de la brèche informatique, la technique de chiffrement utilisée (WEP) était concurrencée par un mode de chiffrement nouveau (WPA) et jugé plus sécuritaire. Dans leur rapport d'enquête, les Commissariats précisent notamment que :

> « Il est essentiel que les organisations non seulement mettent en place de multiples couches de protection, mais aussi qu'elles suivent les progrès technologiques de manière à s'assurer que leurs mesures de protection ne sont pas périmées ou facilement contournables ».

Bien que les commissariats susvisés n'aient pas prononcé une pleine sanction envers l'entreprise[222], ils ont retenu deux griefs : l'absence de nécessité de collecter autant de données sensibles, et le défaut de mise à jour des systèmes de sécurité[223].

Autrement dit, les commissariats se sont interrogés dans un premier temps, sur la nécessité de collecter les données sensibles, au risque qu'elles fassent l'objet de fuites. Leur conclusion était que pour qu'un procédé corresponde aux exigences de sûreté, la collecte d'une telle quantité de données sensibles ne doit plus faire l'objet d'une simple mitigation, mais constitue un risque à éviter. Il s'agit des principes de nécessité et de minimisation des données : la gestion du risque de fuite de données sensibles commandait donc d'en collecter le moins possible, uniquement si elles étaient nécessaires. L'accumulation de données sensibles dispensables constitue donc une première méconnaissance de l'obligation de sécurité. Dans un second temps, l'enquête des commissariats s'est tournée vers les moyens mis en place dans la protection desdites données sensibles. Dans la continuité de l'affaire précédente, les institutions ont conclu que la sensibilité

[222] Les mesures correctrices avaient été prises en partie, les autres recommandations étaient acceptées par TJX/WMI. Néanmoins, l'excès de données sensibles collectées a conduit à une sévérité de la décision finale.

[223] Notons ici que le cumul de ces griefs a probablement joué dans la sévérité de la décision retenue. En effet, au moment du bris de sécurité des renseignements personnels, l'industrie bancaire n'avait pas encore migré vers le WPA, estimant ne pas avoir encore en éprouvé la robustesse. Cependant, compte tenu du nombre de renseignements sensibles conservés par TJX/WMI, les commissariats ont estimé que l'adoption de la technologie la plus fiable était nécessaire, même si la fiabilité n'était alors que théorique.

des données collectées appelait également une obligation de sécurité renforcée. En l'espèce, il s'agissait de se doter de la technologie de chiffrement la plus récente. Cette seconde exigence n'ayant pas été respectée, dans les deux affaires, la conclusion portait sur une violation de l'obligation de sécurité.

En outre, ces deux motifs de sanction dans l'affaire *TJX/WMI*, lus conjointement, peuvent être rapprochés des griefs formulés à l'encontre de *HireVue* et de ses IA de recrutement. En effet s'agissant des IA de *HireVue*, si d'aventure on les jugeait fiables[224], il serait légitime d'interroger la nécessité de collecter des informations aussi sensibles que la stabilité apparente d'une personne candidate, ou encore de ses émotions supposées au cours d'un entretien. De même, quelles seraient les mesures de nature à protéger les personnes candidates contre les biais et discriminations pouvant entacher un tel procédé de recrutement ? Quand bien même des mesures de protection renforcée seraient mises en œuvre, compte tenu de la sensibilité des informations produites et collectées par les IA de recrutement, à l'instar de l'affaire TJX/WMI, ne devrait-on pas préférer l'évitement à la mitigation du risque ? L'obligation de sécurité en contexte d'IA de recrutement ne serait-elle pas mieux respectée en évitant tout simplement de collecter de telles données ?

Enfin, on peut remarquer que dans les deux affaires examinées[225] la jurisprudence identifie l'obligation de protection comme constituant une obligation de diligence renforcée. Plus encore, dans les deux décisions, le progrès technologique renforce la proactivité attendue des entreprises

[224] Rappelons que parmi les nombreuses critiques formulées à l'encontre de l'IA de *Hirevue*, les scientifiques interrogés parlent de « pseudo science » quant à la capacité de l'IA à qualifier des émotions et des comportements, là où l'entreprise revendique une capacité de sa technologie à estimer si la personne reçue en entretient manifeste les signes d'une instabilité mentale. Drew HARWELL, « EPIC's FTC complaint about HireVue », *Washington Post*, sect. Technology (6 novembre 2019), en ligne : <https://www.washingtonpost.com/context/epic-s-ftc-complaint-about-hirevue/9797b738-e36a-4b7a-8936-667cf8748907/> (consulté le 3 novembre 2020); D. HARWELL, préc., note 188; D. HARWELL, préc., note 82.

[225] *The T.J Hooper*, préc., note 220; *Rapport d'enquête TJX Companies Inc./Winners Merchant International L.P.*, préc., note 221.

débitrices. Cela se manifeste aussi bien dans l'exigence de se munir d'une radio, que dans celle d'intégrer une technologie de chiffrement plus récente et réputée plus fiable[226].

En conclusion, les obligations de sécurité et de sûreté constituent des obligations de diligence accrues : en présence de moyens techniques technologiques évolutifs, l'entreprise aurait cette obligation étendue de se mettre à jour des moyens technologiques assurant une protection des individus selon les moyens les plus efficaces et actualisés. Dans le cas d'une IA de recrutement, la base de données demeurant le premier vecteur de discriminations à l'embauche, la protection des droits fondamentaux des candidats passe par la mitigation du biais de données[227]. L'application de l'obligation de diligence se matérialiserait ainsi par l'obligation de mitiger le risque de biais de données. En d'autres termes, toutes les méthodes accessibles permettant la réduction du risque de tels biais, ou permettant de les corriger en amont, devrait faire l'objet d'une application « raisonnable »[228] afin d'assurer la plus grande intégrité possible aux processus de recrutement augmentés.

2. La mitigation du biais de données par le renforcement d'obligations existantes

Après avoir démontré que l'obligation de moyens renforcée s'applique à l'IA de recrutement, nous verrons maintenant par quels outils l'obligation de moyens renforcée se matérialise dans une stratégie de mitigation du risque. Pour ce faire, nous étudierons les différents mécanismes

[226] Rappelons que dans les deux espèces considérées, le risque météorologique était évitable en s'équipant de radio, tandis que dans l'affaire TJX, le risque de fuite de données sensibles était considéré comme réduit en cumulant une limitation des données collectées, mais également en chiffrant les données sensibles nécessaires à l'aide du protocole de chiffrement le plus récent. *The T.J Hooper*, préc., note 220; *Rapport d'enquête TJX Companies Inc./Winners Merchant International L.P.*, préc., note 221.

[227] Mitigation du biais de données qui ne concernerait donc que les données nécessaires. Dans l'hypothèse de données non nécessaires qui seraient collectées, à l'instar des développements dans l'affaire TJX/WMI précitée, ce n'est pas la mitigation mais une approche d'évitement du risque qui serait retenue.

[228] La référence au caractère raisonnable des mesures permet de maintenir cette obligation dans la catégorie des obligations de diligence, et non de résultat. En effet, les organisations ne disposent pas toutes des ressources nécessaires pour se doter des technologies efficaces mais coûteuses. « Raisonnable » renvoie donc à une proportion du risque et des moyens mis en place pour limiter sa réalisation, tout en tenant compte des contraintes du débiteur de l'obligation de sûreté.

juridiques reposant sur l'obligation de diligence accrue, en conservant dans les prochains développements les exemples de *HireVue*, Google Photos et Amazon.

a) Le concept de responsabilité démontrable, un renforcement de l'obligation de documentation

La responsabilité démontrable forme un concept proposé par le Commissariat à la vie privée dans sa Consultation sur les propositions du Commissariat visant à assurer une réglementation adéquate de l'intelligence artificielle[229] : « [La] responsabilité démontrable exigerait des organisations qu'elles soient en mesure de prouver, sur demande, qu'elles se conforment aux exigences de la loi. »

La responsabilité démontrable constituerait une responsabilité accrue des organisations ayant recours à l'IA. En effet, le Commissariat précise les moyens de cette responsabilité démontrable :

> « Il existe diverses méthodes permettant d'obtenir une responsabilité démontrable, comme l'exigence de traçabilité, le droit à l'explication et l'évaluation des facteurs relatifs à la vie privée et aux droits de la personne, comme nous l'avons déjà mentionné. Une exigence en matière de tenue de dossiers serait également nécessaire pour permettre au Commissariat d'effectuer des inspections proactives plus facilement. »[230]

En substance, l'autorité régulatrice susmentionnée propose un modèle de contrôle et d'audit similaire à ce que l'on peut retrouver en finance :

> « Nous proposons que la loi exige également une vérification indépendante par un tiers tout au long du cycle de vie du système d'IA. Les vérificateurs pourraient être passibles de sanctions financières s'ils agissent avec négligence en approuvant des pratiques qui, en fait, ne sont pas conformes. »

[229] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, *Consultation sur les propositions du Commissariat visant à assurer une réglementation adéquate de l'intelligence artificielle*, coll. Consultations, Commissariat à la protection de la vie privée du Canada, 2018, proposition 10, en ligne : <https://www.priv.gc.ca/fr/a-propos-du-commissariat/ce-que-nous-faisons/consultations/consultation-ai/pos_ai_202001/> (consulté le 15 mars 2020).

[230] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101.

Il s'agit donc de créer de la documentation et des supports permettant les contrôles et vérifications de conformité de l'IA avec les textes protecteurs des droits fondamentaux et de la vie privée.

Cette notion de responsabilité démontrable devrait d'ailleurs faire l'objet d'une lecture conjointe avec la Directive sur la prise de décision automatisée du Conseil au Trésor du Canada[231]. En effet, la proposition 5 réfère explicitement à l'évaluation d'impact algorithmique non seulement comme outil de contrôle des biais des IA, mais également comme document central d'examen de la responsabilité démontrable d'un organisme. Une telle obligation, si elle était adoptée par le législateur fédéral, réaffirmerait la nécessité de documenter le cycle de vie des données[232], voire la définition d'une obligation de documentation exigeante.

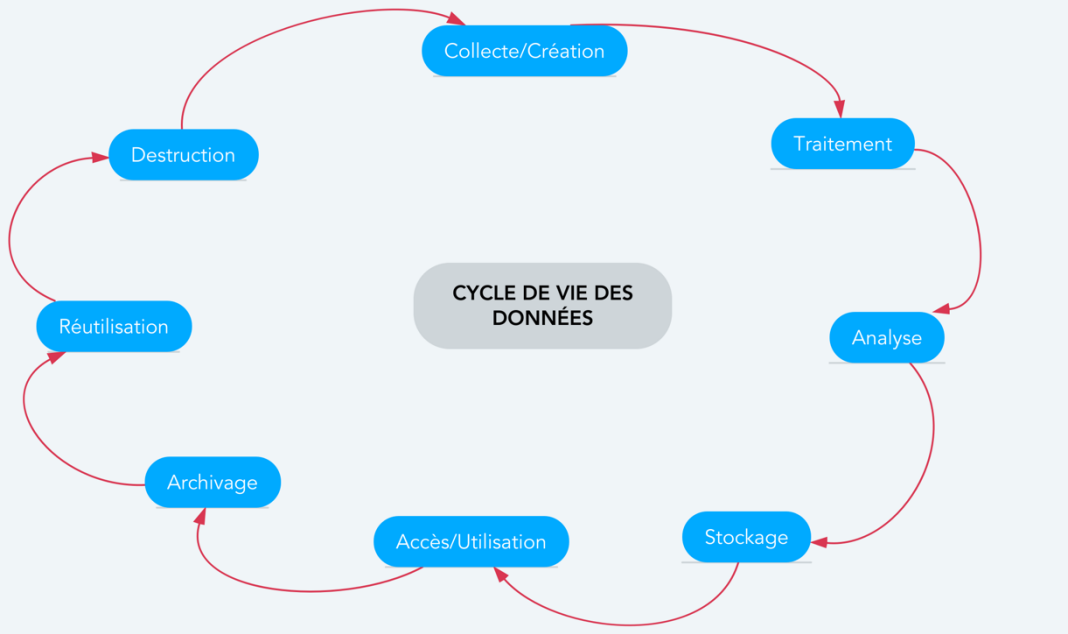

Figure 2. – Le cycle de vie des données

[231] *Directive sur la prise de décision automatisée*, préc., note 79. Cette directive contient une obligation de réaliser une évaluation de l'impact algorithmique des IA que le gouvernement et ses démembrements envisageraient de développer ou d'adopter.

[232] Figure 2 — Le cycle de vie des données.

En effet, la documentation constitue une pratique recommandée par les standards de gestion de la sécurité de l'information comme les normes ISO 27000[233]. La documentation consiste à renseigner le cycle de vie des informations collectées et traitées par une organisation. On retrouve ici l'idée de la responsabilité démontrable : par cette documentation, l'autorité compétente en matière de protection des informations sera en mesure de contrôler le respect de l'obligation de diligence. L'obligation de documentation consisterait donc à exiger d'une organisation de répertorier tous les actes par lesquels elle a assuré la confidentialité, la disponibilité et l'intégrité des informations qu'elle gère.

Dans le cas de l'IA de recrutement, le risque de discrimination à l'embauche entraînerait une exécution de cette obligation par l'évaluation d'impact algorithmique au cours des différentes phases de développement de l'IA. Plus spécifiquement, il s'agirait donc d'auditer la manière dont la base de données a été constituée pour assurer une réduction des risques de contaminations de l'IA par les biais cognitifs humains.

Par voie de conséquence, l'adoption de la responsabilité démontrable dans les législations d'encadrement des IA aurait pour effet de créer une obligation légale de documentation. Cette obligation légale de documentation constituerait une obligation de diligence renforcée.

En effet, le Commissariat à la vie privée préconise que l'évaluation d'impact algorithmique fasse l'objet d'un dépôt obligatoire auprès des autorités de protections de la vie privée. Ce dépôt serait rendu nécessaire en vertu de deux pouvoirs des autorités régulatrices : leur capacité à contrôler les IA développées, ainsi que la sanction des organisations négligentes quant aux impacts sur les droits fondamentaux des citoyens[234]. Le défaut de soumission de l'évaluation d'impact algorithmique exposerait ainsi l'organisation concernée à des sanctions pécuniaires[235].

---

[233] *ISO/IEC 27000*, 2018, en ligne : <https://www.iso.org/obp/ui/#iso:std:iso-iec:27000:ed-5:v1:fr> (consulté le 8 novembre 2020).

[234] Rappelons ici que le Commissariat prône la reconnaissance de la vie privée comme un droit fondamental, au même titre que le droit à l'égalité.

[235] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101. Dans la proposition 11, le commissariat préconise un accroissement de ses pouvoirs de sanction : « Parmi les améliorations nécessaires à la LPRPDE, mentionnons l'habilitation du Commissariat à rendre des ordonnances exécutoires et à imposer des sanctions corrélatives en cas de non-conformité à la loi. […] De véritables pouvoirs […] d'appliquer des sanctions financières

Appliquée au cas précité de l'entreprise *HireVue*, une obligation de documentation renforcée entraînerait notamment l'obligation pour l'entreprise de soumettre à l'autorité protectrice de la vie privée compétente une évaluation d'impact algorithmique de son IA d'analyse des caractéristiques comportementales des personnes candidates à un emploi. De même, l'opacité choisie par *HireVue* l'exposerait à des sanctions pécuniaires.

Se conformer à l'exigence d'une évaluation d'impact algorithmique constituerait donc un moyen d'exiger une gestion du risque de discrimination par les entreprises utilisatrices des IA de recrutement. Autrement dit, avec une obligation d'évaluation d'impact algorithmique, des compagnies comme *HireVue* — et leurs entreprises clientes — auraient l'obligation de se soumettre à cet exercice d'évaluation et de mitigation des risques de discrimination par leur IA de recrutement.

En outre, parallèlement à l'exigence de documentation, la proactivité exigée dans la mitigation des biais algorithmiques pourrait constituer le fondement d'une obligation de sûreté étendue.

b) L'obligation de sûreté renforcée, l'impératif d'une adoption proactive des moyens naissants de mitigation du risque de biais dans les IA de recrutement

Comme cela a été vu précédemment[236], l'obligation de diligence accrue implique une exigence de proactivité, d'anticipation et d'efficacité dans l'adoption de moyens raisonnables de mitigation des risques des processus de recrutement augmentés. Nous avons vu que l'obligation de documentation réaffirmée représentait un des moyens de mitigation du risque, néanmoins d'autres solutions existent. Nous verrons que ces dernières pourraient entrer dans le spectre de l'obligation de sûreté rattachée aux IA de recrutement.

Le *data cleansing*, ou « nettoyage de données » est défini par l'Office québécois de la langue française comme constituant :

permettraient aux Canadiens d'obtenir des règlements plus rapidement et les rassureraient quant à leur capacité de participer en toute confiance au marché numérique. »

[236] Voir supra, b*) L'obligation de sûreté de la base de données, une obligation de moyens renforcée par essence.*

> « [L'] opération par laquelle sont détectées, corrigées, remplacées ou éliminées les données d'un ensemble stocké, qui ne sont pas conformes à certaines règles, qui présentent des anomalies de forme ou qui sont incomplètes, de manière à assurer leur cohérence. »[237]

Dans le cas Google Photo vu précédemment, le nettoyage des données aurait consisté notamment à retirer les images de personnes africaines ou afrodescendantes identifiées comme des gorilles. En d'autres termes, le nettoyage des données permet de corriger la base de données en la purifiant de ses données incohérentes, fausses ou erronées.

Pour autant, si dans le cas de Google Photos d'aucuns peuvent considérer un tel procédé accessible, il n'est pas difficile d'imaginer la complexité de l'exercice dans le cas d'Amazon. En effet, rappelons que la base de données d'entraînement était biaisée, de telle sorte que cela s'est répercuté sur l'apprentissage machine[238]. Dans le cas d'Amazon le nettoyage des données n'aurait donc pas été approprié. Il aurait peut-être fallu se tourner vers une autre méthode : le *fair learning*.

Le *fair learning*, ou apprentissage juste, consisterait à réduire les angles morts de l'apprentissage machine en assurant une diversité des données d'apprentissage. On retrouve notamment cette idée à travers les principes d'équité et d'inclusion de la dite diversité[239].

---

[237] « nettoyage de données », dans Office Québécois de la Langue Française, coll. Vocabulaire du traitement de données, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8873874> (consulté le 27 avril 2020).

[238] Pour rappel, les biais se sont nichés dans le vocabulaire appris puis recherché par l'IA d'Amazon. Les corriger appellerait donc une correction des étiquettes associées à chaque terme que l'IA aurait associé au genre masculin.

[239] *La Déclaration de Montréal pour un développement responsable de l'Intelligence Artificielle*, Montréal, part. 6 et 7, en ligne : <https://www.declarationmontreal-iaresponsable.com/la-declaration>.

« 6 PRINCIPE D'ÉQUITÉ

1. Les SIA doivent être conçus et entraînés de sorte à ne pas créer, renforcer ou reproduire des discriminations fondées entre autres sur les différences sociales, sexuelles, ethniques, culturelles et religieuses.

2. Le développement des SIA doit contribuer à éliminer les relations de domination entre les personnes et les groupes fondées sur la différence de pouvoir, de richesses ou de connaissance. […]».

« 7 PRINCIPE D'INCLUSION DE LA DIVERSITÉ

1. Le développement et l'utilisation de SIA ne devraient pas conduire à une uniformisation de la société par la normalisation des comportements et des opinions.

Concernant le cas d'Amazon il aurait donc fallu constituer une base de données regroupant des profils de demandeurs d'emploi pluriels et variés, et non se contenter des profils des salariés déjà embauchés par ce géant du numérique. En effet, si l'un des intérêts supposés de l'IA de recrutement consiste à s'affranchir des biais humains[240], limiter au maximum la transmission de ceux-ci à la machine constitue un enjeu crucial et corolaire. On peut remarquer à ce propos que le manque d'inclusion et de diversité des données d'apprentissage des IA de recrutement a mené à une prise de conscience des limites de l'outil dans le secteur du recrutement[241]. Plus la composition de la base de données sera inclusive, et moins on retrouvera d'angles morts dans les décisions de l'IA de recrutement. Par conséquent, dans le cas d'Amazon, le biais sexiste de l'IA de recrutement reflète uniquement le manque d'inclusion des équipes de l'entreprise.

Au sujet de l'incidence du manque de diversité des données d'apprentissage, un rapprochement avec l'IA de reconnaissance faciale permet de prendre la mesure de l'incidence du manque de diversification dans les données d'apprentissage.

Récemment, le *National Institute of Standards and Technology* (ci-après « NIST ») a publié son rapport d'analyse portant sur le taux d'exactitude de 189 IA de reconnaissance faciale, produites par plus de 99 développeurs différents[242].

---

2. Le développement et le déploiement des SIA doivent prendre en considération les multiples expressions des diversités sociales et culturelles, et cela dès la conception des algorithmes.

3. Les milieux de développement de l'IA, aussi bien dans la recherche que dans l'industrie, doivent être inclusifs et refléter la diversité des individus et des groupes de la société. […] »

[240] Kimberly HOUSER, *Can AI Solve the Diversity Problem in the Tech Industry? Mitigating Noise and Bias in Employment Decision-Making*, SSRN Scholarly Paper, ID 3344751, Rochester, NY, Social Science Research Network, 2019, en ligne : <https://papers.ssrn.com/abstract=3344751> (consulté le 31 août 2019). « […] AI can improve employment decision-making, but also can be used to root out and correct discriminatory results in AI stemming from human biases.» ; R. DEMICHELIS, préc., note 16. En 2019, la Directrice d'acquisition de Talents chez L'Oréal France indiquait : « L'[IA] nous permet d'avoir une plus grande diversité académique, puisqu'il permet d'évaluer les candidatures au-delà du CV […]. Sinon, face à de gros volumes de candidatures, un recruteur risque d'aller à la facilité et ne retenir que les étudiants de quelques écoles. Ce sont dans ces situations que les biais les plus importants apparaissent. »

[241] R. DEMICHELIS, préc., note 16. La Directrice d'acquisition de Talents de L'Oréal précisait encore « On a passé la phase où on a tous cru que l'IA allait être parfaitement objective. On met en place des garde-fous ; au moment du lancement de l'outil pour chaque pays, les recruteurs vont systématiquement évaluer les candidatures ayant les scores les plus élevés et les moins élevés pour s'assurer de la pertinence de la prédiction. »

[242] Patrick GROTHER, Mei NGAN et Kayee HANAOKA, *Face recognition vendor test part 3: demographic effects*, NIST IR 8280, Gaithersburg, MD, National Institute of Standards and Technology, 2019, DOI : 10.6028/NIST.IR.8280.

Les tests ont porté sur deux procédés de reconnaissance faciale : la comparaison de deux images « one to one search », et la reconnaissance d'une image parmi une multitude d'autres, « one to many search ». L'intérêt de distinguer ici ces deux procédés de reconnaissance réside dans l'effet des erreurs. Les résultats de l'étude identifient deux types de biais au détriment des groupes racialisés.

> « "In a one-to-one search, a false negative might be merely an inconvenience — you can't get into your phone, but the issue can usually be remediated by a second attempt," Grother said. "But a false positive in a one-to-many search puts an incorrect match on a list of candidates that warrant further scrutiny." ».[243]

Le premier biais apparaît dans le *one-to-one match* : les hommes Asiatiques-Américains et les Afro-américains, sont plus sujets aux faux positifs comparés aux Américains blancs. Autrement dit, les erreurs des IA de reconnaissance faciale conduisent plus souvent à une identification à tort des personnes racialisées. Plus encore, les technologies américaines reflètent des biais encore plus importants à l'égard des personnes autochtones.

De même, le deuxième biais concerne les femmes afro-américaines, bien plus touchées par les faux positifs dans le procédé *one-to-many*. De tels résultats peuvent particulièrement inquiéter lorsque l'on sait que la reconnaissance faciale fait l'objet d'un intérêt — et d'une application très concrète — dans les activités policière et judiciaire. En vertu des résultats de l'étude du NIST, il n'est pas exagéré de craindre que des IA biaisées puissent aboutir à l'identification à tort de personnes suspectes, et ce, uniquement parce que les algorithmes fonctionnent moins bien sur leurs groupes sociaux, *a fortiori* lorsqu'il s'agit de femmes.

Ici également, ces discriminations involontaires des algorithmes de reconnaissance faciale résultent d'un entraînement de l'algorithme fondé sur des données d'apprentissage manquant d'hétérogénéité. Lorsque les IA de reconnaissance faciale sont entraînées sur les données des forces de l'ordre, ou des tribunaux, les mécanismes de discriminations systémiques qui ont cours

[243] Ici sont comparé l'effet d'une erreur de l'IA de reconnaissance faciale pour déverrouiller son téléphone, et celle d'une IA censée identifier des suspects. NIST, préc., note 17.

au sein de ces institutions, sont transmis aux algorithmes qui, à leur tour, les reproduisent. C'est notamment ce qu'a pu révéler ProPublica dans son rapport de 2016[244].

À la lumière des discriminations subséquentes aux biais des IA, l'expérimentation en laboratoire semble significativement différer d'un déploiement dans des circonstances réelles. Nonobstant les efforts des développeurs, et la méticulosité de la phase de test de l'algorithme, il serait légitime de s'interroger : en reconnaissant cette marge d'erreur des algorithmes, n'est-il pas possible de les corriger au fur et à mesure de leur utilisation ? Si la réponse à cette question repose principalement sur l'état de la science, il semble que les risques de discriminations demeurent difficiles à corriger au stade du déploiement de la technologie. Bien entendu, alors qu'une telle stratégie de mitigation des risques *a posteriori* pose des enjeux éthiques non négligeables[245], il s'agit surtout de constater la grande difficulté à corriger un algorithme déjà biaisé.

Désapprendre un biais semble complexe au point qu'Amazon et Microsoft aient préféré mettre un terme à leurs expériences. En effet, dans le cas de l'IA de recrutement d'Amazon, après tentative, l'entreprise a annoncé l'arrêt de ses travaux pour corriger son IA de recrutement et l'abandon du projet. Cela rappelle le cas de Tay de Microsoft : après une pollution massive de la base de données, il semble que les corrections n'aient pas été envisageables.

Dès lors que peut-on retenir de ces exemples ? En l'état actuel de l'art, il est impératif de faire preuve d'une grande rigueur au cours de la constitution de la base de données. Cette rigueur assurerait à l'IA une qualité des résultats, ainsi qu'un risque de biais réduit. Nous avons précédemment démontré qu'un bon échantillon de données d'apprentissage ne repose pas uniquement sur la masse de données, mais également (surtout) sur leur qualité. Or, si le big data

[244] J. ANGWIN, J. LARSON, S. MATTU, L. KIRCHNER, et PROPUBLICA, préc., note 17.

[245] En effet, faire le choix de déployer des IA dont on n'aurait pas contrôlé et corrigé les biais *a priori*, engendrerait un risque de discrimination accru pour les personnes minorisées ou discriminées. Dans la mesure où on sait que les personnes dites majoritaires ne sont les plus sujettes aux conséquences négatives des biais algorithmiques, cela reviendrait à exposer sciemment des personnes racialisées et LGBTQI+ par exemple. Ces personnes seraient donc dans la position de cobayes involontaires dans le but d'améliorer une solution technologique après son déploiement. Une telle approche supposerait donc que, selon l'utilisation de l'IA, on privilégierait le potentiel d'amélioration *a posteriori*, à la préservation des personnes face aux risques de discrimination.

répond à l'impératif quantitatif, s'agissant de la qualité en revanche, il semble que des données aussi diverses que possibles permettent d'assurer à la base de données une qualité renforcée. Dans le cas d'Amazon, il aurait fallu par exemple constituer une base de données paritaire, comportant des proportions représentatives — voire équivalentes — de personnes dites minorisées, sujettes aux oppressions systémiques qui opèrent dans nos sociétés.

Ainsi, entraîner les IA de recrutement avec des données les plus diverses possibles, leur assure un meilleur apprentissage ; couvrant par la même une multitude de situations que l'IA de recrutement apprendra à lire et à appréhender de manière plus équitable. Une base de données d'apprentissage composée d'autant de candidatures de personnes racialisées, que de personnes blanches, d'autant d'hommes que de femmes, de personnes valides que de personnes en situation de handicap, de personnes jeunes que de personnes plus âgées, une telle base présentera moins de risques discriminatoires que les bases de données monochromes, homogènes, ne reflétant qu'un seul type de profil. La diversité des données constitue donc un premier levier qualitatif des IA de recrutement. Nonobstant la reconnaissance de cet impératif qualitatif des données d'apprentissage, la question de savoir comment y répondre demeure. Il convient de relever, qu'à l'instar du principe de *privacy by design*[246], il serait probablement pertinent de se diriger vers un principe de *fair by design* en matière d'IA, *a fortiori* dans le recrutement. Le principe de *fair by design* aurait pour conséquence une appréhension des enjeux de biais discriminatoires dès la conception de l'IA. Autrement dit, l'élaboration de l'échantillon de données d'apprentissage s'opèrera en y intégrant les enjeux d'inclusion et de diversité.

Cette idée concorde par ailleurs, avec une piste de solution proposée par Cathy O'Neil. Cette dernière propose que les *data scientist*[247] soient soumis à un pendant du serment

[246] *Privacy by design* est un principe selon lequel dès la conception des algorithmes et des IA, les concepteurs de l'outil incluraient la gestion des renseignements personnels dans l'architecture (le code) qu'ils développent. Il s'agit donc de travailler à la conformité de l'IA avec les législations de protection des renseignements personnels dès la conception.

[247] OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, « Expert en science des données », dans Grand dictionnaire terminologique, coll. Informatique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26527138>

d'Hippocrate[248]. En d'autres termes, un engagement éthique et déontologique serait partie intégrante de la formation et de la pratique des *data scientist*, premiers acteurs des bases de données d'apprentissage des IA. Un tel engagement éthique serait ainsi une façon de traiter le risque de biais directement au niveau des personnes qui constituent les bases de données d'entraînement, ou qui conçoivent les algorithmes. Suivant la comparaison avec les médecins, cet engagement engagerait la responsabilité professionnelle et pénale des *data scientist*.

Finalement, on pourrait conclure que l'obligation de moyens renforcée présente une utilité non négligeable dans la gestion du risque de discrimination des processus de recrutement augmentés. Qu'il s'agisse d'obligation de documentation — à travers l'évaluation d'impact algorithmique ou de l'EFVP[249] —, de l'obligation de sûreté des bases de données, ou encore des autres méthodes émergentes de mitigations du risque de biais de données (nettoyage et diversité des données), ces différents outils peuvent s'intégrer aux pratiques des organisations même en l'absence de législation. Par ailleurs, l'ensemble de ces procédés adossés à l'obligation de moyens renforcée entraîne nécessairement une supervision humaine de l'IA de recrutement utilisée. Ainsi, les erreurs de l'algorithme pourraient faire l'objet de contrôles et corrections par un être humain. À défaut, et parce que l'humain est lui aussi influencé par ses propres biais, l'imputabilité de la décision à la personne responsable du recrutement ouvre au moins la voie à la neutralisation des biais du processus *a posteriori*.

Toutefois, une question demeure : comment est appréhendé le contentieux impliquant l'IA de recrutement ? Nous verrons dans un dernier chapitre la problématique de l'accès à la justice pour les personnes demandeuses d'emploi lésées par les processus de recrutement augmentés.

(consulté le 10 novembre 2020). « Personne spécialisée dans l'exploration, l'analyse et l'interprétation des données, et qui a pour tâche d'orienter les actions et les prises de décisions d'une organisation. »

[248] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016, ch. Conclusion. « How do we start to regulate the mathematical models that run more and more of our lives ? I would suggest that the process begin with the modelers themselves. Like doctors, data scientist should pledge a Hippocratic Oath, one that focuses on the possible misuses and misinterpretations of their models ». Nous y reviendrons plus en détail dans le prochain chapitre.

[249] Évaluation des facteurs de la vie privée

CHAPITRE 3. LES ENJEUX DE L'ACCÈS À LA JUSTICE EN CONTEXTE D'IA DE RECRUTEMENT

De l'émergence d'un cadre juridique applicable, en passant par des mécanismes juridiques facilitant la gestion du risque, les précédents chapitres nous offrent une vision relativement globale des marges de manœuvre juridiques en matière de gestion du risque de discrimination des IA de recrutement.

Pour autant, l'absence de dispositions légales spécifiques à l'IA soulève la question de l'accès à la justice. En d'autres termes, face à une IA de recrutement, comment peut-on faire valoir son droit à l'égalité ? Comment contester la décision algorithmique discriminatoire ?

On ne peut aborder la sanction juridique d'une discrimination à l'embauche par IA, sans identifier une source majeure de difficultés : le manque de transparence de l'IA de recrutement.

Or, en l'absence d'une législation globalement adaptée à l'IA, nous verrons que la personne écartée se heurte à plusieurs obstacles pour faire valoir ses droits. Ainsi, dans ce dernier chapitre nous nous consacrerons à l'effet entravant de l'opacité de l'IA sur la contestation du candidat écarté (A), avant de présenter des pistes d'évolution du cadre juridique applicable à l'IA de recrutement, et à l'IA en général (B).

A. L'opacité de la décision algorithmique : une source d'obstacles à la contestation des rejets de candidatures discriminatoires

Comme cela fut évoqué précédemment, la candidature écartée par une IA biaisée constitue une atteinte au droit fondamental à l'égalité. Dès lors, la personne écartée de façon discriminatoire devrait être en mesure de saisir une juridiction ou une autorité administrative compétente afin que soit sanctionnée cette atteinte.

L'accès à la justice dans un tel contexte implique donc la question de la résilience et de l'efficience des normes juridiques applicables à l'IA de recrutement. Plus encore, ce sont les normes

juridiques qui protègent le droit à l'égalité qui nous intéressent. Or, tant par son contexte d'application — le recrutement — que par sa nature complexe, l'IA de recrutement se caractérise par un manque de transparence : les critères à l'origine de la décision d'accueillir ou d'écarter un profil demeurent inconnus. Si dans le recrutement sans IA, la preuve d'une discrimination à l'embauche représente déjà un défi, face à une IA l'opacité de la décision algorithmique d'écarter une candidature soulève davantage de questionnements juridiques.

Premièrement en effet, il est important de rappeler que la personne candidate se soumet à un traitement automatisé de ses renseignements personnels. Ceci implique donc, en théorie, son consentement à cette technologie et à ses effets[250].

Deuxièmement, le traitement automatisé susmentionné donne lieu à une décision algorithmique déterminant si une candidature est retenue ou bien classée, ou non. La contestation de cette décision implique donc l'identification du motif de la discrimination alléguée : s'agit-il d'une discrimination à l'embauche fondée sur le genre, la race, l'origine, la couleur des yeux, le timbre de la voix, le lieu de résidence ? C'est en fonction du motif de discrimination qu'il sera possible de prouver l'existence d'un traitement automatique différencié, et illégal, de la candidature écartée.

Qu'il s'agisse du recueil du consentement ou de la motivation des décisions de rejet de candidature, ces deux mécanismes permettraient donc, d'une part, de s'assurer *a minima* que la personne candidate accepte le risque inhérent à l'IA de recrutement, et d'autre part, l'identification des biais discriminatoires qui affecteraient la décision de tri. Il s'agit donc là de deux mécanismes protégeant à la fois le droit à la vie privée[251] et le droit à l'égalité des candidats soumis à une IA de recrutement. Or ces deux mécanismes, *a priori* protecteurs, sont tributaires du niveau de transparence du processus augmenté.

---

[250] *RGPD*, préc., note 88, art. 22; COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, part. III. Intérêts commerciaux légitimes; *Loi n°64*, préc., note 20, art. 14; *Projet de loi C-11*, préc., note 20, art. 15(1). « 15 (1) Sauf disposition contraire de la présente loi, l'organisation qui recueille, utilise ou communique des renseignements personnels doit d'abord obtenir le consentement valide de l'individu concerné. »

[251] Rappelons ici que le Commissariat à la vie privée revendique depuis plusieurs années maintenant l'érection du droit à la vie privée au rang de droit fondamental, en ce qu'il conditionne le respect aux autres droits fondamentaux. L'atteinte au droit à la vie privée dans le contexte des technologies de l'information, conduit donc à l'atteinte d'autres droits fondamentaux, notamment le droit à l'égalité en matière de recrutement.

Dans la gestion de risques discriminatoires d'un processus de recrutement par IA, la transparence favorise l'identification des tenants d'une décision de rejet, et par voie de conséquence, le décèlement de potentielles discriminations. À l'inverse de la transparence, l'opacité de l'IA — voire son invisibilité — pour les postulants nuit à leur capacité de contester une décision de l'IA de recrutement. L'accès à la justice des personnes candidates à un emploi dépend donc de la transparence de l'IA de recrutement. On comprend donc que le consentement, autant que la motivation des décisions algorithmiques de tri de candidatures, est directement affecté par le niveau d'opacité du processus augmenté. En effet, plus ce dernier sera opaque, et moins lesdits mécanismes pourront produire leurs effets.

Voyons donc brièvement les débats que suscite la notion de transparence en IA (1), avant de nous attarder sur les effets de l'opacité de l'IA de recrutement sur ces deux garde-fous théoriques : le consentement (2) et la motivation des décisions de tri des candidatures (3).

1. La notion de transparence dans le domaine de l'IA

La définition courante de la transparence désigne « ce qui est visible par tous »[252]. Toutefois la notion revêt des acceptions différentes selon le domaine concerné.

Dans un contexte organisationnel, la transparence renvoie ainsi à la « qualité d'une organisation qui informe sur son fonctionnement, ses pratiques, ses intentions, ses objectifs et ses résultats. »[253] Cette dernière définition correspond davantage à un contexte de recours à l'IA. Dans cette matière, le principe de transparence de l'IA implique ainsi celui de l'explicabilité des résultats des algorithmes. À ce titre, l'explicabilité comporte deux volets : l'explication technique et l'explication due aux personnes affectées par une IA. Selon le Groupe d'experts constitué par la Commission européenne :

> « L'explicabilité concerne la capacité d'expliquer à la fois les processus techniques d'un système d'IA et les décisions humaines qui s'y rapportent (par exemple, domaines

[252] « Transparence », dans Le Petit Robert.

[253] « Transparence », dans Grand dictionnaire terminologique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8361425> (consulté le 20 novembre 2020).

> d'application d'un système d'IA). L'explicabilité technique suppose que les décisions prises par un système d'IA puissent être comprises et retracées par des êtres humains. »

Or, les attentes en matière de transparence diffèrent en fonction du public auquel on s'adresse. Ainsi, par la transparence des systèmes d'IA, les régulateurs vérifient la conformité des organisations aux normes en vigueur, les programmeurs d'IA recherchent des indicateurs de performance de leurs algorithmes, et les citoyens veulent une preuve de fiabilité de ces technologies[254]. Il y a donc autant de réponses à la transparence que de parties prenantes. Dès lors, toute explication[255] ne répond pas nécessairement aux attentes de tous les publics. Pour autant, nous retiendrons la définition précitée dans le contexte organisationnel. Non seulement cette acception permet d'intégrer la dimension organisationnelle du recrutement, mais elle met également l'accent sur l'importance de l'information. Or, comme nous le verrons ultérieurement, l'information constitue une composante essentielle du consentement valide.

2. L'IA de recrutement et le consentement : une incompatibilité intrinsèque

Hubert Reid définit le consentement comme constituant « l'accord donné par une personne à une proposition »[256]. Seul le consentement libre et éclairé répond à l'exigence de validité[257] ; or, les recruteurs ne sont pas soumis à une obligation de divulguer leurs techniques de recrutement auprès des candidats, y compris lorsqu'ils recourent à une IA pour trier les candidatures. Ce postulat entraîne donc deux situations distinctes concernant le consentement

[254] Michael RIDLEY, *Explainable AI: Explanations, Expectations, and Options*, Montréal, Graduate Law and Artificial Intelligence Conference - Fostering Empowerment through Artificial Intelligence: Where Do We Go from Here?, 25 février 2019, en ligne : <https://www.cyberjustice.ca/2019/01/22/graduate-law-and-artificial-intelligence-conference-fostering-empowerment-through-artificial-intelligence-where-do-we-go-from-here/> (consulté le 20 novembre 2020).

[255] Les termes transparence, explicabilité et explication seront utilisés comme synonymes.

[256] H. REID et S. REID, préc., note 25, « Consentement ».

[257] *Code civil du Québec*, L.Q. 1991, c. 64, art. 1399 [*C.c.Q*]; GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 154, p. 6.

des personnes candidates : soit le défaut d'information des candidats entrave le consentement, soit son caractère contraint l'invalide.

a) Le consentement entravé par le défaut d'information

Nous vivons une époque de numérisation accrue du marché du travail. La recherche d'emploi, comme la recherche de talents, s'effectue de plus en plus par l'intermédiaire de plateformes ou de réseaux sociaux professionnels. Lorsqu'une personne répond à une offre d'emploi en ligne, il n'y a aucune garantie de recevoir de l'information sur l'utilisation d'une ou plusieurs technologies de recrutement. En réalité, hormis dans le cas d'entretiens virtuels automatisés, le tri des candidatures peut s'effectuer par IA à l'insu des personnes concernées.

Par exemple, il est possible que l'IA opère un tri des candidatures en fonction de mots clés. Il est également possible que l'IA classe les candidatures en fonction de leur proximité avec le modèle appris par l'algorithme[258]. Dans tous les cas, la sophistication de la technologie, aussi bien que ses effets, demeure invisible aux personnes candidates.

Or, sans information, l'une des deux conditions du consentement valide fait défaut. En l'occurrence cela tombe sous le sens : si une personne candidate X ignore qu'elle est soumise à une IA de recrutement, comment pourrait-elle y consentir ? L'exigence de recueil du consentement sous-tend donc celle d'informer les personnes candidates. De plus, par l'extension des notions de traitement et de décision automatisés, l'obligation d'informer les personnes candidates pèserait sur les employeurs recourant aux IA de recrutement. En effet, au Canada comme dans l'UE[259], l'obligation d'information suppose une communication de l'existence de l'IA dans le processus de recrutement.

> « Nous recommandons aussi que les individus se voient garantir deux droits explicites en ce qui concerne la prise de décision automatisée, à savoir : le droit d'obtenir une explication valable justifiant les décisions prises à leur endroit découlant de la prise de

[258] C'est de ce type d'IA de recrutement dont il s'agit dans le cas Amazon étudié précédemment.

[259] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 86, p. 35, Annexe 1.

> décision automatisée, et le droit de contester ces décisions sous le régime de la LPRPDE. »[260]

Cette obligation d'information permettrait donc de mettre un terme à une pratique qui a pour effet de contourner le consentement des candidats en les gardant dans l'ignorance.

Par ailleurs, il faut également mentionner une autre hypothèse dans laquelle le caractère éclairé du consentement interroge : si les personnes candidates sont informées de l'existence d'une IA de tri, mais qu'elles n'en comprennent pas les enjeux, alors le consentement n'est toujours pas considéré comme étant éclairé.

Ainsi, lorsqu'il est clair qu'une IA est impliquée dans le recrutement[261], le recueil du consentement implique la transmission d'une information intelligible aux postulants. À ce titre, il faut rappeler que l'obligation d'information serait satisfaite en précisant d'une part l'utilisation d'une IA dans le processus d'embauche, mais également en assurant une compréhension du fonctionnement et des effets de l'IA utilisée. Cette exigence constituerait par ailleurs un facteur déterminant de l'exercice du droit de contester la décision automatisée[262].

> « [L] es personnes devraient disposer du droit de contester les décisions découlant de la prise de décision automatisée. Cela s'appliquerait aussi bien aux cas où une personne a donné son consentement au traitement de ses renseignements personnels qu'à ceux où une exception aux exigences du consentement a été invoquée par l'organisation. Il s'agit d'un prolongement du droit d'obtenir une explication. »[263]

L'obligation d'information, qui se retrouve également dans l'article 12 du RGPD[264], fait l'objet d'un renforcement : plus qu'informer sur l'existence et la nature de la technologie utilisée, il est

---

[260] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, part. « Dispositions propres à la prise de décision automatisée ».

[261] C'est notamment le cas si les candidats interagissent avec des IA à l'occasion d'entretiens téléphoniques, tel que Véra, ou vidéo, comme avec *HireVue*. Voir supra *chapitre 2, A., b) L'étiquetage, un vecteur insidieux de biais discriminatoires*.

[262] Un tel droit d'opposition constitue le pendant du consentement libre : si je suis libre de consentir à un processus de recrutement, je suis également libre de m'y opposer. La liberté du consentement ne saurait donc exister sans droit d'opposition. Les problématiques relatives à la liberté du consentement sera étudiée dans le c).

[263] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, part. II. Un droit à la contestation.

[264] *Règlement Général sur la Protection des Données personnelles*, 88 (2016), 2016/679, art. 12, en ligne : <https://eur-lex.europa.eu/eli/reg/2016/679/oj> (consulté le 20 février 2020). « 1. Le responsable du traitement prend des mesures appropriées pour fournir toute information visée aux articles 13 et 14 ainsi que pour procéder à toute communication au titre des articles 15 à 22 et de l'article 34 en ce qui concerne le traitement à la personne

également demandé au responsable du traitement d'expliquer la logique, et les conséquences envisagées de ce processus augmenté. On pourrait y voir le pendant de l'obligation d'information en droit européen de la consommation : la personne professionnelle est tenue de fournir une information complète et accessible aux consommateurs qui, par définition, ne disposent pas d'une expertise suffisante pour comprendre les conséquences de leurs achats.

Ainsi, dans le contexte du recrutement augmenté, les employeurs auraient d'abord l'obligation d'informer les personnes candidates à l'embauche afin de recueillir leur consentement éclairé. Néanmoins, pour que le consentement soit valide, la délivrance de l'information peut bien respecter une intelligibilité raisonnable, encore faut-il que la personne dispose de sa liberté de refuser le traitement de sa candidature par une IA.

b) Le contexte de l'embauche : une contradiction manifeste avec un consentement libre (l'exemple des tests de dépistage de drogue)

> « Le consentement doit être libre et éclairé.
>
> Il peut être vicié par l'erreur, la crainte ou la lésion. »[265]

« [L] ibre et éclairé » ; ces deux caractéristiques cumulatives du consentement, conduisent à son invalidité dès lors que l'une d'elle fait défaut. Les vices du consentement ont d'ailleurs vocation à invalider le consentement obtenu par le biais de l'erreur, de la crainte, ou encore de la lésion. La liberté du consentement amène donc également celle de ne pas consentir.

À ce propos, le Canada reconnaît l'importance d'un droit au refus de se soumettre à un processus décisionnel algorithmique. Le droit à l'opposition fait en effet partie des recommandations de l'autorité protectrice de la vie privée[266]. De plus, la reconnaissance d'un tel droit signifie, a

---

concernée d'une façon concise, transparente, compréhensible et aisément accessible, en des termes clairs et simples […] »

[265] QUÉBEC (PROVINCE), Alain ROY, Benoît MOORE, Élise M CHARPENTIER et Sébastien LANCTÔT, *Code civil du Québec: annotations, commentaires*, (1991), L.Q, art. 1399 [*Code civil du Québec*].

[266] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101, Proposition 3; COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, sect. « II. Le droit à la contestation ».

contrario, que le recours à une décision algorithmique est subordonné au consentement des personnes qui y seront soumises.

> « Si nous voulons protéger efficacement la vie privée en tant que droit de la personne dans un contexte numérique où des systèmes d'IA sont actifs, l'un des droits qui doit être envisagé est la capacité de s'opposer aux décisions prises par les ordinateurs et de demander une intervention humaine. Les lois de plusieurs pays prévoient le droit de ne pas être soumis à la prise de décision automatisée, ou un droit analogue de contester le traitement automatisé des données personnelles, ainsi qu'un droit de ne pas être soumis à des décisions fondées uniquement sur l'automatisation. »[267]

La manifestation parfaite de la liberté de consentir dépend donc de celle de ne pas consentir : la personne candidate reste libre de refuser qu'on lui applique une IA de recrutement.

Or dans le contexte d'un processus d'embauche, il existe une contrainte naturelle : le rapport de force au bénéfice de l'employeur. Afin de mettre en exergue cette vulnérabilité du candidat face au recruteur, nous ferons appel à la littérature portant sur les tests de dépistage de drogue dans les processus d'embauche. À la fin des années 90, les tests de dépistage de drogues en matière d'emploi ont fait couler beaucoup d'encre.

En effet, l'entreprise recruteuse ou employeuse demande à ses salariés, tout comme aux demandeurs d'emploi, de se soumettre à un test de dépistage de drogues. Dès lors, plusieurs questions juridiques émergent : d'abord celle de la liberté de consentir ou non à de telles demandes intrusives, mais également celle de la source du pouvoir de l'employeur dans ce type de demande. S'agissant de cette dernière interrogation, Andrée Lajoie distingue le pouvoir disciplinaire de l'employeur sur ses salariés, du pouvoir contractuel de l'employeur vis-à-vis des demandeurs d'emploi.

> « Dans cette situation, l'employeur exerce les pouvoirs contractuels attribués à son entreprise : ce sont donc les fondements et la portée de ce pouvoir contractuel qu'il convient d'examiner pour en déterminer l'étendue. »[268]

C'est en vertu de ce pouvoir contractuel que l'employeur serait en position d'exiger des tests de dépistage de drogues. Or, ce pouvoir contractuel en appelle un autre : un pouvoir économique.

---

[267] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 101, Proposition 3.

[268] A. LAJOIE, préc., note 155, p. 42.

Dans la relation contractuelle naissant lors d'un processus de recrutement, la nature économique de l'enjeu demeure centrale pour la personne candidate puisque son emploi conditionne son salaire, et par là même, ses moyens de subsistance. Dès lors, dans un processus d'embauche il y a d'abord un déséquilibre contractuel qui, de plus, profite à l'employeur en position de force économique. Alors comment s'assurer que cette vulnérabilité des candidats ne fait pas l'objet d'une exploitation opportuniste des employeurs ?

L'exigence de bonne foi constituerait ici une limite à l'exercice de leur pouvoir contractuel. Opérant une lecture conjointe des articles 6, 7 et 1375 du Code civil, A. Lajoie en conclut que :

> « Dans ces circonstances, les employeurs ne peuvent pas abuser du pouvoir économique dont ils disposent lors d'une offre d'emploi pour exiger de leur futur employé un consentement à des conditions non reliées à la tâche de ces employés […] : ce serait là exercer leurs droits de manières excessive et déraisonnable à l'encontre des exigences de la bonne foi, au moment de la naissance de leur obligation contractuelle. »[269]

L'exigence de bonne foi susmentionnée ferait l'objet d'une rigueur d'autant plus importante que la liberté du consentement des candidats paraît uniquement théorique.

> « la règle semble d'autant plus impérative qu'elle s'applique à un moment où la liberté contractuelle de l'employé, surtout dans une conjoncture économique caractérisée par le chômage […], n'est qu'une fiction juridique formelle : qui, aujourd'hui est vraiment libre de négocier les conditions que lui impose une offre d'emploi ? »[270]

La bonne foi contractuelle amènerait donc les recruteurs à ne pas abuser de leur position de domination économique, en imposant aux candidats d'acquiescer à la transmission de renseignements personnels. En effet, même à l'égard des salariés, une telle capacité des employeurs a déjà été considérée abusive et contraire aux chartes[271].

---

[269] *Id.*, p. 44.

[270] *Id.*, p. 45.

[271] *Syndicat des travailleurs de l'industrie de la fibre de Chambly Inc. et Bennett Fleet Inc.*, (T.A., 1990-04-30) SOQUIJ AZ-90141114, D.T.E. 90T-799, [1990] T.A. 470, en ligne : <https://soquij.qc.ca/portail/recherchejuridique/Selection/4561138> (consulté le 3 août 2021). « Quant à l'examen médical, une telle exigence va à l'encontre des droits à l'intégrité et à la liberté de la personne et au respect de la vie privée garantis par la Charte des droits et libertés de la personne. Le droit d'un employeur d'obliger un salarié à subir un examen médical doit donc être interprété restrictivement, celui-ci n'étant exigible, en l'absence de dispositions légales ou conventionnelles, que s'il y a un motif sérieux, individuel et exceptionnel »

S'agissant spécifiquement des tests de dépistage de drogues, en 2002 la Commission québécoise citait d'ailleurs son homologue fédérale :

> « À ce propos, la Commission canadienne des droits de la personne a adopté la position suivante à l'égard du dépistage de drogue : "[ni] la nature du problème de la drogue au Canada en général ni son incidence dans des professions particulières ne justifient la mise en œuvre de telles mesures. Il existe pour garantir la sécurité d'un organisme, des moyens différents et bien moins importuns d'obtenir des renseignements pertinents sur l'aptitude d'une personne à occuper un emploi." »[272]

Par analogie avec les tests de dépistage de drogues à l'embauche, on s'aperçoit d'une similarité avec l'enjeu du consentement à l'IA de recrutement. Le déséquilibre contractuel entre la personne candidate et l'employeur s'observe de la même façon : le refus de se soumettre à une IA de recrutement, risque de mettre fin au processus d'embauche.

De plus, à l'instar du test de dépistage de drogues, certaines IA de recrutement impliquent une collecte et une analyse excessives des renseignements personnels collectés. C'est le cas de l'IA de recrutement proposée par *HireVue* : la personne candidate doit transmettre des renseignements personnels qui font l'objet d'un traitement algorithmique abusif. En comparaison, dans le cas d'Amazon, il s'agissait d'un traitement biaisé qui conduit à rechercher et écarter les marqueurs du genre féminin. Toutefois, ce traitement reposait sur des informations que l'on retrouve habituellement dans des curriculums vitae. À l'inverse, dans le cas de *HireVue*, il s'agit en revanche de procéder à un traitement, certes fallacieux, des réactions émotionnelles et psychologiques des candidats par le biais de l'IA. Cela ne serait pas possible sans la collecte et l'analyse de données sensibles[273]. Le parallèle avec les tests de dépistage de drogues met ainsi en exergue les conséquences directes de l'abus du pouvoir contractuel des employeurs.

---

[272] Claire BERNARD et COMMISSION DES DROITS DE LA PERSONNE ET DES DROITS DE LA JEUNESSE, *La position de la Commission des droits de la personne et des droits de la jeunesse du Québec face aux tests de dépistage de drogue en milieu de travail*, Montréal, CDPDJ, 2002, p. 34, en ligne : <http://www4.banq.qc.ca/pgq/2006/3220898.pdf> (consulté le 31 janvier 2021).

[273] Pour rappel, l'IA de *HireVue* est en mesure de capter les mouvements oculaires, expressions faciales et intonations de la voix des personnes candidates. Au-delà du caractère intrusif d'une telle collecte, l'analyse qui en est faite est problématique aussi bien parce qu'elle est excessive, que par sa nature biaisée.

En revanche, et contrairement aux tests de dépistage, le candidat n'a pas obligatoirement l'occasion de s'opposer à l'IA de recrutement. En effet, au moment où la demande de dépistage est formulée, il se peut que le processus d'embauche soit enclenché, et qu'ils disposent d'un délai pour présenter ledit test. Cela suppose donc de la transparence sur le critère de la consommation de drogues : la personne candidate sait qu'elle doit se soumettre à un test de dépistage qui entre dans les critères de sélection des candidatures.

Or, qu'il s'agisse d'une IA de tri de candidatures comme celle d'Amazon, ou d'une IA autonome comme celle de *HireVue*, la simple mise en relation entre l'entreprise et le candidat peut s'effectuer par l'intermédiaire d'une de ces IA. Lorsque l'on dépose sa candidature sur une plateforme numérique, ou qu'on l'envoie au format numérique, rien n'indique que c'est un être humain qui les reçoit ou qui les étudie. Il se peut que la réception et le tri des candidatures passent directement par l'IA[274]. En d'autres termes le candidat soumis à une IA de recrutement, transmet obligatoirement les renseignements personnels nécessaires au tri des candidatures, que cela se fasse ou non à son insu. Dès lors, pour s'opposer à un tel procédé seul un retrait de candidature, ou l'abandon du processus de recrutement, restent à disposition.

La liberté du consentement dans le contexte d'une embauche s'apparente donc davantage à une fiction théorique plutôt qu'à un droit effectif et applicable en pratique. Cette fiction ne fonctionne donc pas.

En outre, si la littérature québécoise ne découvre pas l'antinomie entre consentement libre et recrutement, il faut remarquer que l'UE partage très clairement ce constat d'incompatibilité.

> « Au vu de la dépendance résultant de la relation employeur/employé, il est peu probable que la personne concernée soit en mesure de refuser de donner son consentement à son employeur concernant le traitement de ses données sans craindre ou encourir des conséquences négatives suite à ce refus. […] le G29 considère-t-il

[274] L'IA précitée Vera permettait par exemple de programmer des rendez-vous téléphoniques avec les personnes candidates, sans que celles-ci soient préalablement averties.

Q. PÉRINEL, préc., note 16.

> problématique que les employeurs traitent les données à caractère personnel de leurs employés actuels ou potentiels en se fondant sur leur consentement, dès lors qu'il est peu probable que celui-ci soit donné librement »[275].

Au regard de tous ces éléments, considérer que la personne candidate est de facto en incapacité d'exprimer un consentement libre, constituerait d'abord un constat réaliste. Le mécanisme du consentement perd tout son sens lorsque l'individu se situe dans un rapport de force en sa défaveur. Or, dans la mesure où ce mécanisme du consentement servait de bouclier aux autres droits de la personne, il convient de conclure que dans un processus de recrutement augmenté, le consentement n'a plus aucune utilité dans la protection des droits de la personne[276].

Par ailleurs, si d'aventure le consentement conservait un peu de sa substance dans la préservation des droits fondamentaux des postulants, les orientations législatives (fédérales comme provinciales) l'enterraient définitivement. Désormais, il existe des exceptions au consentement, notamment une au bénéfice des intérêts du secteur privé : l'intérêt commercial légitime.

c) L'intérêt légitime de l'employeur : une érosion des droits fondamentaux corrélée à celle du consentement

Dans sa publication intitulée *Un cadre réglementaire pour l'IA : recommandations pour la réforme de la LPRPDE*, le Commissariat à la vie privée choisit un paradigme contradictoire concernant le consentement.

---

[275] GROUPE DE TRAVAIL « ARTICLE 29 », préc., note 154, p. 6 à 8.

[276] Karim BENYEKHLEF, « L'IA et nos principes de justice fondamentale », *Policy Options* Ethical and Social Dimensions of AI (15 février 2018), part. Le partage des données, en ligne : <https://policyoptions.irpp.org/fr/magazines/february-2018/lia-et-nos-principes-de-justice-fondamentale/> (consulté le 18 décembre 2020). « […] les cadres législatifs actuels au Québec et au Canada sont complètement dépassés. Une réflexion majeure s'avère nécessaire. Que vaut en effet le consentement dans un écosystème (Internet, Internet des objets) qui offre des services gratuits en échange d'une collecte de données personnelles et d'un profilage algorithmique de la personne ? Que valent les principes individuels de traitement et de gestion des données personnelles dans un écosystème où les données deviennent collectives et dans lequel l'objectif est de recueillir un maximum d'information, alors que les principes des lois de protection des données s'articulent autour de la notion de minimisation (collecte des seules données nécessaires au but poursuivi), de limitation dans le temps (effacement des données une fois le but poursuivi atteint) ou, encore, de finalité ? »

Il réaffirme d'abord l'importance d'une obligation de le recueillir auprès des individus soumis à l'IA. Puis, il nuance cette réaffirmation en adhérant malgré tout à l'idée que le recueil du consentement présenterait des incompatibilités avec un encadrement de l'IA[277]. C'est en tenant compte de ces lacunes du consentement que le Commissariat recommande d'y déroger dans le cadre de trois exceptions :

- les fins de recherche et de statistiques,
- les cas où la finalité du traitement demeure compatible avec « l'objectif initial de la collecte »[278],
- et pour terminer les intérêts commerciaux légitimes.

Si les deux premières exceptions peuvent paraitre bienvenues, la troisième en revanche interpelle. En effet, dans une volonté d'encadrer l'IA afin que ses bénéfices présumés profitent à la société, on peut entendre l'intérêt d'autoriser le recours à des algorithmes pour constituer des bases de données fournies, les analyser et ainsi faire avancer l'état de la connaissance. À ce titre, la première exception paraît pertinente. De même, considérant l'imprévisibilité de l'évolution technologique des IA, il parait intéressant qu'on puisse analyser des données déjà collectées. Dans ces cas, la personne gestionnaire des données n'aurait pas besoin de redemander leur consentement aux intéressés, puisque la finalité de collecte initiale coïnciderait avec la finalité du traitement par IA. Cette deuxième exception aurait donc pour effet d'étendre le consentement des personnes à des modes de traitement qui n'étaient pas envisagés au départ.

Cette vision extensive du consentement l'affaiblit déjà, mais pourrait également constituer un compromis relatif : solliciter les personnes à chaque technologie utile au traitement de leurs

[277] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111. « En 2018, le Commissariat a mené des travaux visant à renforcer le rôle du consentement, puis a publié des Lignes directrices pour l'obtention d'un consentement valable. Ces améliorations sont importantes, mais il est capital de dire qu'en 2020, la protection des renseignements personnels ne peut reposer uniquement sur le consentement. ». « « L'IA fait ressortir les lacunes du principe du consentement, autant pour ce qui est de la protection de la vie privée des individus que de la matérialisation de ses avantages »

[278] *Id.* Le Commissariat vise notamment l'utilisation des renseignements personnels à titre de données d'apprentissage des IA.

données pourrait provoquer une lassitude[279]. Cependant, l'intérêt légitime amène des inquiétudes d'un autre ordre. Par cette dernière exception, le Commissariat recommande des limites légales au phénomène mouvant et évolutif que représente l'IA :

> « Le consentement demeurerait la règle, mais cette exception définie en termes généraux, similaire au libellé du [RGPD] offrirait la souplesse nécessaire pour autoriser l'utilisation des renseignements personnels à des fins raisonnables imprévues ».[280]

En d'autres termes par cette troisième exception d'intérêt légitime, le Commissariat permet non seulement aux entreprises de collecter des renseignements personnels des individus, mais il autorise également le traitement automatisé de ces renseignements sans consentement, aucun, des personnes concernées. Les exceptions sont justifiées de la sorte :

> « Comme il est impossible de prédire tous les usages futurs de la technologie, même à plus ou moins court terme, nous pensons que la meilleure manière de remédier à cette situation est de définir dans la loi les usages autorisés et les droits applicables de façon large, en utilisant des termes généraux qui peuvent ensuite être interprétés, en fonction du contexte au moment voulu.
>
> Compte tenu de ce qui précède, nous pensons que dans le contexte de la LPRPDE, soit une loi qui régit des activités commerciales, le meilleur moyen de circonscrire les usages autorisés est de prévoir une exception au consentement pour des intérêts commerciaux légitimes. »[281]

Le Commissariat recherche donc une actualisation de la LPRPDE qui la dote d'une résilience telle qu'elle demeurerait applicable à des usages de l'IA et des TI, aujourd'hui imprévisibles. Cet objectif bienvenu paraîtrait tout à fait audible s'il ne se concrétisait pas par des recommandations attentatoires aux droits fondamentaux. Ce qui interpelle, c'est le silence du Commissariat sur la définition d'un intérêt commercial légitime. Plus encore, la notion volontairement floue

---

[279] « L'abus de demandes de consentements et de notifications dévalorise leur signification ». Cette fatigue pourrait être comparée à celle rencontrée lors de la navigation internet : les demandes d'autorisation des cookies apparaissant sur chaque page consultée, ces traceurs sont souvent acceptés sans lecture véritable de leurs caractéristiques. Simon-Pierre DIAMOND, *Les fiducies de données - Un véhicule juridique pour rétablir l'équilibre entre la protection des données personnelles et l'essor de l'intelligence artificielle*, 28 octobre 2019, p. 5 et 6, en ligne : <https://crdm.ulaval.ca/files/20191128_Les_fiducies_de_donnees_-_Un_vehicule_juridique_pour_retablir_lequilibre_entre_la_protection_des_donnees_personnelles_et_lessor_de_lintelligence_artificielle.pdf>.

[280] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111.

[281] *Id.*, part. III. Intérêts commerciaux légitimes.

justifierait une mise en balance des droits fondamentaux par la seule revendication d'un intérêt commercial. Au surplus, il n'y a pas de critères suggérés afin de distinguer l'intérêt légitime de l'intérêt illégitime. C'est à se demander si le caractère légitime repose sur l'acceptabilité d'une atteinte : si cette dernière est estimée acceptable, considère-t-on l'intérêt commercial légitime ?

Ainsi, on pourrait se passer du consentement des personnes, pourtant exposées aux biais discriminatoires des IA de recrutement, au motif que l'obsolescence du consentement ferait de moins en moins de doute en matière d'IA. Pourtant, plutôt que de rechercher un mécanisme juridique nouveau, ou différent, le Commissariat le conserve comme règle de principe :

> « Le consentement demeurerait la règle, mais cette exception définie en termes généraux, similaire au libellé du Règlement général sur la protection des données (RGPD) et donc renforçant l'interopérabilité des lois, offrirait la souplesse nécessaire pour autoriser l'utilisation des renseignements personnels à des fins raisonnables imprévues. Il est préférable de procéder ainsi plutôt que d'étirer ou de déformer le concept du consentement implicite au point de lui faire perdre sa pertinence. »[282]

L'actualisation de la LPRPDE proposée par le Commissariat consiste donc d'abord à reconnaître que l'IA affaiblit le principal mécanisme de protection des droits fondamentaux. Puis, tirant les conséquences de ce constat, il s'agirait d'éviter « d'étirer ou de déformer le concept du consentement implicite » pour éviter de dénaturer ce mécanisme. Sans nul doute, faut-il rappeler que deux des exceptions recommandées par le Commissariat consistent justement à dénaturer le consentement. Non seulement l'extension du consentement à de nouveaux traitements mène à considérer le caractère implicite de l'accord donné, mais la seule reconnaissance d'un intérêt commercial légitime consiste à se passer de consentement, et ce, au profit d'entreprises privées. Plus encore, il semble que les contours de l'intérêt légitime soient extensibles en fonction du contexte :

> « nous pensons que la meilleure manière de remédier à cette situation est de définir dans la loi les usages autorisés et les droits applicables de façon large, en utilisant des termes généraux qui peuvent ensuite être interprétés, en fonction du contexte au moment voulu. »[283]

---

[282] *Id.*

[283] *Id.*

Cela signifierait donc que la légitimité de l'intérêt allégué serait interprétée et analysée à l'aune du contexte. Appliquée aux processus d'embauche augmentés par l'IA, l'exception de l'intérêt commercial légitime poserait donc un problème majeur : les gains de temps et d'argent qu'offre l'IA sont d'ores et déjà reconnus dans le secteur du recrutement. Alors qu'est-ce qui empêcherait la reconnaissance de la légitimité d'un processus de recrutement augmenté ? La légitimité de l'inquiétude apparaît d'autant plus que dans leurs projets de loi respectifs le gouvernement fédéral et celui du Québec prennent des directions, certes différentes, mais toutes les deux particulièrement libérales.

> « L'absence de droit à la révision d'une décision automatisée constitue selon nous une lacune importante du projet de loi C-11. D'ailleurs, il convient de mentionner que ni le projet de Loi n° 64 ni le projet de loi C-11 ne reconnaissent un droit de ne pas faire l'objet d'une décision automatisée comme le fait l'article 22 RGPD. »[284]

Plus encore, nous avons déjà évoqué l'ignorance des candidats quant à l'utilisation d'une IA de tri ; la modernisation de la LPRPDE consisterait-elle à valider cette pratique consistant à maintenir dans l'ignorance les personnes qui sont en position de vulnérabilité ? La modernité irait-elle donc dans le sens d'un recul des droits et libertés fondamentaux ?

En dépit de ces critiques, remarquons néanmoins que le Commissariat conditionne ces exceptions à la mise en place de protections additionnelles[285]. Destinées à compenser l'effacement ponctuel du consentement, ces trois « mesures protectrices » constituent l'Évaluation des facteurs relatifs à la vie privée (ci-après « l'EFVP »), la désidentification, ou encore la responsabilité démontrable[286].

---

[284] Simon DU PERRON, « Le temps des réformes : cinq comparaisons entre le projet de loi n° 64 et le projet de loi C-11 », *Projet AJC | ACT Project* (24 novembre 2020), part. Décisions automatisées, en ligne : <https://ajcact.openum.ca/2020/11/24/le-temps-des-reformes-cinq-comparaisons-entre-le-projet-de-loi-n-64-et-le-projet-de-loi-c-11/> (consulté le 27 novembre 2020).

[285] Simon DU PERRON, « Projet de loi n° 64 : échos de commission parlementaire », *Laboratoire de cyberjustice* (15 octobre 2020), en ligne : <https://www.cyberjustice.ca/2020/10/15/projet-de-loi-n-64-echos-de-commission-parlementaire/> (consulté le 17 novembre 2020); COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111.

[286] Notons ici que l'EFVP remplirait le même rôle que l'évaluation d'impact algorithmique exigée dans l'administration fédérale.

S'agissant en premier lieu de l'EFVP, la définition de la CAI[287] offre quelques éclaircissements sur ce que constitue cette première mesure de protection :

> « […] un processus permettant de déterminer si certains projets impliquant l'utilisation de renseignements personnels posent des risques en matière de protection de la vie privée. »[288]

Ce type de processus constitue donc un outil de vérification et de contrôle de la conformité légale de la gestion des renseignements personnels. En outre, il faudrait remarquer que tant la CAI que le Commissariat fédéral à vie privée recommande le recours à ce type d'évaluation. De plus, bien que ce processus s'applique initialement à la gestion des données personnelles, en matière d'IA, le Commissariat fédéral précité entend s'appuyer sur les EFVP afin de contrôler la conformité des IA aux droits de la personne. Cette approche s'inscrit dans la droite ligne de sa recommandation, régulièrement rappelée, d'ériger le droit à la vie privée au rang des droits fondamentaux de la personne. De ce fait, contrôler les IA par les EFVP permettrait de vérifier les impacts de l'algorithme sur les droits individuels, et plus largement sur la société. En outre, cette évaluation interviendrait à différentes étapes : avant la mise en fonction de l'IA en question, mais également après son déploiement dans des conditions réelles[289].

En outre, l'EFVP pourrait intégrer des tests de vérification des biais discriminatoires des IA de recrutement, avant ou après leur déploiement dans des entreprises. Il semble que la gestion de risque soit l'approche retenue par les gouvernements fédéral et québécois dans leurs projets de loi respectifs. Pour autant, il faudrait remarquer une divergence du gouvernement fédéral ne faisant pas mention de l'EFVP[290], au profit d'un pouvoir de contrôle sur les processus de gestion

---

[287] Commission d'accès à l'information du Québec, cette autorité provinciale assure la protection des données personnelles de la population québécoise.

[288] CAI, *Évaluation des facteurs relatifs à la vie privée : Savoir détecter et atténuer les risques d'atteinte aux renseignements personnels*, janvier 2018, p. 1, en ligne : <https://www.cai.gouv.qc.ca/documents/CAI_FI_efvp.pdf> (consulté le 5 juillet 2020).

[289] Rappelons ici qu'une EFVP intervenant avant le déploiement de l'IA dans ses conditions réelles, aurait surtout vocation à prévenir des risques algorithmiques que l'on pourrait identifier dès le développement et les tests effectués en laboratoire. Néanmoins, la marge d'erreurs peut varier selon que l'IA fonctionne en laboratoire (avec des jeux de données d'entrainement construits), ou en conditions réelles (avec des situations potentiellement nouvelles.

[290] S. DU PERRON, préc., note 284, part. Responsabilité des organisations.

du risque[291]. Or, recommandée par le Commissariat, en cas d'invocation de l'intérêt commercial légitime de l'employeur, l'EFVP permettrait la détection et donc la prévention des discriminations issues d'une IA de recrutement. Il semble que seul le Québec ait pris la décision de suivre la recommandation du Commissariat à la vie privée.

L'EFVP représenterait donc la première mesure de protection compensatoire face à l'exception au consentement que constitue l'intérêt commercial légitime : on évalue l'impact de la technologie *a priori*, cela pallierait donc le contournement du consentement.

En second lieu, le Commissariat entend se reposer sur une deuxième compensation au recul du consentement lorsque les exceptions suscitées[292] sont invoquées : la désidentification des renseignements personnels collectés sans consentement. Est désidentifié, ou dépersonnalisé, le « renseignement qui ne permet plus d'identifier directement la personne concernée »[293]. La désidentification constitue donc une forme d'anonymisation[294]. Cette deuxième alternative est intéressante lorsqu'il s'agit des exceptions liées à la recherche, aux statistiques, ou encore à

« L'un des éléments phares de la réforme québécoise est l'obligation pour les entreprises d'effectuer une évaluation des facteurs relatifs à la vie privée (EFVP) de tout projet de système d'information impliquant la collecte, l'utilisation, la communication, la conservation ou la destruction de renseignements personnels (art. 3.3 LPRPSP[1]). […] Or, le projet de loi C-11 ne fait aucune mention de l'EFVP ce qui s'avère pour le moins étonnant considérant que son adoption était recommandée par le Commissariat à la protection de la vie privée du Canada […] ».

[291] *Projet de loi C-11*, préc., note 20, art. 9 et 77.

« 9 (1) L'organisation met en œuvre un programme de gestion de la protection des renseignements personnels qui comprend les politiques, les pratiques et les procédures qu'elle a mises en place afin de respecter les obligations qui lui incombent sous le régime de la présente loi […] »

« 77 (1) Toute entité peut, de la manière prévue par règlement, demander au commissaire d'approuver un programme de certification comprenant les éléments suivants :

a) un code prévoyant des pratiques qui permettent de mettre en place une protection des renseignements personnels équivalente ou supérieure à tout ou partie de celle prévue sous le régime de la présente loi; b) des lignes directrices sur l'interprétation et la mise en œuvre du code de pratique […] ».

[292] On parle ici des trois exceptions du consentement précitées, soit : l'extension implicite du consentement dans la limite des fins initialement acceptées, la recherche et les fins statistiques, et surtout, l'intérêt commercial légitime.

[293] *Loi n°64*, préc., note 20, art. 19.

[294] *Id.*, art. 28. « un renseignement concernant une personne physique est anonymisé lorsqu'il ne permet plus, de façon irréversible, d'identifier directement ou indirectement cette personne ».

l'extension du consentement à des procédés de traitements inconnus au moment du recueil de l'accord des individus.

Cependant, appliquée à l'IA de recrutement, la désidentification ne présente que peu d'intérêt. Lors d'un processus d'embauche, les recruteurs ont besoin de déterminer les profils des personnes dont les candidatures sont retenues, ou écartées, afin de leur adresser une réponse. Dans un tri de candidatures, il est donc impératif d'identifier les personnes candidates. Cette mesure protectrice n'a donc pas lieu de s'appliquer aux processus d'embauche augmentés. Dans la mesure où l'intérêt légitime justifierait l'usage d'une IA de recrutement, sans recueillir le consentement préalable des candidats, la désidentification serait donc caduque. Or, cette mesure est destinée à compenser l'entrave au consentement des individus, au même titre que l'EFVP. Il nous faut donc nous intéresser à la troisième mesure protectrice compensatoire que recommande le Commissariat fédéral à la vie privée.

Enfin, dans la continuité de l'EFVP et de la désidentification, la responsabilité démontrable constitue la troisième mesure protectrice nécessaire à l'abandon légitimé[295] du consentement.

> « Par conséquent, une telle approche devrait être conjuguée à un renforcement du rôle de l'autorité de réglementation, afin de garantir que la responsabilité est démontrée et que, au bout du compte, les droits sont protégés. »[296]

La responsabilité démontrable constitue un concept et l'aboutissement d'une réflexion visant à délaisser le mécanisme du consentement en faveur d'une responsabilisation du secteur privé. Ce dernier plaide en effet en faveur d'un mécanisme de responsabilité, plutôt que celui du consentement reposant sur l'autorisation. En d'autres termes, le secteur privé souhaite transformer un système de régulation qui l'oblige à demander l'autorisation de collecter et traiter des données, en un système qui lui donne librement accès à ces opérations impliquant des données personnelles.

---

[295] Dès lors que l'intérêt commercial est réputé légitime pour faire exception au consentement, l'abandon de ce dernier ne serait légitime, qu'à la condition de le compenser par les trois mesures protectrices que nous étudions en ces lignes.

[296] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, part. Responsabilité démontrable.

Si on peut s'accorder sur un décalage manifeste entre les exigences de recueil du consentement, et les usages des technologies de l'information, la seule bonne foi du secteur privé ne suffit pas à compenser le système d'autorisation actuel. À ce propos, la méfiance du Commissariat à l'égard du secteur privé est claire :

> « […] si l'on met davantage l'accent sur la responsabilité, cela laisse une plus grande latitude ou liberté aux organisations quant à l'utilisation des renseignements personnels, parfois de manière douteuse. »[297]

Compte tenu d'abus répétés de certains GAFAM en matière d'IA, d'éthique et d'utilisation des renseignements personnels, on ne peut qu'acquiescer à cette méfiance légitime[298].

> « […] les individus ne peuvent pas compter de manière absolue sur les organisations pour traiter correctement leurs renseignements personnels, en particulier si des décisions automatisées sont en jeu […] et si l'organisation en cause ne fait pas nécessairement toujours preuve de transparence en ce qui concerne ses pratiques. Cette situation est aggravée par le profond déséquilibre de pouvoirs entre les individus et les organisations qui recourent à l'IA, caractérisé notamment par une asymétrie entre les connaissances et ressources des uns et des autres. »[299]

S'il fallait interpréter l'articulation de ce mécanisme avec l'EFVP, il faudrait peut-être considérer la responsabilité démontrable comme un contrôle de conformité des IA, complémentaire à celui

---

[297] *Id.*

[298] Rappelons ici une série d'abus et d'erreurs ayant défrayé la chronique : AGENCE FRANCE PRESSE, « Google | Une deuxième chercheuse en éthique renvoyée par le géant américain », *La Presse*, sect. Entreprises (19 février 2021), en ligne : <https://www.lapresse.ca/affaires/entreprises/2021-02-19/google/une-deuxieme-chercheuse-en-ethique-renvoyee-par-le-geant-americain.php> (consulté le 3 août 2021); Zoe SCHIFFER, « Timnit Gebru was fired from Google — then the harassers arrived », *The Verge* (5 mars 2021), en ligne : <https://www.theverge.com/22309962/timnit-gebru-google-harassment-campaign-jeff-dean> (consulté le 3 août 2021); note 17; D. HARWELL, préc., note 82; Julia ANGWIN, Ariana TOBIN et Madeleine VARNER, « Facebook (Still) Letting Housing Advertisers Exclude Users by Race », *ProPublica*, sect. Machine Bias (21 novembre 2017), en ligne : <https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin> (consulté le 7 juillet 2020); Julia ANGWIN et Terry Parris JR, « Facebook Lets Advertisers Exclude Users by Race », *ProPublica*, sect. Machine Bias (28 octobre 2016), en ligne : <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race?token=lZ_nPrh6oVJEnMzcTH1Jr59Ibe3K8XZC> (consulté le 7 juillet 2020); Thierry NOISETTE, « Facebook permet toujours des publicités filtrées par couleur de peau ou religion », *L'Obs*, sect. Tech (23 novembre 2017), en ligne : <https://www.nouvelobs.com/tech/20171123.OBS7749/facebook-permet-toujours-des-publicites-filtrees-par-couleur-de-peau-ou-religion.html> (consulté le 7 juillet 2020); Sheera FRENKEL et Matthew ROSENBERG, « Facebook Sued by District of Columbia Over Cambridge Analytica », *The New York Times*, sect. Technology (20 décembre 2018), en ligne : <https://www.nytimes.com/2018/12/19/technology/dc-sues-facebook-cambridge-analytica.html> (consulté le 20 décembre 2018).

[299] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, part. Responsabilité démontrable.

de l'EFVP. L'entreprise qui invoquerait son intérêt commercial légitime devrait donc démontrer qu'elle est conforme aux exigences légales et aux droits fondamentaux par le biais de ces deux mécanismes. De plus, à l'appui de cette interprétation, on pourrait rappeler que le Commissariat considère que l'efficacité de la responsabilité démontrable passerait par un renforcement de ses pouvoirs de contrôle et de sanction[300]. Dans de telles circonstances, l'EFVP pourrait même constituer un des éléments « démontrables » de la responsabilité attendue des entreprises. Cet élément permet d'ailleurs de rappeler que la responsabilité démontrable fait peser une obligation de documentation renforcée sur les entreprises : pour se prévaloir de l'exception au consentement, il faudra prouver la conformité à la loi et la garantie des droits fondamentaux.

> « En plus des mesures décrites ci-dessous, il faudrait l'assortir d'une exigence de tenue de registres semblable à celle que prévoit l'article 30 du RGPD. Cette exigence serait nécessaire pour renforcer la capacité du Commissariat à mener des inspections proactives en vertu de la LPRPDE et pour permettre aux individus d'exercer leurs droits en vertu de la Loi. »[301]

En matière de recrutement, la responsabilité démontrable signifie que l'employeur souhaitant recourir à une IA devra tenir un registre recensant toutes les informations, et les documents, qui prouvent que son processus de recrutement augmenté est conforme à la loi et aux droits fondamentaux. Concrètement, cela signifie que le recruteur doit démontrer le respect de son obligation de moyens renforcée de sûreté vis-à-vis des personnes candidates. Une telle obligation de sûreté sera donc considérée comme démontrable, documentée, si l'employeur peut apporter la preuve qu'à défaut du consentement des candidats, il a tout mis en œuvre pour assurer la protection de leurs droits fondamentaux au cours de son processus d'embauche augmenté par IA. On en revient à l'obligation de documentation renforcée.

Si en théorie cette mesure protectrice parait très contraignante, plusieurs facteurs conditionnent, en revanche, son efficacité.

---

[300] *Id.* « […] Les individus doivent pouvoir compter sur une autorité de réglementation qui jouit de pouvoirs d'application efficaces pour leur permettre de profiter des avantages de l'IA de façon sécuritaire. »

[301] *Id.*

Premièrement, le Commissariat à la vie privée du Canada — ainsi que son homologue québécois — ne pourra examiner ladite documentation probante qu'à la condition qu'il soit doté de ressources suffisantes à l'accomplissement de ce type de contrôles. Or, cet aspect dépend entièrement de choix politiques et des priorités budgétaires définies.

Deuxièmement, en matière d'IA de recrutement, il reste à préciser sur qui reposerait cette responsabilité démontrable : est-ce l'entreprise qui développe l'IA de recrutement qui devra se soumettre à cette obligation de documentation, ou alors celle-ci reposerait sur l'employeur ? En fonction de la réponse du législateur, il se pourrait que la responsabilité démontrable fasse l'objet d'une application hétérogène en fonction de l'accessibilité des ressources d'une entreprise. Afin d'éviter des disparités dans l'accès aux technologies, ou aux ressources nécessaires à la responsabilité démontrable, peut-être que des adaptations complexifieront l'application de cette mesure protectrice nouvelle.

Pour conclure, face aux enjeux d'opacité de l'IA et d'obsolescence des mécanismes existants (le consentement ici), on assiste à une validation des intérêts commerciaux privés sous réserve de compatibilité avec les droits fondamentaux. Cette position de compromis entraîne une sorte d'équilibre théorique précaire entre un constat d'échec du consentement, et l'émergence de mécanismes compensatoires exceptionnels. Ainsi, quand bien même on voudrait redonner au consentement son rôle protecteur, reconnaissant peut-être par là sa désuétude, le Commissariat recommande de lui asséner le coup de grâce. Il est dommage que la logique n'ait pas été menée à son terme : l'obsolescence du consentement ne signifie-t-elle pas qu'il faille changer de paradigme de régulation ? Pourquoi rattacher ces nouveaux mécanismes à un principe qu'ils contredisent intrinsèquement ? En actant l'obsolescence du consentement en IA, tout en refusant de l'abandonner, on assiste à une contradiction fondamentale : la dévaluation de l'ensemble des droits protégés par le consentement. En effet, en considérant que l'exigence de recueil du consentement demeure le principe, cela implique que le candidat serait libre de consentir. Donc théoriquement cela justifierait le contournement d'une contestation de la décision algorithmique : ce maintien du consentement de principe apporte beaucoup de confusion sur la valeur du consentement des individus. Pourquoi considérer le consentement primordial, lorsque les utilisations des IA comportant le plus de risques pour les droits

fondamentaux seront dispensées de consentement ? Il s'agit là d'une reconnaissance inachevée de l'insuffisance du consentement.

D'autre part, cette confusion s'accentue par la reconnaissance de l'intérêt commercial légitime de l'employeur. Ses intérêts privés justifieraient-ils que l'on contourne le principe ? Que signifie ce raisonnement quant à la prééminence des droits de la personne sur les intérêts privés ? Bien que les nouvelles mesures constituent des garde-fous, qu'en est-il d'une clarification de la supériorité des droits de la personne sur des intérêts mercantiles ?

Aujourd'hui, les usages de l'IA servent principalement la surveillance et la manipulation de masse des populations. On peut citer le logiciel espion Pegasus révélé au cours de l'été 2021[302] ; de même que l'affaire Cambridge Analytica[303], désormais célèbre, démontre le pouvoir de manipulation des algorithmes, au point d'influencer les résultats d'une élection présidentielle.

En matière d'emploi, la tendance se manifeste par du profilage et de l'analyse intrusifs au risque de généraliser, et d'accélérer, une machine à discriminer. La régulation de l'IA de recrutement appelle très certainement à se saisir d'une occasion de réaffirmer la prééminence des droits de la personne, piliers du caractère démocratique d'une société.

À l'aune de ces enjeux, une position plus audacieuse dans les recommandations du Commissariat à la vie privée clarifierait la place que l'on veut donner aux droits fondamentaux face à des technologies dont les usages liberticides explosent. La question est d'autant plus impérative, qu'en parallèle d'une difficulté de prévention des atteintes excessives aux droits des candidats, l'opacité de l'IA de recrutement soulève également des questionnements sur la motivation des

[302] Dana PRIEST, Craig TIMBERG et Souad MEKHENNET, « Private Israeli spyware used to hack cellphones of journalists, activists worldwide », *Washington Post* (18 juillet 2021), en ligne : <https://www.washingtonpost.com/investigations/interactive/2021/nso-spyware-pegasus-cellphones/> (consulté le 3 août 2021); « « Projet Pegasus » : révélations sur un système mondial d'espionnage de téléphones », *Le Monde.fr* (18 juillet 2021), en ligne : <https://www.lemonde.fr/projet-pegasus/article/2021/07/18/projet-pegasus-revelations-sur-un-systeme-mondial-d-espionnage-de-telephones_6088652_6088648.html> (consulté le 3 août 2021).

[303] K. GRANVILLE, préc., note 2; S. FRENKEL et M. ROSENBERG, préc., note 298.

décisions algorithmiques de rejet de candidatures. Sans justification de la décision, il est en effet difficile de motiver sa contestation.

3. L'opacité des motifs de la décision algorithmique : une entrave à la contestation des rejets discriminatoires des personnes candidates

La question de la motivation de la décision algorithmique d'écarter une candidature implique celle de la transparence de la décision. Or en matière d'IA, l'explicabilité des algorithmes constitue le principal enjeu de régulation des usages de l'IA. Si la définition de la transparence suscite des débats entre les experts de disciplines différentes, les désaccords portent également sur le réalisme de ladite transparence. En effet, d'une part les professionnels de l'IA réfutent l'hypothèse d'une explication possible des IA, lorsque juristes et professionnels de l'éthique rappellent l'importance d'y parvenir. Nous verrons donc que la motivation des décisions d'écarter des candidats est étroitement liée aux défis de l'explicabilité, mais que plusieurs travaux apportent des méthodes prometteuses d'explication ou de contrôle des biais qui sont applicables à l'IA de recrutement.

a) La motivation des décisions algorithmiques de rejet : un processus tributaire des défis de l'explicabilité (l'exemple de *HireVue*)

Le Groupe d'experts de haut niveau sur l'intelligence artificielle formé par la Commission européenne apporte une définition du principe d'explicabilité[304] qu'on pourrait rapprocher de la notion fédérale « d'explication valable » retenue par le Commissariat à la vie privée[305]. De même,

[304] « […] les capacités et la finalité des systèmes d'IA doivent être communiquées ouvertement, [et] les décisions […] doivent pouvoir être expliquées aux personnes directement et indirectement concernées. » GROUPE D'EXPERTS DE HAUT NIVEAU SUR L'INTELLIGENCE ARTIFICIELLE, *Lignes directrices en matière d'éthique pour une IA digne de confiance*, Commission Européenne, 2019, p. 16, en ligne : <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (consulté le 10 avril 2019).

[305] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, part. I. Le droit à une explication valable. « Le droit d'obtenir une explication valable s'inscrit dans le prolongement des principes qui se trouvent actuellement dans la LPRPDE, soit l'exactitude, la transparence et l'accès aux renseignements personnels. Ce droit permettrait aux individus de comprendre les décisions rendues à leur sujet et faciliterait l'exercice d'autres droits, tels que celui de faire corriger les renseignements personnels inexacts, y compris ceux fondés sur des inférences. Ce droit

si la *Déclaration de Montréal* parle de « justifiabilité »[306], ce principe se retrouve sous différentes terminologies dans la plupart des corpus éthiques d'IA[307].

En effet, la transparence permet d'effectuer un contrôle d'intégrité des décisions algorithmiques. L'intégrité de la décision d'écarter des candidatures revêt une importance essentielle quant à la confiance des individus dans la technologie, et l'amélioration des performances des IA de recrutement. En effet, ce n'est qu'à partir de données authentiques, exactes et vérifiées que l'on peut tendre vers une décision algorithmique fiable et légitime. Qu'ils soient augmentés ou non par une IA, les processus d'embauche sont bien souvent opaques quant aux critères de sélection de la meilleure candidature. Dans un processus mené uniquement par des êtres humains, l'intégrité du processus de recrutement est tributaire de la gestion des risques de biais cognitifs des recruteurs.

De même, l'intégrité des décisions algorithmiques de recrutement pose également la question de l'identification et de la correction des biais affectant l'apprentissage machine. La différence fondamentale entre un processus dit classique, et un processus augmenté c'est que la gestion du risque de biais cognitifs demeure accessible à travers plusieurs dispositifs internes[308], mais également par la contrainte légale[309]. En revanche, en matière d'IA de recrutement il n'existe pas à ce jour d'extension de tels dispositifs à l'IA. Nous avons d'ailleurs vu dans le chapitre 2 que s'il

s'apparenterait à celui énoncé à l'alinéa 15(1) h) du RGPD, qui exige que le responsable du traitement des données fournisse « des informations utiles concernant la logique sous-jacente » des décisions. »

[306] préc., note 239, part. 5 « Principe de participation démocratique ».

[307] On le retrouve notamment parmi les 23 principes d'Asilomar, dans les lignes directrices européennes pour une IA digne de confiance, ou encore dans les principes de Partnership on AI regroupant notamment les GAFAM, Alexandra BENSAMOUN et Grégoire LOISEAU, *Droit de l'intelligence artificielle*, coll. Les Intégrales, 2110-9680, 15, Issy-les-Moulineaux, LGDJ, 2019, p. 15, 16 et 25.

[308] On peut remarquer plusieurs procédés et pratiques dans le secteur du recrutement, qui ont vocation à favoriser un recul des discriminations. Citons entre autres, un engouement autour des formations sur les discriminations systémiques et les biais, ou encore des procédures visant à assurer le plus de neutralité possible. Notamment, les CV dits neutres ou dont les facteurs de discriminations sont gommés (sans la photo, le nom, le genre ou l'âge des personnes candidates); il existe aussi les grilles de critères et barèmes appliqués aux candidatures, les épreuves et tests d'évaluation des compétences etc…

[309] *Loi sur l'équité salariale*, préc., note 72, art. 4 al. 2, 31 et suiv.; *Loi sur l'équité en matière d'emploi*, préc., note 63, art. 2 et 10; *Loi Canadienne sur les droits de la personne*, préc., note 44, art. 7 et suiv.

existe des moyens de prévenir les risques de biais, leur correction quant à elle représente une tâche ardue, voire quasi impossible.

Par conséquent, si les dispositifs de prévention des biais algorithmiques s'avèrent insuffisants[310], la discrimination qui n'a pu être évitée par la prévention — donc *a priori* — appelle la sanction du droit fondamental à l'égalité *a posteriori*. De ce fait, la question de l'accès à la justice pour les candidats discriminés sous-tend celle de la capacité de notre système judiciaire à sanctionner *a posteriori*, une discrimination à l'embauche par IA.

La transparence représente donc un enjeu central quant à l'avenir des IA de recrutement. Or, les débats relatifs à la transparence de l'IA ne se limitent pas à l'acception du terme. La faisabilité d'une telle transparence algorithmique constitue également un objet de discussion.

Nonobstant la pertinence d'un tel principe, son application demeure complexe et soulève des controverses. En effet, là où d'aucuns considèrent l'explication nécessaire et incontournable, les experts de l'IA assurent que l'état de l'art ne permet pas une explication complète des résultats des algorithmes. De plus, du fait du caractère concurrentiel du marché des technologies de l'information, les entreprises qui développent les systèmes d'IA expriment une réelle frilosité vis-à-vis de ce principe.

S'agissant des limites techniques à l'explication des décisions algorithmiques, la « boîte noire »[311] constitue le principal point d'achoppement.

Défini comme un « Élément logiciel ou matériel dont le fonctionnement est connu […], mais dont la structure interne est inconnue »[312], cet angle mort implique deux observations ; premièrement,

[310] Rappelons ici que le risque 0 n'existe pas : « à l'impossible, nul n'est tenu »

[311] BENOÎT GEORGES, « Les boîtes noires du « deep learning » », *Les Echos* (27 août 2018), en ligne : <https://www.lesechos.fr/tech-medias/intelligence-artificielle/les-boites-noires-du-deep-learning-137363>.

[312] « Boîte noire », dans Grand dictionnaire terminologique, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8374149> (consulté le 18 novembre 2020).

lorsqu'on est confronté à un système d'IA, il est possible de connaître les données d'entrée[313] et les données de sortie[314]. Deuxièmement, ce qui se passe entre les données d'entrée et de sortie demeure un mystère[315], *a fortiori* en présence de réseaux de neurones artificiels[316]. L'Autorité canadienne de protection de la vie privée estime d'ailleurs que cette grande inconnue des systèmes d'IA pose un problème lorsqu'il s'agit de vérifier l'existence de biais discriminatoires appris par l'algorithme.

> « Les algorithmes utilisés pour prendre une décision concernant une personne peuvent être comme une boîte noire, laissant la personne dans l'ignorance quant à la manière dont la décision a été prise. Il est également admis que les données ne sont pas intrinsèquement objectives […]. Les décisions automatisées risquent d'être inéquitables, biaisées et discriminatoires. »[317]

Dans le cas d'IA de recrutement, la boîte noire représente donc un premier obstacle à la motivation d'une décision automatisée discriminatoire. En effet, lorsque la boîte noire empêche de saisir la logique de la décision de tri X, comment transmettre l'explication relative à la candidature X ? De même, comment y identifier un biais discriminatoire ? L'opacité de l'IA de

[313] Soit l'ordre ou la question adressés à l'IA.

[314] C'est-à-dire le résultat de l'IA ou la décision algorithmique.

[315] Ce fut notamment le cas avec l'IA Alphago qui a vaincu le champion humain au jeu de Go : les concepteurs ne savent pas par quel procédé l'IA a pu conclure à une combinaison gagnante : on pouvait lire la réaction du Pr. Stuart Russel : « malheureusement, nous ne savons pas exactement ce qu'il fait, ni d'ailleurs ses créateurs» in AGENCE FRANCE PRESSE, « La victoire d'Alphago contre le champion du jeu de go restera dans l'Histoire », *Les Affaires*, sect. Technologies de l'information (14 mars 2016), en ligne : <https://www.lesaffaires.com/techno/technologie-de-l-information/la-victoire-d-alphago-contre-le-champion-du-jeu-de-go-restera-dans-l-histoire/585999> (consulté le 23 novembre 2020); William AUDUREAU et Florian REYNAUD, « Jeu de go : victoire décisive de l'intelligence artificielle contre Lee Sedol », *Le Monde*, sect. Pixels (12 mars 2016), en ligne : <https://www.lemonde.fr/pixels/article/2016/03/12/jeu-de-go-victoire-decisive-de-l-intelligence-artificielle-contre-lee-sedol_4881624_4408996.html>.

[316] « Réseau de neurones artificiels », dans Grand dictionnaire terminologique, coll. Intelligence Artificielle, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8386038> (consulté le 26 novembre 2020); Yoshua BENGIO, « La révolution de l'apprentissage profond - Interstices », 2019, en ligne : <https://interstices.info/la-revolution-de-lapprentissage-profond/> (consulté le 19 juillet 2019). « Ensemble organisé de neurones artificiels interconnectés, créé dans le but de pouvoir effectuer des opérations complexes ou de résoudre des problèmes difficiles grâce à un mécanisme d'apprentissage lui permettant d'acquérir une forme d'intelligence. […] À l'origine, les créateurs de réseaux de neurones artificiels se sont inspirés du fonctionnement du système nerveux, lequel est organisé en fonction des liaisons qui s'établissent entre des neurones biologiques. »

[317] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111.

recrutement représente d'autant plus de risques pour les candidats que le secteur du recrutement présente déjà des imperfections particulièrement problématiques.

Mesuré comme étant l'un des plus discriminatoires (raciste et sexiste notamment)[318], le secteur du recrutement ne peut pas compter uniquement sur la nature humaine des recruteurs pour corriger les biais des IA. Il faudrait donc des méthodes de mitigation du risque de discrimination à l'embauche, qui sont complémentaires et qui contribuent à la réduction des risques, tant au niveau des recruteurs humains que des IA de recrutement. En effet, si par le biais de travaux en équité diversité et inclusion, les êtres humains gèrent les risques issus de leurs biais cognitifs, les biais des IA de recrutement, eux, risquent de persister.

Aussi, l'explication de la décision algorithme occupe une place centrale dans la recherche et la neutralisation éventuelle des biais discriminatoires d'une IA de recrutement.

Une définition de l'explication s'impose donc afin de lever toute ambiguïté. Le collège de spécialistes du Centre Berkman de l'Université de Harvard définit l'explication comme suit : « explanation, as required under the law, […] is about answering how certain factors were used to come to the outcome in a specific situation ».

De cette définition est tirée l'idée d'une distinction entre le système d'explication, et l'explication intrinsèque de l'IA. Selon cette logique, le système d'explication est conçu dans le but d'identifier les facteurs déterminant une décision donnée, alors que le système d'IA est conçu pour effectuer des prédictions. Aussi, pour rendre une décision algorithmique humainement compréhensible, il ne serait pas nécessaire de savoir par quels procédés les algorithmes d'IA procèdent aux prédictions. Appliquée à l'IA de recrutement, cette idée signifierait qu'il n'est pas nécessaire de décortiquer les modèles prédictifs de l'IA ; ce qui importe c'est de connaître d'une part les données d'entrée, et d'autre part, de vérifier l'existence de biais.

[318] Myrlande PIERRE, Manon POIRIER, Jean-Philippe BEAUREGARD et Paul EID, *Conférence - Discrimination à l'embauche : que nous disent les recherches sur le testing de CV ?*, 11 février 2021.

Deux méthodes permettraient ainsi de parvenir à l'explication d'une décision algorithmique : l'explication locale, c'est-à-dire expliquer une décision spécifique, et l'équité contrefactuelle qui consiste à tester les résultats de l'IA.

Ainsi, l'explication locale consisterait à identifier les intrants qui influencent le plus la décision. L'explication locale en IA peut se définir comme suit :

> « In the AI world, explanation for a specific decision, rather than an explanation of the system's behavior overall, is known as local explanation […]. This explanation is local in the sense that the important factors may be different for different instances. For example, for one person, payment history may be the reason behind their loan denial, for another, insufficient income. »[319]

Dans un cas d'IA de recrutement, cette méthode d'identification des biais permettrait donc de savoir ce qui influence le plus le classement des personnes postulantes : la formation, l'expérience professionnelle, la réussite de cas pratiques, les qualités relationnelles, etc.

S'agissant en outre de l'équité contrefactuelle, celle-ci permet de tester les biais à travers les résultats de l'IA. En modifiant les intrants (ou donnée d'entrée), on est en mesure d'observer les variations de résultats de l'IA. Cette méthode d'explication est également défendue par des scientifiques de l'Institut Alan Turing.

> « Our definition of counterfactual fairness captures the intuition that a decision is fair towards an individual if it is the same in (a) the actual world and (b) a counterfactual world where the individual belonged to a different demographic group. »[320]

En effet, le concepteur d'une IA souhaite produire des prédictions qui se rapprochent le plus de la réalité. L'entraînement de l'IA permet d'ailleurs de développer la capacité de généralisation de l'IA pour qu'elle produise de bonnes prédictions[321]. L'équité contrefactuelle consiste donc à vérifier la variation des prédictions lorsque l'on change une variable x parmi les intrants.

---

[319] Finale DOSHI-VELEZ, Mason KORTZ, Ryan BUDISH, Christopher BAVITZ, Samuel J. GERSHMAN, David O'BRIEN, Stuart SHIEBER, Jim WALDO, David WEINBERGER et Alexandra WOOD, « Accountability of AI Under the Law: The Role of Explanation », *SSRN Electronic Journal* 2017, 7, DOI : 10.2139/ssrn.3064761.

[320] « Counterfactual Fairness », *Microsoft Research*, en ligne : <https://www.microsoft.com/en-us/research/video/counterfactual-fairness/> (consulté le 3 novembre 2020); Matt KUSNER, Joshua LOFTUS, Chris RUSSELL et Ricardo SILVA, « Counterfactual Fairness », 2018, en ligne : <https://arxiv.org/abs/1703.06856>.

[321] P. BERTAIL, D. BOUNIE, P. WAELBROECK et S. CLÉMENÇON, préc., note 165, sect. « 3.2 Les biais algortihmiques ». « la règle de décision déterminée par un algorithme d'apprentissage […] ne doit pas seulement pouvoir prédire le passé,

Par exemple, dans un processus de recrutement augmenté, que se passe-t-il lorsque le genre des postulantes est supprimé des candidatures, le classement des profils est-il différent ? De même, le classement varie-t-il en fonction de certains marqueurs ethnoculturels ? Ces questions mènent aux tests d'équité contrefactuelle. En changeant certains intrants, comme le groupe ethnoculturel, l'âge, le genre ou la race, on peut observer si l'algorithme de recrutement produit un classement différent. L'idée est de vérifier les variations des données de sortie en fonction des intrants susceptibles de fonder une discrimination à l'embauche.

En conclusion, par ces deux méthodes, la transparence des décisions algorithmiques de rejet de candidature serait donc possible sans ouvrir la boîte noire. On ne saura toujours pas par quel « raisonnement »[322] l'IA de recrutement classe ou écarte des candidatures, mais on pourra observer son comportement face à des intrants portant sur des motifs de discriminations.

L'équité contrefactuelle et l'explication locale ayant vocation à tester les biais discriminatoires transmis à la machine, il serait ainsi possible d'obtenir la preuve que le processus de recrutement serait discriminatoire envers certains postulants soumis à l'IA utilisée. De ce fait, lorsque des biais raciaux, de genre ou ethnoculturels seront constatés, cela signifie qu'on pourra identifier un groupe de personnes présentant les mêmes caractéristiques dévaluées par l'IA. Dès lors, l'ensemble des postulants appartenant à ce groupe sont susceptibles d'avoir été écartés sur des motifs discriminatoires. Ces deux solutions représentent donc un mode de contrôle *a posteriori* des biais d'une IA de recrutement.

Toutefois, si ces deux méthodes ouvrent la voie à la constitution de preuves de biais discriminatoires des IA de recrutement, elles ne répondent pas à la question très concrète de l'accès à ces preuves par les personnes candidates. Par conséquent, ces solutions à l'opacité des

mais permettre de prédire efficacement, lorsqu'elle sera déployée, le label Y associé à de nouvelles données d'entrée X, non encore observées. On dira le cas échéant que la règle prédictive a alors de « bonnes capacités de généralisation ».

[322] À défaut d'un vocabulaire dédié aux machines, on parle ici de raisonnement par analogie au cheminement intellectuel déductif ou inductif d'un être humain.

IA de recrutement présentent non seulement un aspect contraignant, mais appellent également des alternatives.

b) Le dépassement de l'explicabilité de l'IA de recrutement : la perspective prometteuse des preuves de biais algorithmiques

Nous avons vu précédemment que la preuve d'une discrimination à l'embauche par IA est *a priori* possible avec deux méthodes : l'explication locale et l'équité contrefactuelle. Toutefois, plusieurs limites contraignantes rendent ces méthodes difficiles à appliquer pour des personnes lésées au cours du processus d'embauche augmenté.

Tout d'abord, les mêmes universitaires du Centre Klein Berkman précité, soulèvent des limites techniques et éthiques aux deux voies exposées. Aussi bien l'explication locale que l'équité contrefactuelle reposent sur l'identification d'intrants déterminants de la donnée de sortie. Appliquées à l'IA de recrutement, cela implique, dans ces deux méthodes, une identification des intrants qui influencent la décision d'écarter ou de retenir une candidature.

Or l'identification des intrants repose sur deux prérequis : une cartographie des intrants, ainsi qu'une représentation mathématique de concepts complexes intrinsèquement humains. Autrement dit, il faudrait d'abord répertorier l'ensemble des données d'entrée, mais au surplus, il faudrait que ces données d'entrée contiennent des concepts humains complexes, traduits en équation mathématique. En effet, ce n'est que par le langage mathématique que l'on pourrait enseigner à l'algorithme des concepts tels que l'équité, le genre, l'égalité ou même la neutralité. Le défi, qui n'a pas encore été relevé de façon satisfaisante, est donc de taille.

Ensuite, dans l'hypothèse où l'on parviendrait à une traduction mathématique fonctionnelle de ces concepts, on serait confronté à un nouveau défi, en particulier avec l'équité contrefactuelle. En effet, si la cartographie des intrants ne semble pas représenter de défi particulier[323], l'équité contrefactuelle, en revanche, implique un travail plus délicat : cette

[323] F. DOSHI-VELEZ et al., préc., note 319, 8. « For example, self-driving cars may have multitudes of sensors, each with highdimensional range and vision inputs; the human brain already converts its visual inputs into higher-level concepts such as trees or street signs ». Globalement, la cartographie reposera donc sur l'étiquetage de données numériques

dernière méthode suppose la présence de filtres appris à l'IA tels que la race, le genre, l'orientation sexuelle ou le handicap.

> « In summary, to build AI systems that can provide explanation […], we must both list those terms and allow the AI system access to examples to learn them. System designers should design systems to learn these [filters], and also store data from each decision so that is possible to reconstruct and probe a decision post-hoc if needed. »[324]

Or, même si les personnes conceptrices d'IA parviennent à implémenter des filtres aussi sensibles et complexes, le test d'une IA implique l'accès aux données toutes aussi sensibles des personnes soumises à l'IA. Vérifier l'équité contrefactuelle d'une IA de recrutement nécessite donc la collecte de données sociodémographiques[325] des personnes candidates, puis leur traitement par l'IA, et pour finir un contrôle des biais liés à ces filtres par la méthode de l'équité contrefactuelle. Pour révéler des biais discriminatoires de l'IA de recrutement par cette méthode, il faudrait effectivement comparer les résultats de classement lorsqu'on applique certains filtres sociodémographiques.

Pourtant, de tels filtres comportent leur part de risque dans la mesure où même sans intention de les enseigner négativement à la machine, ils peuvent produire des effets invisibles même aux yeux de ses programmeurs : les systèmes d'IA ont souvent la capacité de les créer de manière autonome au cours de leur apprentissage[326]. Éthiquement, il y a un enjeu à arbitrer : faut-il créer de tels filtres pour identifier des discriminations au risque de les faire exploser ? Pour le moment, et en dépit d'un sens éthique douteux, l'expérience de Facebook conduirait à beaucoup de prudence[327].

---

de sorte que ces données numériques soient présentées en des termes de langage humain : les pixels X seraient traduits comme représentant un chat, un ours, ou un ballon.

[324] *Id.*, 9.

[325] Le groupe ethnoculturel, la race, le genre, le handicap, le lieu de vie (quartiers à fort taux de criminalité par exemple).

[326] Voir le cas des biais sexistes de l'IA de recrutement de Amazon, supra *Chapitre 2, A. LA base de données de l'IA de recrutement, un vecteur de transmission des biais humains à la machine : le cas Amazon*.

[327] T. NOISETTE, préc., note 74.

Par ailleurs, au-delà des imperfections techniques de ces méthodes de constitution des preuves, les personnes candidates n'ont aucun moyen d'accéder à de tels modes de preuves. En effet, ces méthodes sont développées par des laboratoires de recherche en IA et disposent très probablement des ressources nécessaires à l'évaluation d'IA de recrutement. Bien que l'on puisse se réjouir de cet état de fait, les candidats n'auront pas les moyens de prouver par ces méthodes que le processus de recrutement auquel ils ont participé leur a été préjudiciable. On se heurte donc à un obstacle concret de taille : toutes les méthodes du monde ne facilitent pas l'accès à la justice des candidats discriminés si ces derniers n'y ont pas accès.

En attendant que l'on tranche les questions et débats éthiques, il faut des réponses à fournir aux demandeurs d'emploi d'aujourd'hui, d'ores et déjà soumis à des IA de recrutement. De même, les employeurs de bonne foi, ou attachés à leur image de marque, auront eux aussi besoin d'accéder aux informations leur permettant de gérer les risques discriminatoires de leurs IA de recrutement.

Par conséquent, si l'explicabilité demeure le procédé le plus logique pour prouver la discrimination algorithmique à l'embauche, en pratique elle présente un certain nombre de contraintes. Or, de nombreuses IA de recrutement sont d'ores et déjà en fonction, et *a priori* sans configuration préalable compatible avec un système d'explication. Il faut donc des alternatives à l'explication dans l'accès à la preuve.

c) Les limites des preuves de biais discriminatoires : une difficulté d'accès à la preuve algorithmique pour la personne candidate lésée

Les alternatives développées par le collège d'experts susmentionné s'articulent autour de l'évaluation globale de la fiabilité de l'IA. Ces alternatives permettraient ainsi un accès à la preuve pour des IA de recrutement auxquelles on ne pourrait pas appliquer les méthodes d'explication susvisées.

En premier lieu, les garanties théoriques constituent une première voie d'accès à la preuve de discrimination.

> « Theoretical guarantees are a form of perfect accountability that only AI systems can provide, and ideally will provide more and more often in the long term […] »[328]

Il s'agit de processus dont la fiabilité est prouvée en laboratoire : par exemple, les systèmes de cryptage, de votes électroniques, ou encore de loteries. Ces procédés démontreraient un fonctionnement dont la fiabilité n'est pas à démontrer pour chaque résultat fourni par les IA.

> « For example, we trust our encryption systems because they are backed by proofs; neither explanation or evidence are required. Similarly, if there are certain agreed-upon schemes for voting and vote counting, then it may be possible to design a system that provably follows those processes. Likewise, a lottery is shown to be fair because it abides by some process, even though there is no possibility of fully explaining the generation of the pseudo-random numbers involved. »[329]

Dans le cas du recrutement augmenté, on pourrait imaginer des systèmes d'IA dont les processus permettraient de garantir la fiabilité des résultats. L'exemple de vérification telle que celle opérée chez L'Oréal pourrait ainsi se développer pour parvenir à une gestion du risque discriminatoire satisfaisante[330].

Au Canada, l'obligation de responsabilité démontrable et l'EFVP (évaluation des facteurs relatifs à la vie privée) pourraient également compter parmi les supports de garanties théoriques : l'EFVP documenterait ces procédés dignes de confiance que les commissaires à la vie privée pourraient contrôler, voire certifier. Alors que ces garanties théoriques offrent une piste de solution intéressante, notons cependant que l'équipe précitée de Harvard doute de l'application de telles garanties dans le monde réel et complexe : « […] however, these guarantees require very cleanly specified contexts that often do not hold in real-world settings »[331]. Nous rejoignons ces doutes dans le cas de l'IA de recrutement. Il faudrait donc se tourner vers la seconde alternative proposée par les chercheurs : la bonne vieille preuve statistique.

---

[328] F. DOSHI-VELEZ et al., préc., note 319, 11.

[329] *Id.*

[330] Une responsable de l'Oréal témoignait à cet effet « On met en place des garde-fous ; au moment du lancement de l'outil pour chaque pays, les recruteurs vont systématiquement évaluer les candidatures ayant les scores les plus élevés et les moins élevés pour s'assurer de la pertinence de la prédiction. » R. DEMICHELIS, préc., note 16.

[331] F. DOSHI-VELEZ et al., préc., note 319, 11.

La preuve statistique consisterait à mesurer le taux d'erreurs en fonction des caractéristiques personnelles des individus. À l'instar des audits menés par le NIST sur la reconnaissance faciale, des audits pourraient mettre en exergue les groupes sociodémographiques surreprésentés dans les erreurs des IA de recrutement. De telles observations provoquent l'inversion du fardeau de la preuve en faveur des personnes candidates, mais ne permettent pas de prouver que la décision concernant la candidature de la personne X est affectée du biais statistiquement observable. En dépit de cette limite, cette preuve peut s'obtenir sur tout système d'IA de recrutement déployé et a déjà été jugée recevable dans le contentieux en matière d'emploi[332], voire particulièrement éclairante[333]. Donc, y recourir permettrait *a priori* d'éviter la question de la recevabilité de la preuve ; de même, par la preuve statistique, on s'épargne peut-être les conséquences d'une littératie numérique embryonnaire du monde judiciaire.

Ainsi, ces alternatives aux méthodes d'explication entraînent une appréciation de la confiance qu'on peut avoir dans une IA de recrutement, et ce, aussi bien à travers les garanties théoriques que les preuves statistiques.

Malgré les promesses de l'ensemble de ces voies d'accès à la preuve, celles-ci ne répondent pas à un impératif d'accessibilité pour les victimes de discriminations.

En effet, même une personne extrêmement motivée, ne pourrait se procurer de telles preuves par sa seule détermination. Le *testing* de l'IA de recrutement implique soit l'application d'une méthode d'explication précitée[334], soit de recourir à une alternative[335]. Pourtant la preuve

---

[332] Aidan R. VINING, David C. MCPHILLIPS et Anthony E. BOARDMAN, « Use of Statistical Evidence in Employment Discrimination Litigation », (1986) 64-4 *Can. B. Rev.* 660-702, en ligne : <https://heinonline.org/HOL/P?h=hein.journals/canbarev64&i=666> (consulté le 21 mai 2020).

[333] *Canada (Procureur Général) c. Walden*, 2010 Cour fédérale, par. 114 à 116, en ligne : <http://canlii.ca/t/2c43v> (consulté le 2 décembre 2020); *Canada (Human Rights Commission) c. Canada (Department of National Health and Welfare)*, 1998 Cour fédérale, par. 20 et 21 [*Chopra*], en ligne : <http://canlii.ca/t/4c05> (consulté le 2 décembre 2020); *Khiamal c. Canada (Commission des droits de la personne)*, 2009 Cour fédérale, par. 98 à 102, en ligne : <http://canlii.ca/t/26vcn> (consulté le 2 décembre 2020).

[334] L'explication locale ou l'équité contrefactuelle.

[335] Ici, les garanties théoriques ou encore le *testing*.

statistique, tout comme les garanties théoriques, implique des contrôles et vérifications tenant de l'audit interne ou externe. Or, même si une entreprise effectuait de tels contrôles, celle-ci n'a aujourd'hui aucune obligation de divulguer ses processus internes. De même, comment les candidats auraient-ils connaissance de ces contrôles ?

À l'aune de ces limites, ces voies de preuves demeurent toutes inaccessibles à une personne injustement écartée du processus d'embauche. Cette première difficulté ne permet donc pas de contourner la vulnérabilité intrinsèque dans laquelle les candidats se trouvent. Condition favorable à la normalisation des contrats d'adhésion[336], tout comme au recul de la valeur des droits fondamentaux, cette vulnérabilité rend le consentement inefficace[337].

En outre, de cette difficulté d'accès à la preuve en émerge une autre : très prosaïquement, les chances qu'une personne en recherche d'emploi engage une procédure contre un recruteur sont réduites.

La personne postulante qui conteste judiciairement le processus d'embauche fera face à plusieurs défis : la lourdeur de la procédure, le coût et les délais de cette procédure, mais également le caractère énergivore et parfois éprouvant psychologiquement d'une procédure judiciaire. Or, même en obtenant gain de cause, cela pourrait empiéter sur sa recherche d'emploi. Cela pourrait également l'exposer à une publicité non désirée auprès d'autres recruteurs. Pour le formuler trivialement, il y aurait fort à parier que beaucoup de candidats ayant connaissance de la discrimination, ou des soupçons de discriminations, considèrent que la contestation n'en vaut pas la peine.

---

[336] *Code civil du Québec*, préc., note 265, art. 1379. « le contrat est d'adhésion lorsque les stipulations essentielles qu'il comporte ont été imposées par l'une des parties ou rédigées par elle, pour son compte ou suivant ses instructions et qu'elles ne pouvaient être librement discutées. »

[337] S.-P. DIAMOND, préc., note 279, p. 6. La désuétude du régime fondé sur le consentement entraînerait plusieurs conséquences : « Favorise le contrat d'adhésion ; L'abus de demandes de consentements et de notifications dévalorise leur signification ; Crée une industrie de collecte de données personnelles et de son commerce ; Activités commerciales qui favorisent l'oligopole. ». De même sur le plan juridique, les effets constatés sont que cela : « Favorise le consentement non éclairé ; Creuse la disproportion des pouvoirs entre le citoyen et les grands détenteurs des données personnelles ; Accentue la précarité de la protection à la vie privée et aux droits de la personne. »

Bien que ces voies d'accès à la justice des personnes candidates demeurent des pistes intéressantes, ces méthodes d'explication et leurs alternatives échappent certes toutes à la perfection. Mais il convient de remarquer qu'en l'absence d'un cadre juridique clair et adapté, ces différentes solutions ne pourront pas compenser les principaux obstacles au droit de contestation des personnes candidates. En effet, pour conclure nos développements autour de l'opacité de l'IA de recrutement et l'accès à la preuve de discrimination, plusieurs constats s'imposent. Qu'il s'agisse d'autoriser le profilage et le tri des personnes candidates, ou bien d'obtenir la preuve d'une discrimination par IA, l'opacité des IA de recrutement constitue une entrave mettant en échec les mécanismes juridiques sur lesquels le législateur entend actuellement se reposer.

Or, nous l'avons vu, cette opacité de l'IA de recrutement se conjugue à une faiblesse de la personne candidate ; elle ne dispose pas véritablement de moyens de contester une discrimination par IA de recrutement. Au surplus de l'opacité de la technologie, la vulnérabilité des personnes postulantes constitue donc un frein dont découlerait l'impossibilité de consentir, et d'accéder aux informations nécessaires à l'exercice d'un droit d'ester contre une décision algorithmique de tri de candidature. La réponse à la problématique du droit de contestation se situe donc ailleurs. Pour que les candidats lésés puissent exercer leur droit de contester une discrimination à l'embauche par IA, ils doivent être en mesure de se saisir des 4 voies de preuves[338] de biais discriminatoires étudiées. Pour aboutir à cela, des changements de paradigmes s'imposent dans la régulation de l'IA de recrutement.

[338] Rappelons que ces quatre voies d'accès à la preuve ont été étudiées en distinguant, d'une part, celles qui reposent sur le principe de l'explication de la décision algorithmique, — soit l'explication locale et équité contrefactuelle (c.f supra *b) Le dépassement de l'explicabilité de l'IA de recrutement : la perspective prometteuse des preuves de biais algorithmiques* pp. 141 à 143) — de celles qui, d'autre part, ont vocation à contourner l'explication pour se concentrer sur la fiabilité de l'IA — soit les garanties théoriques et la preuve statistique (c.f supra, *c) Les limites des preuves de biais discriminatoires : une difficulté d'accès à la preuve algorithmique pour la personne candidate lésée* pp. 143 à 146).

B. La correction des entraves à la contestation des décisions de rejet de l'IA de recrutement : une refonte nécessaire du cadre légal émergent

Dans cette seconde partie, il s'agira de proposer les correctifs susceptibles de neutraliser les effets de l'opacité de l'IA de recrutement, et par voie de conséquence, les entraves à l'exercice du droit d'ester contre la décision d'écarter un profil. Trois approches complémentaires seront proposées. L'une visant à intervenir sur la vulnérabilité de la personne en recherche d'emploi (1).

Ensuite, une deuxième approche consisterait à définir un cadre juridique qui responsabilise les acteurs de l'IA de recrutement : les personnes conceptrices, utilisatrices, ou vendeuses de l'IA de recrutement, devraient être soumises à des exigences claires, suivant un principe de gradation de la contrainte en fonction du droit fondamental exposé (2).

Enfin, une telle définition du cadre juridique de l'IA de recrutement devrait s'accompagner d'une clarification des réponses juridictionnelle et institutionnelle au contentieux de l'IA de recrutement (3).

1. La nécessité d'un changement de paradigme face à l'IA de recrutement : l'exigence de mécanismes de rééquilibrage du rapport de force entre employeurs et demandeurs d'emploi

Nous avons établi la vulnérabilité des personnes candidates, de leurs difficultés à consentir, comme à saisir une juridiction pour défendre leurs droits en cas de discriminations.

Faire l'économie de mécanismes de rééquilibrage de cette relation contractuelle entre candidats et employeurs, aurait donc pour conséquence de subordonner le respect des droits fondamentaux aux variations de la conjoncture économique. Il est donc essentiel de penser des mécanismes juridiques susceptibles de faciliter l'accès à la justice des candidats discriminés. Cela passe par des leviers palliant cette vulnérabilité économique des personnes demandeuses d'emploi.

a) Du consentement des candidats vulnérables à des mécanismes d'autorisation par un tiers expert

Nous l'avons vu précédemment, le consentement ne représente plus un mécanisme de protection des droits fondamentaux des individus soumis à l'IA de recrutement. En effet, la situation de vulnérabilité du demandeur d'emploi concourt à son impossible liberté de consentir à l'IA de recrutement. Cette vulnérabilité, nous l'avons vu, constitue un terreau fertile aux contrats d'adhésion, à la dévalorisation des droits fondamentaux[339], et donc à celle du droit à l'égalité.

Cependant, que se passerait-il si un tiers compétent protégeait le droit fondamental à la vie privée de l'individu soumis à des IA, y compris dans le contexte de recrutement ? Dès lors qu'il s'agit de quelqu'un d'extérieur au rapport de force entre employeur et candidat, ce tiers serait en mesure d'évaluer le risque de discrimination posé par une IA. Ce tiers expert défendrait les droits fondamentaux de l'individu face à l'IA de recrutement de l'employeur. L'intervention d'un tiers expert contrebalancerait ainsi le rapport de force dans l'accès à l'information par exemple. Le paradigme ici proposé s'appuie notamment sur deux pistes de réflexion : la fiducie des données, et un système d'autorisation des autorités de protection de la vie privée. Ces deux pistes de réflexion consistent à transférer la charge du recours pesant sur les candidats vulnérables, vers un ou plusieurs tiers experts, qu'ils soient professionnels ou institutionnels.

S'agissant de la fiducie de données, il s'agit d'un transfert de responsabilité à un tiers professionnel, chargé de protéger les données et les droits du candidat vulnérable. En effet, la fiducie de données consiste à appliquer le système de la fiducie aux données personnelles.

> « La fiducie résulte d'un acte par lequel une personne, le constituant, transfère de son patrimoine à un autre patrimoine qu'il constitue, des biens qu'il affecte à une fin

[339] C.f supra, note 337, concernant l'effet de la désuétude du consentement sur la protection des droits fondamentaux p.146

> particulière et qu'un fiduciaire s'oblige, par le fait de son acceptation, à détenir et à administrer. »[340]

Malgré les difficultés théoriques qu'entraîne la notion de fiducie[341], le concept de fiducie de données consisterait à créer une fiction juridique : une universalité fictive constituée des droits, renseignements et données personnelles d'un individu A ; appelons-la l'universalité numérique. Cette universalité numérique de la personne A serait administrée par un tiers expert B, fiduciaire de ses données. L'intérêt du concept de fiducie dans le cadre de l'IA, c'est le statut de fiduciaire. En effet, en droit civil, le fiduciaire a une qualité juridique impliquant une série d'obligations.

> « […] le fiduciaire est soumis à une série d'obligations (prudence, diligence, honnêteté, loyauté) qui sont la contrepartie nécessaire des pouvoirs juridiques qu'il détient sur les biens mis en fiducie. La nature de sa mission justifie également que le fiduciaire doit rendre compte de son administration aux bénéficiaires de la fiducie […]. En cas de manquement à ses obligations, le fiduciaire engage sa responsabilité personnelle (art. 1308 C.c.Q.). »[342]

Ainsi, par analogie, le fiduciaire de données B se verrait soumis aux mêmes obligations qu'un fiduciaire classique. La seule différence notable résiderait dans la compétence du fiduciaire : en droit civil, toute personne peut être nommée fiduciaire, mais en matière d'universalité numérique, cette qualité ne serait accessible qu'aux personnes compétentes. Cette condition conduirait ainsi à une responsabilité professionnelle du tiers expert, et non à une responsabilité personnelle. La fiducie des données répondrait donc à deux objectifs : non seulement la vulnérabilité des candidats ne constituerait plus une situation source d'abus dans les processus de recrutement, mais en plus, par l'existence du fiduciaire des données, le candidat n'aurait plus

[340] *C.c.Q*, préc., note 257, art. 1260.

[341] La fiducie s'appliquant à un patrimoine autonome, il s'agit d'un instrument de gestion des biens. Or, une telle perspective impliquerait de considérer les données personnelles comme des biens également. La question serait de savoir si une donnée constitue un bien, et par voie de conséquence, si on peut constituer un patrimoine autonome de données. Cette question de la nature des données n'étant pas tranchée, nous nous attarderons sur l'idée d'une sorte de patrimoine constitué exclusivement des données personnelles, droits et libertés individuelles applicables aux activités numériques, ainsi qu'aux technologies de l'information.

[342] Anne-Sophie HULIN, « Introduction à la fiducie québécoise de données », *Laboratoire de cyberjustice* (26 novembre 2020), part. 1. Qu'est-ce qu'une fiducie de données ?, en ligne : <https://www.cyberjustice.ca/2020/11/26/introduction-a-la-fiducie-quebecoise-de-donnees/> (consulté le 27 novembre 2020).

à assumer le poids administratif, économique[343] ou judiciaire d'une contestation de la décision algorithmique discriminatoire. En effet, la protection de ses droits étant déléguée à un tiers expert, la personne candidate n'aurait plus qu'à se reposer sur l'expertise du tiers, et ainsi se concentrer sur sa recherche d'emploi. De même, en cas de détection d'une discrimination, la personne candidate délègue encore l'exercice et le suivi du recours à ce même tiers expert. Là encore, cela lui permet de dédier son énergie à sa recherche d'emploi, et ce, même si un ou plusieurs recours sont intentés en son nom. L'accès à la justice du candidat soumis à l'IA de recrutement serait ainsi facilité, en dépit de sa vulnérabilité.

Dans la recherche d'un meilleur équilibre contractuel entre employeurs et candidats, la fiducie de données pourrait donc constituer une piste intéressante dans la lutte contre les discriminations par IA de recrutement. Toutefois, l'une des réserves majeures qu'il nous faut relever, repose sur la nature intrinsèque d'une fiducie : elle porte sur des biens. De ce fait, le concept même de fiducie de données personnelles, supposerait que les données constituent des biens à part entière. La perspective, certes vertigineuse, d'en faire des biens et des objets de commerce pourrait, et devrait, amener les mêmes interrogations que celles entourant la disponibilité du corps humain : peut-on disposer de manière absolue de nos données personnelles, y compris à des fins commerciales ? Peut-on interdire certains usages des données personnelles à l'individu qu'elles concernent ? Peut-on parler de propriété des données personnelles[344] ?

[343] Le qualificatif « économique » réfère ici non pas au coût d'un recours, mais au fait que la personne candidate lésée se trouve en recherche d'emploi. Il s'agit donc d'une situation économique vulnérable, de sorte que le temps et l'énergie consacrés à un recours pourraient produire des effets sur la qualité de la recherche d'emploi de l'individu. Avec un tiers expert, l'incidence du recours sur la recherche d'emploi du candidat serait ainsi considérablement atténuée.

[344] CRDP, *Entre propriété et liberté: 30 ans de protection des données personnelles. Regards croisés Europe/Amérique*, Colloque, Montréal, 6 novembre 2018, en ligne : <https://crdp.openum.ca/nouvelles/2018/11/20/videos-de-la-conference-entre-propriete-et-liberte-30-ans-de-protection-des-donnees-personnelles/> (consulté le 12 novembre 2021).

Dès lors, tel qu'évoqué précédemment[345], il conviendrait certainement de trancher la question de la nature, et le cas échéant, de la propriété des données personnelles, avant une transposition du mécanisme aux données personnelles.

En outre, nous l'avons relevé précédemment, le Commissariat à la vie privée milite en faveur d'un renforcement de ses pouvoirs d'enquête et de sanction afin d'adapter son œuvre à la réalité de l'IA. Au Québec, la CAI pourrait également suivre cette évolution de ses pouvoirs et ressources. En effet, le Commissariat fédéral entend se fonder sur l'EFVP (évaluation des facteurs relatifs à la vie privée) pour contrôler en amont l'impact qu'une IA aura à la fois sur les droits de la personne, et sur la société de manière générale.

Il s'agit donc d'une approche fondée sur la gestion du risque que représenterait un algorithme d'IA, quel que soit son domaine d'application. Aussi, inspiré de la perspective de l'UE, le Commissariat à la vie privée du Canada recommande, pour sa part, un renforcement des obligations en fonction du caractère sensible de l'utilisation de l'IA.

Or en Europe, comme au Canada, l'utilisation d'IA dans le recrutement est réputée à « haut risque » pour le droit à l'égalité des individus[346]. Dès lors, plus qu'une EFVP des IA de recrutement, peut-être faudrait-il envisager un système de labellisation ou d'autorisation par les autorités indépendantes que constituent le Commissariat et ses homologues provinciaux (en l'occurrence la CAI, pour le Québec). Avant son déploiement sur des individus, l'IA de recrutement serait soumise à une EFVP remise au Commissariat ou à la CAI. Ce serait sur le fondement de cette EFVP, et peut-être d'audits complémentaires, que cette dernière autoriserait ou non l'IA de recrutement. Une telle voie parait prometteuse, voire plus réaliste que celle de la fiducie de données. D'abord, parce que l'EFVP interviendrait avant la mise en circulation d'une IA de

[345] C.f supra*, a) Du consentement des candidats vulnérables à des mécanismes d'autorisation par un tiers expert*, p.150, note 341.

[346] C. MULLER, préc., note 148, p. 21. « compte tenu de son importance pour les particuliers et de l'acquis de l'UE sur l'égalité en matière d'emploi, l'utilisation d'applications d'IA dans les procédures de recrutement et dans des situations ayant une incidence sur les droits des travailleurs serait toujours considérée comme étant "à haut risque" […]».

recrutement, ce contrôle permettrait de réduire considérablement l'exposition des candidats aux risques de discrimination à l'embauche. Ensuite, l'EFVP ne suppose pas la réponse à des enjeux éthiques et aussi délicats que l'appropriation des données. Enfin, cette EFVP est à la charge des entreprises qui concevraient ou utiliseraient les IA de recrutement. Par conséquent cet outil, déjà connu des entreprises et professionnels experts, pourrait également constituer une voie à part entière de neutralisation des risques de discriminations. De même, si la fiducie de données devait entrer dans le droit positif, on peut imaginer un arrimage avec l'EFVP : les tiers experts, fiduciaires, pourraient s'appuyer sur les EFVP des IA utilisées par les employeurs, et ainsi, exercer leurs contrôles ou requérir des informations supplémentaires de la part des recruteurs.

Par ailleurs, si l'EFVP permet d'exiger une documentation renforcée des IA de recrutement, le label ou l'autorisation institutionnels entraînent une plus grande lisibilité : il sera possible de distinguer les technologies de recrutement à risque de celles suffisamment fiables pour faire l'objet d'une validation par un tiers institutionnel. Plus encore, un tel procédé rendrait accessibles les méthodes précitées des garanties théoriques, ou de l'équité contrefactuelle.

Enfin, en arrimant ce système au mécanisme de la fiducie des données, le fiduciaire des données aurait la responsabilité de vérifier les autorisations et EFVP récentes de l'IA de recrutement utilisée. Un tel arrimage aurait notamment l'avantage de fluidifier l'étape de la vérification par le fiduciaire, et par là même le processus de recrutement.

b) La délégation de l'exercice du droit d'ester contre les discriminations : une voie vers des procédures accessibles aux personnes candidates lésées

Tant la fiducie des données que le contrôle d'un tiers institutionnel, conduisent à une forme de délégation du droit de contestation des personnes candidates. À tout le moins si le droit de recours demeure celui de la personne candidate discriminée par l'IA de recrutement, c'est ce tiers (professionnel ou institutionnel) qui actionnerait la procédure répressive.

S'agissant de la fiducie des données, le tiers professionnel n'est autre que le fiduciaire.

On peut penser à cet égard aux avocates et avocats spécialistes de la protection des données personnelles, mais également aux personnes auditrices des technologies de l'information. Peut-être aussi qu'un tel système de protection des données personnelles créera de nouvelles professions. Quoiqu'il en soit, l'idée d'une fiducie de données présenterait plusieurs avantages dans le cas d'un recrutement par IA.

Premièrement, ce ne serait plus la personne candidate qui assumerait la responsabilité de consentir ou non au risque de discrimination que présente l'IA de recrutement. C'est le fiduciaire des données qui, après vérifications et contrôle du procédé, serait chargé d'autoriser ou non le transfert des données nécessaires au processus de recrutement par IA. La conséquence directe est que l'employeur ne communiquerait plus uniquement avec un candidat vulnérable : sur l'accès aux données personnelles d'un individu, il traiterait avec le fiduciaire de ses données qui, lui, conserve *a priori* toute sa liberté d'agir dans l'intérêt de la protection des données qui lui sont confiées. Cette première voie de délégation de responsabilité du consentement présenterait donc l'avantage de ne plus se reposer sur la liberté théorique d'un candidat en situation de vulnérabilité économique.

Deuxièmement, en cas de révélation de biais affectant l'IA de recrutement d'un employeur, le fiduciaire serait habilité à engager toutes les procédures nécessaires à la contestation de la décision algorithmique litigieuse concernant le constituant de la fiducie. En d'autres termes, lorsque A subira une discrimination à l'embauche par IA, cela signifiera d'abord que son fiduciaire B avait validé cette transmission des données au cours du processus ; mais cela impliquera ensuite que B sera habilité à engager toute contestation de la décision algorithmique litigieuse. De même, c'est le fiduciaire B qui assurera le suivi de la contestation, ou des procédures judiciaires nécessaires à la protection du droit fondamental à l'égalité de A.

En revanche, il nous faut remarquer qu'un tel mécanisme nécessitera des réponses à plusieurs égards : que se passera-t-il si le candidat vulnérable souhaite communiquer ses informations en contradiction avec l'avis du fiduciaire ? Le constituant demeure-t-il propriétaire de ses données ? Le contrat unissant le constituant au fiduciaire sera-t-il à titre onéreux ? Et dans l'affirmative, les candidats auront-ils un accès égalitaire à ce mode de gouvernance de leur universalité

numérique ? En effet, nonobstant le gain d'équilibre entre un candidat et l'employeur, demeure la question du coût que représenterait ce régime de fiducie. Un tel travail méticuleux de contrôle, d'audit et de vérifications, conduit à s'interroger qui paierait le recours à des tiers experts fiduciaires. Si les postulants devaient assumer seuls le coût de tels services, la protection du droit à l'égalité face à une IA de recrutement se retrouverait conditionnée aux ressources financières des individus. Une telle perspective amplifierait les discriminations et les inégalités, alors même que la fiducie de données aurait pour objectif de les réduire. Aux côtés des questions énumérées précédemment, il s'agit là d'un enjeu qui impliquerait une véritable volonté politique des législateurs québécois et fédéral. Ces questions, non exhaustives, mériteront probablement des réponses claires et cohérentes entre elles afin d'assurer l'efficacité de la fiducie de données.

Aussi, l'intervention d'un tiers professionnel à titre de fiduciaire pourrait apparaître complémentaire, voire préférable, à celle qui repose sur un tiers institutionnel. Dans cette seconde hypothèse, un mécanisme d'autorisation de l'autorité de protection de la vie privée complèterait les procédures actuelles, qui elles, reposent sur le signalement. En effet, aujourd'hui en cas de fuite de données, de pratiques attentatoires au droit à la vie privée ou à la protection des données personnelles, l'entreprise concernée procède à un signalement de l'incident auprès de la CAI au Québec, ou du Commissariat au niveau fédéral. Dans l'hypothèse de pratiques douteuses, ce sont les individus qui effectuent ce signalement. Une fois l'atteinte portée à la connaissance de l'autorité de protection de la vie privée compétente, cette dernière ouvre une enquête. Un tel procédé pourrait être étendu à l'IA de recrutement : en cas de pratique à risque pour le droit à l'égalité des candidats, un signalement de ces derniers pourrait attirer l'attention de l'autorité compétente. Celle-ci effectuerait alors elle-même les vérifications nécessaires, et le cas échéant, elle engagerait des poursuites contre l'employeur. Un tel paradigme s'inscrirait dans la continuité des procédures à l'œuvre en matière de protection des données personnelles, tout en favorisant un accès à la justice des candidats soumis à l'IA de recrutement.

En premier lieu, la vérification des biais potentiels affectant l'IA de recrutement ne reposerait plus sur une personne candidate, mais sur une autorité publique. Dotées des ressources nécessaires

à de telles vérifications, cette autorité disposerait, de plus, du pouvoir d'y contraindre l'employeur. Dans de telles circonstances, l'équité contrefactuelle, l'explication locale, les garanties théoriques et les preuves statistiques constituent autant de moyens de contrôler *a posteriori* l'existence de biais discriminatoires affectant l'IA de recrutement litigieuse. La charge de la preuve ne pèsera donc plus sur la personne lésée, mais sur le tiers institutionnel.

En second lieu, ce système d'autorisation place l'étape de gestion des risques de discrimination en amont de l'utilisation de l'IA biaisée sur des candidats. Un régime d'autorisation pour les IA de recrutement permettrait de s'assurer que les employeurs recourent uniquement à des IA validées, et donc reconnues comme présentant le moins de risques discriminatoires. Puis, la documentation produite au soutien de la demande d'autorisation pourrait fonder un contrôle *a posteriori*. Ainsi, le tiers institutionnel procèderait à un contrôle dual : dans un premier temps, il s'agira de neutraliser le plus de risque de discrimination possible par le biais de l'EFVP et de l'obligation de documentation ; dans un second temps, il s'agira de vérifier si l'IA conserve les conditions favorables à son autorisation, au cours de son déploiement, puis tout au long de son utilisation. Du point de vue des personnes candidates, l'intervention d'un tiers expert dans leur relation avec des employeurs usant d'IA de recrutement présente des bénéfices certains.

En outre, la délégation de l'exercice du droit de contestation leur épargnerait des frais procéduraux coûteux. En comparaison avec un fiduciaire tiers expert, l'avantage du tiers institutionnel réside dans la facilité relative à organiser l'accessibilité de ce système de gouvernance de cette universalité numérique[347]. En effet, si dans un régime de tiers expert, nous avons évoqué le risque d'un accès inégal à ces services pour les candidats, dans le cas du tiers institutionnel, l'accès à la fiducie de données pour les individus ne dépendra que de la volonté politique[348]. Ainsi, sous réserve de ressources allouées à cela, le régime passant par un tiers institutionnel simplifierait peut-être la question du financement de la fiducie de données. Au

[347] Pour rappel, le terme « universalité numérique » désigne l'universalité fictive englobant les droits, renseignements et données personnelles que crée le constituant de la fiducie de données. Voir Supra *a) Du consentement des candidats vulnérables à des mécanismes d'autorisation par un tiers expert.*

[348] Relevons par ailleurs que même sans un régime de fiducie, les droits des individus seront mieux protégés lorsque les autorités de protection de la vie privée disposeront d'un financement à la hauteur de leurs tâches.

surplus de cet avantage, il y aurait également un impact des délais procéduraux moins pesant dans la recherche d'emploi, puis les risques de représailles ou de conséquences pour la réputation professionnelle du candidat seraient restreints. La partie demanderesse serait donc le tiers expert, et non le candidat lésé.

Enfin, du point de vue des organes de contrôles des IA, le système de signalement ou de fiducie de données, leur offre une plus grande assurance d'exercer un contrôle *a posteriori* des IA de recrutement. Contourner la vulnérabilité facilite donc la contestation ou le signalement, ce qui constitue autant de possibilités de mettre un terme aux éventuelles inégalités de traitement entre les candidats soumis à des IA. La prééminence des droits fondamentaux sur les intérêts commerciaux ferait ainsi l'objet d'une réaffirmation claire, préservant ainsi les droits de la personne d'un nivellement de leur valeur par le bas.

En conclusion, l'intervention de tiers experts qu'ils soient des professionnels formés ou une institution de protection de la vie privée, constituerait une alternative au consentement désuet actuel. Pour autant si les deux voies envisagées répondent à l'obstacle de la vulnérabilité des candidats, d'autres mécanismes sont à étudier s'agissant de l'opacité de l'IA de recrutement.

2. L'exigence de conformité aux droits fondamentaux : une source de responsabilisation des acteurs de l'IA de recrutement

Le législateur peut adopter trois approches différentes face à un secteur d'activité évolutif : la régulation, l'autorégulation et la co-régulation.

Il s'agira de définir ces trois approches, avant d'étudier différents mécanismes de responsabilisation des acteurs de l'IA de recrutement.

a) La responsabilisation des acteurs de l'IA de recrutement par la régulation

Après avoir défini et présenté les limites de l'autorégulation et de la co-régulation, nous verrons en quoi l'approche de régulation correspond davantage aux enjeux de l'IA de recrutement.

L'autorégulation consiste à respecter des règles élaborées par et pour les acteurs d'un secteur d'activité[349], alors que la co-régulation résulterait de négociations entre les autorités et les opérateurs.

> « Aussi appelée régulation par le marché, l'autorégulation est un phénomène de privatisation de la production normative […]. Le concept signifie, par essence, que les règles encadrant le comportement dans un marché sont développées, administrées et appliquées par les personnes (ou leurs représentants) dont le comportement doit justement être encadré »[350]

Il faut remarquer ici que le Commissariat à la vie privée exclut fermement la voie de l'autorégulation[351]. À mi-chemin entre l'autorégulation et la régulation, la corégulation, quant à elle, consisterait en une forme de complémentarité entre la régulation par le marché et par l'État.

> « Ce concept permettrait de combiner réglementation étatique et régulation privée et de les rendre complémentaires. C'est une forme de gestion coopérative qui présente, à la fois, des éléments d'autorégulation et des éléments de réglementation associant les acteurs privés aux acteurs publics ; ces derniers ayant pour mission de garantir le respect de certaines règles entre les acteurs, et, en particulier, un équilibre des forces entre sphères économique et sociale. »[352]

---

[349] Bertrand DU MARAIS, « Analyses et propositions pour une régulation de l'Internet », (2002) 7-2 *Lex Electronica*, en ligne : <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/9530> (consulté le 18 novembre 2020); Pierre TRUDEL, « Les Effets Juridiques de l'Autoreglementation », (1988) 19-2 *R.D.U.S.* 247-286, en ligne : <https://heinonline.org/HOL/P?h=hein.journals/rdus19&i=279> (consulté le 18 novembre 2020).

[350] Karim BENYEKHLEF, *Une possible histoire de la norme: les normativités émergentes de la mondialisation*, 2e édition, Montréal, Éditions Thémis, 2015, p. 740-741.

[351] Remarquons cependant que la pratique de la consultation du Commissariat à la Vie privée dénote une tendance à une forme de co-régulation, croisant ainsi les perspectives avant d'influencer l'œuvre législative. Ceci conforte le caractère préparatoire de la co-régulation mentionné par le Pr. Poullet.

[352] K. BENYEKHLEF, préc., note 350, p. 743.

Faisant écho au caractère négocié de la norme co-régulée, le Pr Poullet la qualifie de processus « pré-normatif »[353].

À la différence des deux modes d'encadrement susmentionnés, la régulation résulterait d'un rapport strictement vertical : le législateur impose la norme aux justiciables opérant dans un secteur donné, là où la co-régulation et l'autorégulation relèveraient de rapports horizontaux avec ces justiciables. Alors en quoi la régulation présente-t-elle une meilleure garantie de protection des droits fondamentaux face à l'IA de recrutement ?

Plus qu'un enjeu de collaboration ou d'exercice de l'autorité législative, le choix de l'approche législative découle notamment des différents intérêts à défendre. En l'occurrence l'IA, et plus largement les technologies de l'information, révèlent la contradiction entre d'un côté, la protection des droits fondamentaux, piliers de la société démocratique occidentale, et d'un autre côté, la compétition technologique entre acteurs économiques, inhérente au modèle capitaliste néolibéral. Si le secteur privé se meut souvent en défenseur des intérêts économiques, la défense des droits fondamentaux, elle, ne dépend que de l'attention que le législateur entend lui accorder. Ainsi, les acteurs économiques privés plaident bien souvent l'intervention minimale, voire inexistante de l'État : l'intervention de l'État briderait l'innovation technologique, il faudrait laisser libre cours au marché qui s'équilibre naturellement[354]. L'argument souvent présenté à l'appui d'une réserve du législateur repose sur la volatilité de l'IA : ses évolutions étant rapides et imprévisibles, le rythme du législateur sera toujours plus lent que celui de l'objet sur lequel porte son œuvre. Nous verrons que ces arguments tiennent difficilement la route.

Alors, partant du postulat que les IA de recrutement intéressent le législateur, faut-il créer de nouvelles normes ? Les normes existantes suffiront-elles à encadrer l'IA au gré de l'apparition

---

[353] Yves POULLET, « Technologies de l'information et de la communication et "co-régulation" : une nouvelle approche ? », dans *Liber amicorum Michel Coipel*, Bruxelles, Kluwer, 2004, p. 167, en ligne : <http://www.crid.be/pdf/public/4731.pdf>.

[354] Catherine LARRÈRE, « Montesquieu et le « doux commerce » : un paradigme du libéralisme », *Cahiers d'histoire. Revue d'histoire critique* 2014.123.21-38, en ligne : <http://journals.openedition.org/chrhc/3463> (consulté le 22 avril 2019). « au XVIIe et au XVIIIe siècles, la reconnaissance de la capacité du commerce à réguler les passions violentes, notamment politiques, a favorisé l'acceptation des conduites orientées par le gain et a donc aidé au développement du capitalisme dans l'Europe des Lumières ».

des problématiques qu'elle pose ? En d'autres termes, faut-il légiférer ou appliquer les lois existantes ? C'est le débat « Law of the Horse » qui a opposé le Pr L. Lessig[355] au juge F. Easterbrook à propos de la réglementation du cyberespace.

D'une part, selon F. Easterbrook,[356] plutôt que de légiférer sur les chevaux, on leur applique le droit des biens, le droit des contrats, des courses, etc. Par analogie, le cyberespace ne nécessitait donc pas d'intervention législative : l'autorégulation pouvait se développer dans le cadre des normes existantes.

> « Error in legislation is common, and never more so than when the technology is galloping forward. Let us not struggle to match an imperfect legal system to an evolving world that we understand poorly. Let us instead do what is essential to permit the participants in this evolving world to make their own decisions. That means three things: make rules clear; create property rights where now there are none; and facilitate the formation of bargaining institutions. Then let the world of cyberspace evolve as it will, and enjoy the benefits. »

Cette perspective, particulièrement proche des arguments néolibéraux, omet cependant un élément : la souveraineté de l'État signifie qu'il est seul à définir les limites de son intervention.

S'agissant des données personnelles, les législateurs se sont saisis de la question en dépit d'une évolution technologique rapide. De même, malgré la technicité des processus pharmaceutiques, ou celle des marchés financiers, les États choisissent une régulation stricte de ces activités privées. Aujourd'hui, on voit émerger en outre, des algorithmes d'IA destinés à soutenir la lutte contre l'évasion fiscale des administrations fiscales[357].

La question ne porte donc pas tant sur la capacité du législateur à appréhender les nouvelles technologies, mais bien sur sa volonté de s'y attarder. De plus, s'agissant des IA, il faudrait relever un besoin croissant d'intervention législative, au point que certaines multinationales la sollicitent. Ce fut notamment le cas de Microsoft a encouragé un véritable lobbying en faveur d'une

---

[355] Lawrence LESSIG, « The Law of the Horse: What Cyberlaw Might Teach », (1999) 113-501 *Harvard Law Review*, en ligne : <https://cyber.harvard.edu/works/lessig/finalhls.pdf>.

[356] Frank H. EASTERBROOK, « Cyberspace and the Law of the Horse », *U Chi Legal F 207* 1996.

[357] note 15; A. BERGER, préc., note 15.

régulation de la reconnaissance faciale[358]. Compte tenu de sa sensibilité[359], il ne serait pas irréaliste d'anticiper des demandes similaires s'agissant de l'IA de recrutement.

Ensuite, une autre analogie entre le cyberespace et l'IA de recrutement mérite qu'on s'y attarde : le Pr Lessig relevait que le code informatique contraint et détermine autant le cyberespace, que l'architecture pour la ville.

> « And finally the architecture of cyberspace, or its code, regulates behavior in cyberspace. The code, or the software and hardware that make cyberspace the way it is, constitutes a set of constraints on how one can behave. […] [Code writers] embed certain values, or they make the realization of certain values impossible. In this sense, these features of cyberspace also regulate, just as architecture in real space regulates »[360].

Or, les algorithmes d'IA se composent de code informatique. Nous avons déjà vu que les perceptions du monde des programmeurs et des data scientist se reflètent dans les algorithmes d'IA et leurs données d'apprentissage. Dès lors, la conception de l'IA de recrutement, son « architecture », borne pareillement les effets et la marge de contrôle de la technologie. C'est d'ailleurs tout l'intérêt d'une mitigation des risques de transmission des biais humains aux bases de données des IA. Compte tenu de l'incidence de l'IA de recrutement sur la situation des personnes candidates[361], et sans minimiser les arguments défendus par F. Easterbrook, il faudrait

---

[358] B. SMITH, préc., note 85; Brad SMITH, « Facial recognition: It's time for action », *Facial recognition: It's time for action*, en ligne : <https://blogs.microsoft.com/on-the-issues/2018/12/06/facial-recognition-its-time-for-action/>; Zone Techno- ICI.RADIO-CANADA.CA, « Pétition d'employés de Google contre une collaboration avec le Pentagone », *Radio-Canada.ca*, en ligne : <https://ici.radio-canada.ca/nouvelle/1101433/employes-google-petition-contre-collaboration-pentagone>; note 17.

[359] préc., note 148. Rappelons que l'UE cite l'IA de recrutement dans la catégorie des IA à « haut risque ».

[360] L. LESSIG, préc., note 355, 508.

[361] Adrián TODOLÍ-SIGNES, « Algorithms, artificial intelligence and automated decisions concerning workers and the risks of discrimination: the necessary collective governance of data protection », (2019) 25-4 *Transfer: European Review of Labour and Research* 465-481, 6, note 8, DOI : 10.1177/1024258919876416. « When an algorithm is in command, minorities will tend to be at a disadvantage. […] As there are always fewer data available on minorities (race, religion, sexual orientation, etc.), this will lead the algorithm to understand that making a decision in favour of a minority group is riskier than making one in favour of a majority group (Hardt, 2014). In other words, to select a candidate from a minority group the algorithm will demand (by default) more qualities, aptitudes, knowledge, etc. than if it selects someone from a majority group, simply due to the fact that it is easier to predict (statistically) the behaviour of a candidate belonging to the latter group » on peut lire en complement à la note 8 : « The same happens if the decision is not made by the algorithm but when the algorithm simply classifies workers and the final decision is made by the human resources manager. »

rejoindre les nuances apportées par le Pr Lessig : des normes contraignantes spécifiques à l'IA, et ici, à l'IA de recrutement, embrasseraient mieux les défis et enjeux de conformité aux droits fondamentaux.

En conclusion, le choix de régulation qui répondrait au mieux aux besoins exprimés par les parties prenantes résiderait dans la régulation. Les individus autant que les secteurs académiques et privés souhaitent manifestement une actualisation des législations. En outre, bien que les processus de modernisation de l'arsenal législatif soient lancés, le sens de cette modernisation mérite peut-être des réflexions plus approfondies. Nous verrons donc dans les prochains développements l'intérêt d'une régulation résiliente, mais ferme.

b) Une exigence de conformité en amont par la définition d'une responsabilité professionnelle : l'intégration des principes éthiques dans le droit positif

Dans le chapitre précédent, la principale source de biais discriminatoires a été identifiée : la base de données des IA de recrutement. Qu'il s'agisse de base de données d'entraînement ou de base de données alimentées tout au long de l'activité de l'IA, les biais discriminatoires au cœur de notre sujet émanent de l'être humain. Les travailleurs du clic, data scientist, personnes mathématiciennes ou programmeuses, représentent autant de sources de biais contaminant les IA. Concernant l'IA de recrutement, les professions du recrutement et des ressources humaines évoluent dans un secteur particulièrement discriminatoire[362]. Aussi, des mécanismes préventifs

[362] Myrlande PIERRE, Manon POIRIER, Jean-Philippe BEAUREGARD et Paul EID, *Discrimination à l'embauche : ce que nous disent les recherches sur le testing de CV ?*, 11 février 2021; Au cours de ce panel, les sociologues Paul EID et Jean-Philippe BEAUREGARD précisaient que dans leur *testing* de CV, le secteur des ressources humaines était le plus discriminatoire à l'embauche des personnes racialisées.

Jean-Philippe BEAUREGARD, *Les frontières invisibles de l'embauche des Québécois minoritaires : hiérarchie ethnique, effet modérateur du genre féminin et discrimination systémique : dévoiler la barrière à l'emploi par un testing à Québec*, Thèse de doctorat, Université de Laval, 2020 Accepted: 2020-10-12T07:22:20Z, p. 97, en ligne : <https://corpus.ulaval.ca/jspui/handle/20.500.11794/66769> (consulté le 25 août 2021) « C'est dans la catégorie des emplois en ressources humaines que nous avons observé le plus haut taux net de discrimination à l'encontre des minoritaires, soit 72,7%. Le taux le plus élevé de l'étude de la CDPDJ (40%) avait aussi été mesuré dans l'accès aux emplois en RH. »;

Paul EID, « Les inégalités « ethnoraciales » dans l'accès à l'emploi à Montréal : le poids de la discrimination », (2012) 53-2 *rs* 415-450, 439, DOI : 10.7202/1012407ar "« Les secteurs des ressources humaines et du secrétariat sont ceux où l'écart entre les probabilités de rappel des deux candidats sont les plus grandes […]. Ces deux mêmes secteurs

responsabilisant l'ensemble des professionnels intervenant dans la conception de l'IA de recrutement constitueraient un premier levier de mitigation du risque de discrimination à l'embauche. Le fair learning, le data cleansing, les échantillons de données diversifiées représentent des méthodes émergentes de neutralisation des transmissions de biais humains aux IA de recrutement. Ces méthodes inspirées de principes éthiques reconnus ne représentent cependant aucune contrainte. Le mécanisme doit donc représenter une contrainte qui prendrait sa source dans la norme légale ou réglementaire. Une obligation de mitiger les risques de biais discriminatoires de l'IA de recrutement fonderait un mécanisme de protection en amont de la discrimination à l'embauche. C'est là tout l'intérêt d'inclure de telles obligations dans un régime juridique spécifique : la responsabilité professionnelle des concepteurs d'IA.

Les ordres professionnels encadrent certaines professions libérales afin d'assurer la protection du public. Il s'agit d'assurer l'intégrité du travail de personnes exerçant des métiers susceptibles de causer des dommages importants à la population. On retrouve cette organisation ordinale chez les avocats, les ingénieurs ou encore les architectes. Pour chaque profession structurée par un ordre professionnel, il existe un régime de responsabilité professionnelle[363]. Assurant l'application des lois et règlements applicables à ces professionnels libéraux, les ordres professionnels procèdent notamment au contrôle et à la sanction disciplinaire en cas de manquement de leurs membres. Il s'agit par exemple de contrôler l'accès à la profession, d'assurer une formation continue obligatoire des membres, ou encore d'inspecter les professionnels dont les pratiques susciteraient des doutes. L'ensemble de ces mesures d'encadrement professionnel ont vocation à garantir la sécurité du public dans le cadre des services professionnels qui lui sont proposés.

Appliqué aux concepteurs de l'IA, un tel système d'encadrement des professions permettrait de garantir une haute qualité des pratiques, et plus spécifiquement, de pratiques à

---

d'emploi se démarquent également des autres par leurs taux nets de discrimination plus élevés, soit de 40 % pour les postes en ressources humaines et de 38,8 % pour les postes de secrétaire. ».

[363] INSTITUT CANADIEN DES COMPTABLES AGRÉÉS, *Dictionnaire de la comptabilité et de la gestion financière*, 2006, « responsabilité professionnelle », en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=506448> (consulté le 13 août 2021). « Ensemble des responsabilités que le professionnel libéral peut encourir dans l'exercice de sa profession, à savoir la responsabilité disciplinaire, la responsabilité civile et la responsabilité pénale. »

la pointe de la gestion du risque de discrimination par IA. Les professionnels concepteurs de l'IA de recrutement seraient donc soumis à une obligation de formation continue sur les biais, les nouvelles méthodes de fair learning, ou le data cleansing par exemple.

De même, suivant la proposition de C. O'Neil, ces professionnels de l'IA pourraient être soumis à des lois et règlements spécifiques, mais également à un code éthique ou déontologique. En effet, C. O'Neil propose l'équivalent du serment d'Hippocrate pour les data scientist :

> « How do we start to regulate the mathematical models that runs more and more of our lives? I would suggest that the process begin with the modelers themselves. Like doctors, data scientists should pledge a Hippocratic Oath that focuses on the possible misuses and misinterpretations of their models. »[364]

L'autrice plaide cependant l'insuffisance et les limites d'engagements éthiques qui, de plus, ne s'appliqueraient qu'à des professionnels soumis aux pressions de leurs clients ou employeurs. Autrement dit, l'éthique n'engage à rien.

> « […] the Hippocratic Oath ignores the on-the-ground pressure that data scientists often confront when bosses push for specific answers. To eliminate [Weapons of Math Destruction], we must advance beyond establishing best practices in our data guild. Our laws need to change, too. And to make that happen we must reevaluate our metric of success.
>
> Today, the success of a model is often measured in terms of profit, efficiency, or default rates. It's almost always something that can be counted. »

On retrouve encore ici cette idée d'une complémentarité entre l'existence de principes éthiques et la nécessité de leur donner une force contraignante à travers la régulation. L'objectif d'une responsabilisation des acteurs de l'IA de recrutement ne se réaliserait ainsi qu'à deux conditions : la conformité aux droits fondamentaux doit constituer une exigence stricte, mais doit également établir une exigence légale ou réglementaire. Pour la réalisation de ces deux conditions, le législateur devra choisir la régulation. Notons à ce propos que les recommandations du Commissariat à la vie privée, tout comme les orientations des gouvernements fédéral et québécois, se dirigent vers un encadrement de l'IA par la loi. Il serait souhaitable que la logique

[364] C. O'NEIL, préc., note 248, p. 276.

aboutisse à une réaffirmation de la prééminence des droits fondamentaux sur les intérêts commerciaux et la course aux technologies.

Par ailleurs, en quoi cette régulation a-t-elle intérêt à intégrer les principes éthiques d'IA ? À titre liminaire, il faut rappeler que l'éthique permet de tracer les bornes de comportements et pratiques considérés intègres dans un secteur d'activité donné.

> « Étude des valeurs et principes moraux qui s'appliquent ou qui devraient s'appliquer aux gens d'un milieu, aux personnes exerçant une même fonction ou profession ou aux membres d'un corps. »[365]

L'IA n'a pas échappé à ce besoin d'éthique et d'intégrité. En effet, en réaction aux impacts importants de l'IA sur les individus, et sur la société, plusieurs corpus éthiques ont ainsi émergé sur la scène internationale. Or, malgré la pluralité de corpus éthiques en IA, on y retrouve sensiblement les mêmes principes clés. Compte tenu du consensus émergeant autour de 5 à 7 principes en éthique de l'IA[366], ils paraissent tout indiqués dans l'élaboration d'une régulation de l'IA de recrutement. À ce propos, A. BENSAMOUN et. G. LOISEAU rappellent que l'éthique de l'IA et le droit positif présentent une plus grande proximité qu'il n'y parait[367].

[365] Michel FILION, « Éthique », dans Dictionnaire encyclopédique du Droit québécois, Gaudet Éditeur ltée, en ligne : <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26548265> (consulté le 24 novembre 2020).

[366] Alexandra BENSAMOUN et Grégoire LOISEAU, « Chapitre 1- Intelligence artificielle et l'éthique », dans *Droit de l'intelligence artificielle*, coll. Les Intégrales, 2110-9680, 15, Issy-les-Moulineaux, LGDJ, 2019 aux pages 15-27.

[367] *Id.* aux pages 32-33.

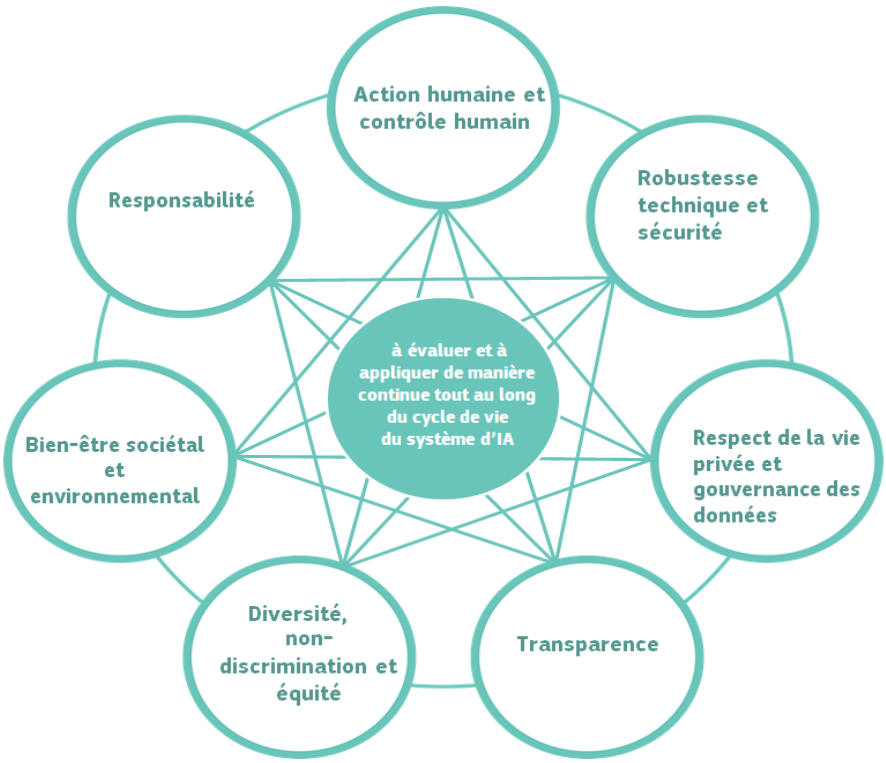

Figure 3. – Interrelation des sept exigences[368]

Citons notamment le principe de prudence de la Déclaration de Montréal[369] qui rappelle le principe de non-malfaisance, qu'on retrouve dans l'UE, qu'il semble intrinsèquement lié à la prohibition constitutionnelle des discriminations[370]. Aussi, compte tenu de leur haute compatibilité avec les droits fondamentaux, les principes éthiques peuvent paver la voie à une régulation résiliente de l'IA.

---

[368] GROUPE D'EXPERTS DE HAUT NIVEAU SUR L'INTELLIGENCE ARTIFICIELLE, préc., note 304, p. 18. « elles revêtent toutes une importance égale, elles se soutiennent mutuellement et devraient être appliquées et évaluées tout au long du cycle de vie d'un système d'IA »

[369] C. ABRASSART et al., préc., note 239, p. 15. « Toutes les personnes impliquées dans le développement des SIA doivent faire preuve de prudence en anticipant autant que possible les conséquences néfastes de l'utilisation des SIA et en prenant des mesures appropriées pour les éviter. […] 3. Avant d'être mis sur le marché, qu'ils soient payants ou gratuits, les SIA doivent satisfaire des critères rigoureux de fiabilité, de sécurité et d'intégrité, et faire l'objet de tests qui ne mettent pas en danger la vie des personnes, ne nuisent pas à leur qualité de vie et ne portent pas atteinte à leur réputation ou leur intégrité psychologique. Ces tests doivent être ouverts aux autorités publiques compétentes et aux parties prenantes concernées. »

[370] *Charte Canadienne des Droits et Libertés*, préc., note 26; *Charte des Droits et Libertés de la Personne*, préc., note 28.

En premier lieu, s'appuyer sur l'éthique des IA autorise la construction d'une régulation dont les notions clés bénéficient d'une large acceptation dans les Big Tech. Leur intégration dans les normes de droit positif présente donc l'avantage de renforcer une transformation dans laquelle le secteur de l'IA s'est déjà engagé[371]. De plus, les GAFAM comme les Big Tech ont activement contribué à la définition de ces corpus éthiques. L'intérêt de leur intégration dans une stratégie de régulation réside d'abord dans l'uniformisation de l'application de ces normes.

Ensuite, en se saisissant de ces principes, le législateur fédéral leur donne une valeur contraignante de nature à engager la responsabilité des entreprises. Concrètement, cela signifie que ces dernières n'auront plus le loisir d'interpréter leurs engagements éthiques comme bon leur semble. Dans le contexte des IA de recrutement, c'est du droit fondamental à l'égalité dont il est question. Soumettre son respect à la casuistique relèverait de l'absurde, *a fortiori* s'il s'agit d'accès à un moyen de subsistance aussi vitale qu'un emploi. À l'échelle du recrutement augmenté par IA, la responsabilité professionnelle, pensée comme un levier de mitigation des risques de biais, implique donc que les candidats seront protégés par les mêmes normes, interprétées de la même façon, sans que le sens éthique de l'employeur puisse donner lieu à une disparité de protection contre les discriminations à l'embauche. La sécurité juridique des personnes candidates, tout comme celle des employeurs recourant à ces technologies, commande en effet la recherche d'une uniformité d'interprétation de la norme.

Plus encore, le contrôle et la sanction de la norme génèrent à la fois protection et prévisibilité aux différentes parties prenantes de l'IA de recrutement. En effet, le secteur de l'IA n'est pas étranger à la casuistique discriminatoire. Les IA de reconnaissance faciale, fortement biaisées, ont conduit à une opposition entre les GAFAM quant à leurs orientations éthiques, voire politiques[372], et ce, alors même que ces entreprises partagent les mêmes engagements éthiques[373]. L'actualité des

---

[371] A. BENSAMOUN et G. LOISEAU, préc., note 366, par. 53 à 56.

[372] B. SMITH, préc., note 85; note 17; RADIO-CANADA, « Google et Microsoft en désaccord sur l'interdiction de la reconnaissance faciale », *Radio-Canada.ca*, sect. Zone techno (21 janvier 2020), en ligne : <https://ici.radio-canada.ca/nouvelle/1483573/reconnaissance-faciale-moratoire-interdiction-banissement-temporaire-5-ans-europe-sundar-pichai-brad-smith> (consulté le 14 août 2021).

[373] Alex HERN, « "Partnership on AI" formed by Google, Facebook, Amazon, IBM and Microsoft », *the Guardian*, sect. Artificial intelligence (28 septembre 2016), en ligne :

deux dernières années démontre, malheureusement, que chaque organisation interprète à sa façon des principes aussi flous que « Le principe de la prévention de toute atteinte »[374]. La conservation de principes comme ceux de non-discrimination ou de prudence dans la sphère éthique expose donc à une multiplication des normativités individuelles[375], sources de casuistique[376].

En second lieu, le caractère « déterritorialisé » des principes éthiques représente un autre atout : leur intégration dans le droit positif offre l'opportunité d'une convergence des lois encadrant l'IA. Qu'elle soit programmée pour faciliter le recrutement, ou qu'elle serve à l'identification de suspects et de personnes disparues[377], l'IA a la particularité de transcender les

<http://www.theguardian.com/technology/2016/sep/28/google-facebook-amazon-ibm-microsoft-partnership-on-ai-tech-firms> (consulté le 15 août 2021). « Google, Facebook, Amazon, IBM and Microsoft are joining forces to create a new AI partnership dedicated to advancing public understanding of the sector, as well as coming up with standards for future researchers to abide by. »

[374] GROUPE D'EXPERTS DE HAUT NIVEAU SUR L'INTELLIGENCE ARTIFICIELLE, préc., note 304, p. 15. « Les systèmes d'IA ne devraient ni porter atteinte, ni aggraver toute atteinte portée[29], ni nuire aux êtres humains d'une quelconque autre manière.[30] Cela englobe la protection de la dignité humaine ainsi que de l'intégrité mentale et physique. Les systèmes d'IA et les environnements dans lesquels ils évoluent doivent être sûrs et sécurisés. Ils doivent être robustes sur le plan technique et il convient de veiller à ce qu'ils ne soient pas exposés à des utilisations malveillantes ».

[375] Les normativités individuelles constituent les normes internes aux entreprises dans le cadre de leur mise en conformité avec les lois. En réponse au renforcement des obligations de diligence et obligation de documentation, se développent ces normes internes, complémentaires aux normes légales ou aux standards de l'industrie. À ce propos, Vincent Gautrais et Christelle Papineau nous disent : « Les conditions sont réunies pour qu'il y ait densification normative, ce phénomène théorisé par Catherine Thibierge au début des années 2000 et qui prend la forme de ce que nous appelons "délégation constatée", à savoir une délégation du législateur à des standards techniques et, conséquemment, à une normativité individuelle » dans Vincent GAUTRAIS et Christelle PAPINEAU, « L'explosion documentaire : la normativité individuelle, nouveau centre de gravité normatif », dans *Lex electronica*, coll. Hors-série « Normes énormes », n°23, Montréal, Centre de recherche en droit public, Université de Montréal, 2018, p. 143-172 à la page 156, en ligne : <https://www.lex-electronica.org/articles/volume-23/numero-23-hors-serie-2018-version-integrale/> (consulté le 13 novembre 2021).

[376] Alexandra BENSAMOUN et Grégoire LOISEAU, « La gestion des risques de l'intelligence artificielle . - De l'éthique à la responsabilité », (2017) 46-1203 *JCP G*, 4. « Le second écueil des règles éthiques renvoie à la privatisation de la norme. La règle de comportement devient le fait des acteurs privés et relève de leur bon vouloir. Que la norme éthique soit à l'initiative des autorités publiques et qu'elle ne doive qu'à la réflexion d'un groupement privé, elle implique les acteurs du secteur dans sa conception et/ou dans sa mise en œuvre. Et encore ne parle-t-on pas de tous les acteurs privés. Car cette voie fait incontestablement primer la loi du plus fort, celle des opérateurs les plus influents qui imposeront leur éthique. Le risque avait d'ailleurs déjà été identifié au sujet de la lex mercatoria »

[377] LAURÈNE CHAMPALLE, « En Inde, un logiciel de reconnaissance faciale pour retrouver les enfants disparus », *leparisien.fr* (19 mai 2018), en ligne : <http://www.leparisien.fr/international/en-inde-un-logiciel-de-reconnaissance-faciale-pour-retrouver-les-enfants-disparus-16-05-2018-7719015.php>; note 17.

frontières. Premièrement, parce qu'elle provient principalement des GAFAM et des Big Tech. Ces multinationales en situation monopolistique bénéficient d'une liberté et d'une influence mondiale telles, que leurs activités, et donc l'IA, entrent dans le champ d'un phénomène transcendant les droits nationaux : le droit global.

> « […] un droit prenant en compte les effets de la mondialisation et offrant à ses acteurs, privés comme publics, des voies possibles de régulation. [Il s'agirait] d'abord [d'] une addition de droits : les droits nationaux, les droits de la personne, le droit du commerce international, les droits transnationaux, les normes alternatives nationales et transnationales, etc. »[378].

En effet, les étapes du développement d'un système d'IA échappent déjà aux paradigmes nationaux : la base de données d'entraînement s'alimente par le travail des microtravailleuses et usagers du web, de même que des programmeurs peuvent co-écrire le code d'une IA tout en étant dispersés sur plusieurs continents. De surcroît, les GAFAM exercent une influence telle, que le débat ne porte plus sur la réalité de leur influence, mais sur les moyens de la contrebalancer.

> « La question est aussi celle de la souveraineté numérique des États, menacée par les "entreprises souveraines de l'Internet". La régulation d'une intelligence artificielle éthique participe ainsi de l'émergence d'un pouvoir périphérique délégué, voire autonome, des "GAFAM". C'est d'ailleurs le sens de l'initiative du Danemark qui envisage de nommer un ambassadeur numérique auprès des géants de la Tech. »[379]

Sans nul doute, au nombre des promesses de l'IA se trouve la capacité d'exploiter le nouvel or noir, lui aussi transnational : les mégadonnées[380]. Aussi, dans la perspective d'un droit global émergent de l'IA, l'avantage d'une intégration des principes éthiques dans les lois réside dans leur capacité à déboucher sur des standards globaux[381], à l'instar de ceux de la *Lex mercatoria*[382]. La différence entre l'IA et la *lex mercatoria* porterait sur le rapport de force à l'origine des normes :

---

[378] K. BENYEKHLEF, préc., note 350, p. 800-801.

[379] A. BENSAMOUN et G. LOISEAU, préc., note 376, 4.

[380] préc., note 4. « Ensemble d'une très grande quantité de données, structurées ou non, se présentant sous différents formats et en provenance de sources multiples, qui sont collectées, stockées, traitées et analysées […] ».

[381] Ces standards émergent d'ores et déjà par le biais des lignes directrices ou recommandation d'organismes influents tels que l'OCDE, l'UE ou encore de collège d'acteurs de l'IA comme Partnership on AI ou la *Déclaration de Montréal*.

[382] Antoine LEDUC, « L'émergence d'une nouvelle lex mercatoria à l'enseigne des principes d'UNIDROIT relatifs aux contrats du commerce international : thèse et antithèse », 2001.23.

en lieu et place des normes fixées par le plus fort économiquement, les États ont ici la possibilité de reprendre leur place de législateurs en jetant les bases du droit global de l'IA.

> « Il convient d'apprécier l'importance considérable prise par ce droit global. Il ne s'agit pas simplement de ces champs de l'activité humaine qui paraissent hors d'atteinte de la loi du souverain (comme les droits de propriété intellectuelle), mais aussi de champs d'activités dont le souverain national décide qu'ils seraient mieux encadrés ou régulés par des acteurs privés. »[383]

L'incorporation de principes éthiques faisant déjà consensus, contribuerait à l'émergence d'un droit global dont les fondements seraient admis par le secteur privé[384], en restant compatible au respect du droit à l'égalité. Cette idée renforce la perspective d'une corégulation, ici éthique, prenant la forme d'une étape pré-normative, ou une forme de norme technique[385] embryonnaire.

Enfin, il convient de relever que la responsabilité professionnelle dans certains corps de métiers s'appuie sur un code de déontologie ou un code éthique. S'agissant de l'IA de recrutement, ces vases communiquant entre éthique et responsabilité professionnelle nourrissent déjà des discours en faveur de la responsabilisation des acteurs de l'IA. Une « fondamentalisation »[386] des principes éthiques permettrait donc d'abord une déclinaison des droits fondamentaux dans la régulation de l'IA ; ensuite, ce processus d'intégration aux droits fondamentaux clarifierait la supériorité des droits fondamentaux sur l'évolution technologique de l'IA. La régulation demeure donc l'approche d'encadrement présentant le plus de garanties de

---

[383] Karim BENYEKHLEF, « Droit global : un défi pour la démocratie », *Revue Projet: comprendre pour agir* À l'heure des multinationales, le retard du droit ? (24 juin 2016), en ligne : <https://www.revue-projet.com/articles/2016-06-benyekhlef-droit-global-un-defi-pour-la-democratie> (consulté le 18 décembre 2020).

[384] A. BENSAMOUN et G. LOISEAU, préc., note 376, 5. « L'implication éthique présente enfin le grand avantage d'habiller une classique et naturelle recherche de profits, en s'achetant une vertu à moindres frais. Elle constitue alors un véritable moyen de communication et participe à l'amélioration de l'image et de la réputation de l'acteur concerné. 13. — "Une soft law pour réguler un soft power". Tel est en définitive l'enjeu de la régulation de l'intelligence artificielle par l'éthique et ses instruments dédiés. »

[385] K. BENYEKHLEF, préc., note 350, p. 746. « La norme technique constitue un modèle de conduite à portée générale, et ce quel que soit son objet car elle propose une solution à des problèmes répétés […]. Les normes techniques apparaissent comme le fruit d'une procédure déterminée et décrivent souvent un résultat. » Ces principes éthiques pourraient, par exemple, se décliner en normes ISO de gestion du risque de biais discriminatoires dans les IA, et ce, à l'instar des normes techniques de gestion des données personnelles.

[386] A. BENSAMOUN et G. LOISEAU, préc., note 366.

rééquilibrer le rapport de force entre les individus et le secteur privé[387] et, dans notre matière, entre les personnes candidates et les employeurs usant de l'IA de recrutement. La responsabilité professionnelle des concepteurs de l'IA de recrutement constituerait ainsi un premier levier de mitigation, en amont, des risques de discrimination.

Pour conclure, — et dans la perspective d'une intégration des principes éthiques de l'IA aux lois —, la responsabilité professionnelle permettrait l'application d'obligations de sûreté renforcées dès la conception des IA ; une forme d'obligation de moyens renforcée de se conformer à un principe *ethical by design*. Par la définition de normes contraignantes issues des principes éthiques[388], les professionnels de l'IA prendraient ainsi l'engagement de réduire les biais des algorithmes de recrutement. Une telle stratégie de régulation autoriserait une gestion du risque de discriminations reposant, en amont, sur la responsabilité professionnelle, et en aval, sur les tiers experts. De tels mécanismes ne produiraient cependant leurs effets qu'à la condition que les personnes dirigeantes soient, elles aussi, soumises à des obligations analogues.

c) Une responsabilité de l'employeur du fait de ses IA de recrutement : l'incitation à une gestion proactive du risque de discrimination

Bien que la responsabilité professionnelle puisse occuper une place centrale dans la stratégie de régulation des IA de recrutement, il faudrait remarquer que, seul, ce mécanisme ne conduirait qu'à une déresponsabilisation des véritables détenteurs du pouvoir au sein d'une

[387] COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, préc., note 111, part. Responsabilité démontrable. « les individus ne peuvent pas compter de manière absolue sur les organisations pour traiter correctement leurs renseignements personnels, en particulier si des décisions automatisées sont en jeu et peuvent perturber le cours de leur vie […]. Cette situation est aggravée par le profond déséquilibre de pouvoirs entre les individus et les organisations qui recourent à l'IA, caractérisé notamment par une asymétrie entre les connaissances et ressources des uns et des autres. »

[388] L'intégration de principes éthiques dans la loi, de même que la fondamentalisation de ces prncipes, évoquée précédemment (supra note 68), rappelle le phénomène d'internormativité. En effet, dans son sens premier, l'internormativité renvoie à « [ces] phénomènes d'absorption de la norme alternative par le droit étatique ; ces points d'entrée par lesquels le droit étatique s'approprie la norme alternative et, ce faisant, la transforme en règle juridique étatique. », K. BENYEKHLEF, préc., note 350, p. 802.

entreprise. Par analogie avec la *Loi n° 64*[389], une responsabilité des dirigeants apporterait un complément à celle des professionnels de l'IA. La notion de responsable du traitement automatique[390], ou de détenteur du contrôle de la technologie pourrait trouver sa déclinaison dans l'IA de recrutement. Que le législateur choisisse de responsabiliser la personne en charge des ressources humaines, ou des technologies, la simple identification d'une personne physique responsable pourrait inciter les employeurs à prêter bien plus d'égards à la conformité de leurs pratiques aux chartes.

En effet, les équipes chargées de la conception ou du déploiement d'IA de recrutement œuvrent bien souvent sous la direction d'une personne fixant les critères et objectifs de l'IA. Dès lors, en les tenant responsables des discriminations issues des IA de recrutement produites, le législateur énoncerait une obligation de moyens forcée portant sur la conformité avec les chartes. Les discriminations à l'embauche seraient ainsi imputables aux dirigeants des entreprises conceptrices ou utilisatrices d'IA de recrutement. Une telle approche permettrait d'activer deux leviers de prévention : la mise en place d'une gestion proactive du risque algorithmique d'une part, et d'autre part, un renforcement des contrôles par les autorités protectrices de la vie privée. Conjuguée à l'EFVP et à l'obligation de documentation, la responsabilité des employeurs du fait de leurs IA assurait une évaluation de l'IA de recrutement tout au long de son cycle de vie : de la conception au déploiement, en passant par l'autorisation de l'IA de recrutement, le régulateur procèderait à un contrôle continu du respect de l'obligation de sûreté.

> « Certains ont proposé la création d'une agence des algorithmes, à l'image de la Federal Drug Administration aux États-Unis qui examine, teste et approuve des médicaments couverts par des brevets et des droits de propriété intellectuelle. Un tel organisme pourrait aussi exercer des fonctions de certification et de standardisation

[389] S. DU PERRON, préc., note 143, part. Sanctions et pouvoirs de la CAI. « Le projet de loi confère à la CAI le pouvoir d'imposer aux entreprises du secteur privé des sanctions administratives pécuniaires pouvant aller jusqu'à 10 000 000 $ ou 2 % du chiffre d'affaires mondial. Le projet de loi habilite également la CAI à engager des poursuites pénales devant les tribunaux en cas de violation de la loi. Les entreprises contrevenantes font face à des amendes se situant entre 15 000 $ à 25 000 000 $ ou 4 % du chiffre d'affaires mondial. »

[390] C.f *Chapitre 1*, B., *c) La modulation des garanties dans l'encadrement de la prise de décision automatique*, p. 56-61

> pour les objets fondés sur l'IA et interagissant dans l'environnement humain (voiture autonome, drone, robot, etc.). »[391]

Une telle agence, nouvelle ou désignée parmi les autorités existantes, paraitrait tout indiquée afin de garantir aux personnes candidates et employeurs une sécurité juridique dans leurs interactions avec l'IA de recrutement. Cette stratégie de régulation offre une prévisibilité à l'entreprise conceptrice ou utilisatrice de l'IA de recrutement : une fois approuvée par l'agence, l'IA de recrutement serait réputée fiable sur le plan technologique, ou technique. Par ailleurs, au surplus de ce gain de prévisibilité, les vérifications effectuées en amont par une telle agence impliqueraient conséquemment une limitation de responsabilité totale ou partielle des entreprises en cas de discriminations à l'embauche[392]. De même, du point de vue des individus, les personnes candidates pourraient s'engager dans des processus d'embauche augmentés avec l'assurance que l'IA respecte les exigences de sûreté. L'agence assurerait une fonction aussi fondamentale que celle des ordres professionnels ou des autorités sanitaires habilitées à autoriser un médicament : le principe de précaution constituerait le cœur de la gestion du risque de discrimination que représentent les IA de recrutement.

Pour conclure, l'accès à la justice des personnes exposées aux discriminations d'IA de recrutement biaisées suppose d'abord des obligations à la charge des employeurs qui garantissent le plus de transparence possible pour les candidates et les autorités de régulation. Sur ce point, il convient d'accueillir positivement les orientations que prend le législateur. Néanmoins, si intéressantes et nécessaires que puissent être les réflexions des autorités et les projets de modernisation des lois, le prisme d'analyse demeure tardif et articulé autour de législations relatives à la vie privée et à la protection des données personnelles. Pourtant, l'exemple de la discrimination par IA de recrutement met en évidence le risque d'un contentieux mixte, à mi-chemin entre la compétence du Commissaire fédéral à la vie privée et celle de la Commission canadienne des droits de la personne. Afin de prévenir de potentiels dénis de justice, il serait peut-être judicieux de penser une stratégie de régulation de l'IA qui assure, certes la

[391] K. BENYEKHLEF, préc., note 276, part. Les bonnes pratiques de développement algorithmiques.

[392] D'ailleurs, une telle perspective faciliterait peut-être l'émergence d'une offre de régimes d'assurance incluant les risques liés aux activités privées reposant sur l'IA, et ce, au-delà du seul domaine du recrutement.

cohérence du cadre normatif existant, mais qui le transcende et pose les fondations d'un cadre à la hauteur des mutations sociales et technologiques que le développement de l'IA amène dans son sillage[393].

3. Les défis de la gestion du contentieux de l'IA de recrutement : la nécessaire clarification du parcours juridictionnel et de la stratégie de régulation

Dans cette dernière partie, il s'agira d'aborder le volet de l'application de la sanction en cas de gestion du risque insatisfaisante, en tout ou partie. À quoi ressemblerait le parcours judiciaire ou administratif d'une personne candidate lésée ? À quelle juridiction confier ce contentieux, et quels sont les enjeux d'administration de la justice ?

a) L'arrimage des autorités de contrôle existantes : la condition sine qua non d'une bonne administration du contentieux de l'IA de recrutement

Au cours des précédents développements, les différentes voies de régulation proposées s'inscrivent dans une perspective d'anticipation et de mitigation du risque de discrimination. Il s'agit de construire, ou d'étendre des mécanismes juridiques à la gestion du risque de discrimination des IA de recrutement. Néanmoins, le risque nul relevant de l'idéal, le risque de biais, ou d'erreurs de l'IA de recrutement demeure, et ce même en cas de respect zélé des obligations de moyens proposées. Aussi, la question de l'émergence d'un contentieux de l'IA de recrutement commande quelques observations ; en cas de discrimination par IA de recrutement, vers qui devrait se tourner la personne lésée par une IA de recrutement ?

Se situant à mi-chemin entre la compétence des commissions des droits de la personne, les tribunaux et les autorités de protection de la vie privée, la contestation d'une décision algorithmique de tri de candidature appellerait probablement un arrimage entre ces différentes institutions. L'étape de l'enquête, tout comme celle du procès contre une décision d'IA de

[393] préc., note 148. En comparaison, les orientations de l'UE semblent bien avancées. Peut-être que le Canada pourrait s'engager dans des réflexions aussi poussées.

recrutement soulève plusieurs questionnements. Dans un contexte québécois est-ce à la CAI de recevoir le signalement d'un traitement automatisé litigieux, puis de mener l'enquête ? L'enquête relève-t-elle plutôt de la compétence de la Commission québécoise des droits de la personne ? Puis à l'issue de cette enquête, est-ce un tribunal de droit commun qui aura la compétence de sanctionner une discrimination par IA de recrutement ? Ces interrogations soulèvent plus largement celle de l'orientation régulatrice du législateur. En matière d'atteinte aux droits de la personne, tout comme en matière de protection de la vie privée, nous l'avons vu précédemment, le législateur québécois — tout comme son homologue fédéral — a choisi une approche mixte : la prévention des comportements et pratiques attentatoires aux droits fondamentaux, et ultimement, la répression en cas d'échec et de mauvaise foi des mesures préventives. On retrouve dans cette dualité une culture de la prévention et de l'éducation du secteur privé, et par conséquent, une mitigation des risques d'atteintes aux droits fondamentaux des individus.

Or, dans la mesure où un contentieux de l'IA de recrutement se situerait au carrefour des enjeux susmentionnés, une approche similaire parait également appropriée dans la promotion d'une proactivité de la gestion de risque appliquée à l'IA de recrutement. Plusieurs formes d'arrimages pourraient s'inspirer de mécanismes connus. N'ayant pas l'ambition de l'exhaustivité, les trois formes d'arrimage présentées pourraient néanmoins constituer les étapes d'un processus de transformation juridictionnelle du contentieux de l'IA de recrutement.

En premier lieu, une première forme d'arrimage pourrait prendre la forme d'un protocole de répartition de compétence entre les commissions des droits de la personne et les autorités de protection de la vie privée. Au Québec, les litiges en matière d'emploi font l'objet d'une répartition de compétence entre la CNESST[394] et la Commission des droits de la personne : les affaires relatives au harcèlement psychologique relèvent de la première, tandis que celles touchant aux discriminations relèvent de la seconde. Une entente entre les deux organismes a permis un arrimage lorsque leurs compétences se concurrencent sur certaines affaires[395].

[394] Commission des normes de l'équité de la santé et de la sécurité du travail

[395] COMMISSION DES NORMES DE L'ÉQUITÉ et COMMISSION DES DROITS DE LA PERSONNE ET DES DROITS DE LA JEUNESSE, *Entente de collaboration entre la CDPDJ et la CNESST concernant leurs interventions en matière de harcèlement | Publications |*

Appliquée à l'IA de recrutement, une répartition de compétence analogue signifierait que, par exemple, le législateur répartirait le type de contentieux qui relèverait de la CAI et de la Commission des droits de la personne québécoise.

Pour autant, il conviendrait de remarquer le risque que ces règles de répartition excluent certains facteurs de discrimination. Le concept d'intersectionnalité des oppressions systémiques met en effet en exergue une réalité qui échappe souvent aux catégories juridiques : les systèmes d'oppressions coexistent, et s'enchevêtrent de sorte que les personnes se situant au carrefour de ces systèmes subissent des discriminations dont les motifs sont polyformes. Ainsi, la femme autochtone est exposée au racisme, au sexisme, mais également à une discrimination hybride. De même, pour les femmes racialisées subissant un sexisme raciste, ou un racisme sexiste.

Dans notre exemple précité de l'IA d'Amazon, les femmes pénalisées dans l'étude de leurs candidatures ont subi une discrimination de genre. Une analyse intersectionnelle des biais de cette IA de recrutement impliquerait donc une comparaison entre le traitement des candidatures de femmes selon leur groupe ethnoculturel, ou leur racialisation. L'audit du NIST précité[396] a d'ailleurs pris soin d'étudier les biais de race et de genre, séparément, puis conjointement. C'est ainsi que le NIST conclu que les IA de reconnaissance faciale présentent non seulement des biais plus importants à l'égard des personnes racialisées, mais également des biais sexistes concernant les femmes racialisées.

> « Among U.S.-developed algorithms, there were similar high rates of false positives in one-to-one matching for Asians, African Americans and native groups (which include Native American, American Indian, Alaskan Indian and Pacific Islanders). The American Indian demographic had the highest rates of false positives […].
>
> For one-to-many matching, the team saw higher rates of false positives for African American females »[397]

CDPDJ, 10 avril 2019, en ligne : <https://www.cdpdjqcc.mywhc.ca/fr/publications/entente-de-collaboration-entre> (consulté le 25 août 2021).

[396] NIST, préc., note 17.

[397] *Id.*

L'existence de ce type de discrimination pose donc la question de savoir si une répartition stricte entre deux organes d'enquête ne risque pas d'exclure des situations graves d'atteintes aux droits fondamentaux. En l'occurrence, face à une IA de recrutement, les biais apparaissent multiples et peuvent influencer simultanément les décisions algorithmiques de tri de candidatures.

En outre, une autre difficulté pourrait émerger d'une entente de répartition de compétence ; la révélation du biais discriminatoire de l'IA passe par une recherche technique des biais. Or, cette recherche implique l'analyse du traitement des renseignements personnels, voire leur mode de collecte. Rappelons que dans le cas d'IA de recrutement comme celle de *HireVue*, c'est dès la collecte des données personnelles des candidats que les candidats sont aux prises avec des atteintes graves à leurs droits fondamentaux. L'atteinte au droit à l'égalité d'accès à l'emploi ne représenterait alors que la conséquence d'un traitement abusif et fallacieux des données collectées. Dans de telles circonstances, l'enquête relative à la discrimination algorithmique à l'embauche reposerait sur celle des atteintes à la vie privée. Aussi, en raison de cette interdépendance entre droit au respect de la vie privée et droit à l'égalité d'une part, et des circonstances particulières de la discrimination par IA de recrutement, d'autre part, une répartition de compétences entre les différentes autorités risque de complexifier l'accès à la justice, au lieu de le faciliter.

Une deuxième forme d'arrimage, plus inclusive, consisterait en une collaboration étroite entre les deux organes mentionnés. Une formation mixte et ponctuelle pourrait se réunir autour d'une enquête conjointe de la CAI et de la Commission des droits de la personne. À la suite de cette enquête, la formation mixte déciderait de l'opportunité des poursuites ou des sanctions. Une telle articulation des autorités entre elles éviterait d'abord une multiplication des procédures judiciaires par une mutualisation des moyens et du temps d'enquête. Ensuite, pour une personne candidate, cette collaboration institutionnelle pourrait signifier un signalement facilité : le signalement (ou la plainte) auprès d'une des deux commissions déclencherait une la création de la formation mixte ponctuelle. L'enquête déterminerait ensuite si la discrimination résulte de biais de l'IA, ou d'une intervention humaine attentatoire au droit à l'égalité. L'arbitrage du législateur serait nécessaire afin de déterminer le degré de collaboration entre ces deux institutions.

Enfin, plutôt que de réunir ponctuellement une formation mixte, l'idée d'une autorité permanente distincte des deux commissions concurrentes, pourrait également constituer une clarification du parcours judiciaire de la contestation d'une décision algorithmique de recrutement. Certains membres des autorités concurrentes pourraient développer une compétence spécifique en IA de recrutement afin d'apporter leur expertise à cette nouvelle entité. Compétente à la fois pour l'enquête, la promotion et la sanction des processus de recrutement augmentés, cet instrument du régulateur se consacrerait à la particularité du contentieux de l'IA de recrutement.

Toutefois, nonobstant les avantages d'un organe dédié à ce contentieux spécifique, le recrutement ne constitue pas le seul secteur d'application de l'IA. La multiplication d'institution spécialisée par contentieux impliquant l'IA pourrait également nuire à la sécurité juridique. L'impact social des technologies de l'information commanderait assurément une structure institutionnelle renouvelée dans laquelle les candidats lésés par une IA de recrutement exerceraient leur droit à la contestation des décisions algorithmiques. Le Canada aurait donc intérêt à développer la stratégie par laquelle le pays entend orienter et réguler l'écosystème de l'IA ; les discriminations résultant des biais de cette technologie ne se limitant pas au recrutement, l'accès à la justice ne saurait représenter un défi dans ce seul cadre. Or, les candidats seront peut-être confrontés à des enjeux similaires dans d'autres circonstances que celles de l'emploi.

b) La lisibilité du cadre juridique des IA : l'impératif d'une stratégie québécoise clairement définie

L'IA de recrutement, bien que soulevant des problématiques spécifiques, constitue avant tout une IA. Indépendamment de l'organisation du contentieux que le législateur déterminera, l'émergence d'un contentieux de l'IA de recrutement générera de la jurisprudence. Pour autant, les autres secteurs d'application des IA feront potentiellement naitre d'autres formes de contentieux. Dès lors, la question d'une cohérence juridique et jurisprudentielle des normes applicables aux IA se posera rapidement.

> « Les gouvernements ont adopté une approche réactive en matière de réglementation et de gouvernance, d'abord en laissant le marché se développer librement, puis en appliquant des mesures progressives face aux menaces et problèmes émergents. Si cette approche fonctionne dans les secteurs évoluant de manière plus graduelle, elle est inadaptée aux enjeux de la gouvernance des données et de l'IA, où de nouveaux modèles d'affaires peuvent prendre de l'expansion très rapidement et rejoindre en très peu de temps des millions d'utilisateurs dans le monde entier. »[398]

L'enjeu de l'harmonisation des normes de régulation de l'IA de recrutement par rapport aux autres applications de l'IA se pose donc, et ce, de manière relativement pressante. La vision du législateur fédéral en matière de régulation de l'IA demeure encore incertaine. Notamment, certaines disparités apparaissent déjà entre les législateurs fédéral et québécois en matière de gestion des données personnelles. À ce propos, le comparatif des projets de loi C-11 et la L*oi n° 64*[399], proposé par Simon Du Perron met en exergue cinq dissemblances. Les gouvernements fédéral et québécois adoptent en effet des orientations de régulation divergentes sur la place de l'EFVP, la notion de dépersonnalisation, les exceptions au consentement, la portabilité des données et les décisions automatisées[400]. Dans le cas de l'IA de recrutement, et de l'IA en général, quatre points paraissent pourtant cruciaux pour une régulation cohérente protectrice des droits fondamentaux des individus[401].

En premier lieu, l'EFVP constituerait la clé de voute de la prévention des discriminations issues des biais algorithmiques. Nous l'avons vu précédemment, cette évaluation effectuée par l'autorité protectrice de la vie privée, constitue un contrôle de conformité aux exigences de sûreté d'une part, et de compatibilité aux droits fondamentaux, d'autre part. Ce processus préventif offre, de surcroît, l'opportunité d'appliquer des méthodes de détection de biais telles que l'équité

---

[398] ELEMENT AI et NESTA, *Fiducies de Données: Un nouvel outil pour la gouvernance des données*, *Livre Blanc*, Atelier international sur les fiducies de données, décembre 2018, p. 10, en ligne : <https://hello.elementai.com/rs/024-OAQ-547/images/Fiducies_de_Donnees_FR_201914.pdf> (consulté le 26 novembre 2020).

[399] La loi est aujourd'hui adoptée, aussi faut-il préciser que le comparatif de Simon Du Perron porte bien sur la version du texte avant son adoption finale.

[400] S. DU PERRON, préc., note 284. « […] le projet de loi C-11 ne fait aucune mention de l'EFVP ce qui s'avère pour le moins étonnant considérant que son adoption était recommandée par le Commissariat à la protection de la vie privée du Canada (CPVP) et que ce mécanisme est déjà implanté dans le secteur public fédéral. ». En effet, le secteur public fédéral se repose sur le mécanisme de l'évaluation d'impact algorithmique.

[401] La portabilité des données touchant à des enjeux de gestion des données personnelles moins directs avec l'IA de recrutement, le propos se concentrera sur les quatre autres points de comparaison relevés par S. Du Perron.

contrefactuelle, le *testing* ou l'explication locale. Or, si le législateur québécois réaffirme l'importance de ce procédé dans la *Loi n° 64*, son homologue fédéral l'ignore[402]. Une telle disparité à propos de l'EFVP étonne, certes, mais implique également deux approches dont les répercussions sur les droits fondamentaux divergent. Dans l'optique fédérale, l'entreprise a la faculté de demander la certification de ses programmes de gestion du risque de ses IA ; tandis que dans l'approche québécoise le contrôle de la stratégie de gestion du risque constitue une exigence. Dans un cas, la bonne foi du secteur privé semble privilégiée, dans l'autre, la régulation québécoise n'entend pas se fier à la seule bonne foi des acteurs privés de l'IA. Bien que nous partagions cette dernière approche, l'existence de deux approches aussi contradictoires risque de nuire à la lisibilité de la régulation des IA.

En second lieu, S. Du Perron relève des définitions distinctes de la notion de dépersonnalisation. Le gouvernement québécois retient en effet le critère de l'irréversibilité de l'identification pour déterminer ce que constitue un renseignement dépersonnalisé[403]. En ce sens, l'acception québécoise de la dépersonnalisation désigne le processus par lequel un renseignement « ne permet plus d'identifier directement la personne concernée »[404], par opposition à l'anonymisation qui désigne une dépersonnalisation irréversible du renseignement[405]. En revanche, la dépersonnalisation au sens du législateur fédéral repose sur deux critères. Premièrement, la dépersonnalisation implique de ne plus identifier directement ou indirectement la personne concernée[406]. Cette dernière acception est d'ailleurs similaire à celle retenue par le Commissaire australien à l'information et à la vie privée.

---

[402] S. DU PERRON, préc., note 284, part. Responsabilité des organisations.

[403] *Loi n°64*, préc., note 20, art. 19. « Pour l'application de la présente loi, un renseignement personnel est dépersonnalisé lorsque ce renseignement ne permet plus d'identifier directement la personne concernée. »

[404] *Id.*, art. 12 al.4. « Pour l'application de la présente loi, un renseignement personnel est : 1° dépersonnalisé lorsque ce renseignement ne permet plus d'identifier directement la personne concernée […]».

[405] *Id.*, art. 23 al.2. « Pour l'application de la présente loi, un renseignement concernant une personne physique est anonymisé lorsqu'il ne permet plus, de façon irréversible, d'identifier directement ou indirectement cette personne ».

[406] *Projet de loi C-11*, préc., note 20, art. 20. Dépersonnaliser. « Modifier des renseignements personnels — ou créer des renseignements à partir de renseignements personnels — au moyen de procédés techniques afin que ces

> « De-identification involves two steps. The first is the removal of direct identifiers. The second is taking one or both of the following additional steps:
>
> the removal or alteration of other information that could potentially be used to re-identify an individual, and/or the use of controls and safeguards in the data access environment to prevent re-identification ».[407]

Deuxièmement, supposant l'impossibilité de garantir une dépersonnalisation irréversible, la distinction avec l'anonymisation n'apparaît pas dans le projet de loi C-11. On pourrait ainsi résumer cette différence des définitions québécoise et fédérale à l'intensité de l'obligation de dépersonnalisation choisie : au Québec, il s'agirait d'une obligation de résultat atténuée par la limitation de l'exigence au caractère direct de la dépersonnalisation, lorsque le fédéral définit une obligation de moyens renforcée en exigeant une dépersonnalisation directe, et indirecte.

Troisièmement, cette variation dans l'intensité des exigences se retrouve également dans les dispositions relatives aux décisions automatiques et aux exceptions du consentement. Compte tenu de l'interdépendance de ces deux éléments, nous les traiterons conjointement.

Tout d'abord, il faut relever que le projet de loi C-11 ouvre la porte à des abus bien plus importants que l'intérêt commercial légitime recommandé par le Commissariat à la vie privée. En effet, le législateur fédéral propose cinq exceptions : les activités d'affaires, le transfert des renseignements personnels aux fournisseurs de services, la soumission des renseignements à un processus de dépersonnalisation, les fins de recherches ou de statistiques, et… à « une fin socialement bénéfique ». Remarquons ici que l'ensemble de ces exceptions ouvre la voie à une transmission des renseignements personnels dépersonnalisés ou non, et ce, à l'insu même des individus concernés. La dernière exception relative aux fins socialement bénéfiques constitue en outre une sorte de concept valise autorisant le gouvernement fédéral à faire du secteur privé son « proxy » pour accéder librement aux données personnelles de sa population.

---

renseignements ne permettent pas d'identifier un individu ni ne puissent, dans des circonstances raisonnablement prévisibles, être utilisés, seuls ou en combinaison avec d'autres renseignements, pour identifier un individu. »

[407] OFFICE OF THE AUSTRALIAN INFORMATION COMMISSIONER, « De-identification and the Privacy Act » (mars 2018), part. What does a de-identification process involve?, en ligne : <https://www.oaic.gov.au/privacy/guidance-and-advice/de-identification-and-the-privacy-act/> (consulté le 20 juin 2020).

> « Pour l'application du présent article, fin socialement bénéfique s'entend de toute fin relative à la santé, à la fourniture ou à l'amélioration des services et infrastructures publics, à la protection de l'environnement ou de toute autre fin réglementaire. »[408]

À l'aune de telles exceptions, définies de manière aussi extensible, la loi présentée comme un moyen de protéger « la vie privée des consommateurs » parait pour le moins ironique. En comparaison le législateur québécois, a non seulement écarté la recommandation relative à l'intérêt commercial légitime, mais les trois exceptions au consentement se limitent « [aux] fins compatibles avec celles pour lesquelles il a été recueilli, [à l'utilisation] manifestement au bénéfice de la personne concernée et [aux] fins d'étude, de recherche ou de production de statistiques [sous réserve de dépersonnalisation] »[409]. Ici, au-delà des orientations relativement opposées entre les deux législateurs précités, il convient de remarquer que le Québec propose une régulation qui, si elle n'abandonne pas le principe du consentement, évite à tout le moins de multiplier des exceptions le vidant de sa substance. Son homologue fédéral opte en revanche pour une direction différente, poussant plus loin l'incohérence des recommandations du Commissariat à la vie privée.

> « […] force est de constater que l'approche poursuivie par Québec vise davantage la protection des renseignements personnels tandis que celle d'Ottawa cherche plutôt à favoriser leur circulation. »[410]

Cette dualité entre les régulations émergentes au Québec et au fédéral se manifeste enfin à travers les perspectives relatives aux décisions automatisées.

> « Alors que le projet de loi n° 64 ne vise que les décisions fondées exclusivement sur un traitement automatisé de renseignements personnels, le projet de loi C-11 s'applique à toute organisation qui utilise un système décisionnel automatisé pour faire une prédiction, formuler une recommandation ou prendre une décision concernant l'individu (art. 2 LPVPC). »[411]

L'approche du projet de loi C-11 recouvre le traitement intégralement automatisé autant que les systèmes d'aide à la décision, la L*oi n° 64* semble circonscrite au traitement entièrement

---

[408] *Projet de loi C-11*, préc., note 20, art. 39.

[409] *Loi n°64*, préc., note 20, art. 12 al.2.

[410] S. DU PERRON, préc., note 284, part. Décisions automatisées.

[411] *Id.*

automatisé. Afin d'assurer une plus grande cohérence avec les dispositions qui se veulent protectrices, le législateur québécois pourrait étendre le champ d'application des dispositions en question aux systèmes d'aide à la décision, ou encore, il pourrait définir des dispositions spécifiques à ces systèmes dont l'utilisation implique des humains. Plus encore, le législateur serait fondé à exiger un contrôle humain sur les IA utilisées, de sorte qu'en toutes circonstances, l'utilisation d'une IA impliquerait l'imputabilité des décisions algorithmiques à un être humain intégré au processus décisionnel appuyé par la technologie.

Enfin, contrairement aux recommandations du Commissariat à la vie privée, tout comme celles du Laboratoire de cyberjustice et de l'OBVIA[412], aucun des législateurs précités ne définit un droit d'opposition clair au traitement automatisé. Nous l'avons déjà relevé, en matière d'IA de recrutement, une telle lacune entérinerait la vulnérabilité des candidats face aux employeurs, et ce, même si les fiducies de données ou l'intervention institutionnelle entraient dans l'arsenal législatif destiné à la protection des droits fondamentaux. Dès lors, les perspectives de régulation actuelles s'enfoncent dans la dissonance cognitive qu'implique un attachement au principe du consentement, tout en élaborant des dispositions nouvelles le vidant entièrement de sa substance. Pour autant le RGPD, source d'inspiration pour les régulateurs occidentaux, prévoit en son article 22 un droit d'opposition explicite[413].

À l'aune de ces comparaisons, l'harmonisation des stratégies de régulation québécoise et fédérale constitue, pour le moment, un échec. Or, au-delà de l'IA de recrutement, l'élaboration d'une stratégie de régulation claire et harmonisée entre les législateurs fédéral et québécois

---

[412] Pierre-Luc DÉZIEL, Eve GAUMOND et Karim BENYEKHLEF, *Repenser la protection des renseignements personnels à la lumière des défis soulevés par l'IA*, avril 2020, p. 24, en ligne : <https://ajcact.openum.ca/publications/repenser-la-protection-des-renseignements-personnels-a-la-lumiere-des-defis-souleves-par-lia/> (consulté le 18 août 2021). « Enfin, il importe de rappeler que le principe de consentement est corollaire au droit d'opposition. L'obtention du consentement est dénuée de valeur si l'individu faisant l'objet d'un traitement de renseignements personnels n'a pas de possibilité réelle de refuser. Ainsi, à l'exclusion des cas d'exceptions au consentement, un individu devrait toujours être en mesure de s'opposer à un traitement de renseignements personnels ».

[413] *RGPD*, préc., note 88, art. 22 al.1. « Article 22 Décision individuelle automatisée, y compris le profilage

1.La personne concernée a le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire. »

risque de déterminer la solidité de la résilience du cadre juridique des IA. Y aurait-il une hiérarchie des valeurs démocratiques variable en fonction du secteur d'activité concerné ? Des normes générales communes à toutes les IA seront-elles ensuite complétées par des normes sectorielles ? Quelle forme prendra l'écosystème institutionnel chargé de l'application de cette stratégie de régulation ? À ce propos, l'UE ici également trace une première voie. L'Europe s'appuie en effet sur une distinction entre des IA à haut risque et des IA à risque modéré, afin d'établir une gradation des exigences et protections requises pour les personnes citoyennes[414]. La distinction reposerait alors sur deux critères cumulatifs :

> « premièrement, l'application d'IA est employée dans un secteur où, compte tenu des caractéristiques des activités normalement menées, des risques importants sont à prévoir […]
> deuxièmement, l'application d'IA dans le secteur en question est, de surcroît, utilisée de façon telle que des risques importants sont susceptibles d'apparaître. Ce second critère prend en considération le fait que toutes les utilisations de l'IA dans les secteurs sélectionnés n'impliquent pas nécessairement des risques importants. […]»[415]

L'UE envisage, en sus de ces deux critères, une exception par laquelle certaines IA seraient considérées à haut risque, par nature. L'IA de recrutement entrerait dans cette catégorie exceptionnelle.[416]

---

[414] préc., note 148, p. 20 à 24.

« En principe, le nouveau cadre réglementaire pour l'IA devrait atteindre ses objectifs avec efficacité sans être excessivement normatif, au risque de créer une charge disproportionnée, en particulier pour les PME. Pour réaliser cet équilibre, la Commission estime devoir suivre une approche fondée sur les risques. […] deuxièmement, l'application d'IA dans le secteur en question est, de surcroît, utilisée de façon telle que des risques importants sont susceptibles d'apparaître. Ce second critère prend en considération le fait que toutes les utilisations de l'IA dans les secteurs sélectionnés n'impliquent pas nécessairement des risques importants […]. »

[415] *Id.*, p. 20.

[416] *Id.*, p. 21. « Nonobstant les considérations qui précèdent, il peut également exister des cas exceptionnels dans lesquels, compte tenu des risques, l'utilisation d'applications d'IA à certaines fins devrait être considérée comme étant à haut risque en soi, c'est-à-dire indépendamment du secteur concerné, et resterait soumise aux exigences [relatives aux IA à haut risque]. […] compte tenu de son importance pour les particuliers et de l'acquis de l'UE sur l'égalité en matière d'emploi, l'utilisation d'applications d'IA dans les procédures de recrutement et dans des situations ayant une incidence sur les droits des travailleurs serait toujours considérée comme étant 'à haut risque' ».

D'autre part, parallèlement à cette proportionnalité des exigences au risque, la régulation de l'IA s'exercerait par le biais d'un écosystème d'autorités régulatrices sectorielles[417] dont la tête de pont serait une structure de gouvernance européenne[418].

> « La structure de gouvernance relative à l'IA et les éventuelles évaluations de la conformité dont il est question ici ne modifieraient en rien les compétences et les responsabilités des autorités compétentes, visées par le droit de l'UE en vigueur, que ce soit dans des secteurs spécifiques ou concernant des questions bien précises (finance, médicaments, aviation, dispositifs médicaux, protection des consommateurs, protection des données, etc.). »[419]

Il faut conclure ici que l'UE dispose déjà d'une stratégie de régulation reposant sur deux volets : l'un constituant une déclinaison juridique des mécanismes de gestion du risque, l'autre reposant sur les autorités régulatrices telles que l'Autorité européenne des marchés financiers, l'Agence européenne du médicament, ou encore le Comité européen de la protection des données.

Au Canada, il serait grand temps que le législateur prenne le leadership sur ces questions. En effet, si l'UE identifie déjà l'IA de recrutement dans la catégorie des IA à haut risque, le Canada ne dispose pas encore d'une classification des IA en fonction des risques auxquels elles exposent les individus.

---

[417] *Id.*, p. 28-29.

[418] *Id.*, p. 28. « Il est nécessaire de mettre en place une structure de gouvernance européenne sur l'IA, qui prendrait la forme d'un cadre pour la coopération des autorités nationales compétentes, pour éviter une fragmentation des responsabilités, renforcer les capacités dans les États membres et faire en sorte que l'Europe se dote progressivement des capacités nécessaires pour tester et certifier les produits et services reposant sur l'IA ».

[419] *Id.*, p. 29.

CONCLUSION

À l'heure où les démocraties occidentales traversent une crise protéiforme, le témoin des mutations sociales actuelles pourrait légitimement s'interroger sur le besoin d'une révision profonde des fondements du droit. Est-il toujours pertinent face aux transformations et « disruptions » que provoquent l'IA ? Ne faudrait-il pas tout revoir, tout réévaluer ?

Tout ? Peut-être pas. Si nous trouvons essentiel de soulever ces interrogations, les développements exposés au cours de ce travail nous conduisent à une réponse nuancée. En effet, bien que l'obsolescence de certains paradigmes et législations guette la société canadienne, la problématique de la gestion du risque de discrimination du recrutement augmenté rassure sur les fondements essentiels de notre système juridique.

En premier lieu, malgré un rythme d'évolution soutenu de l'IA et des technologies de l'information, les droits fondamentaux protégés par les chartes demeurent applicables et effectifs en présence d'IA de recrutement. Les textes constitutionnels offrent une protection résiliente qui transcende les mutations technologiques que nous vivons. Les bornes des droits fondamentaux ainsi que leur définition repose sur des perspectives dynamiques des rapports sociaux. La capacité de notre droit à absorber les défis de l'IA de recrutement, et de l'IA en général, ne repose donc pas tant sur nos valeurs sociales (incarnées par les Chartes), mais sur la volonté à se saisir de ces enjeux de la part de nos législateurs, aussi bien fédéral que québécois.

En effet, en réponse à l'inexistence d'un cadre juridique dédié à l'IA, les législations de protection de la vie privée et des renseignements personnels servent de référentiel. Bien qu'elles ne soient plus adaptées aux réalités technologiques d'aujourd'hui, les étendre à l'IA offre un premier socle d'encadrement en agissant à sa source : le *big data*. Force est de constater que ces dispositions législatives régulent la collecte, le traitement des données, et responsabilisent les gestionnaires vis-à-vis des droits fondamentaux des personnes exposées à la décision algorithmique. La seule extension de ces normes à l'IA de recrutement évite donc un vide juridique. Dans le recrutement, la régulation des technologies de profilage et d'analyses comportementales constitue des améliorations déterminantes en cas de discriminations algorithmiques à l'embauche. En d'autres

termes, les lois de protection des données personnelles et de la vie privée permettent d'ores et déjà de penser la gestion de risque de discrimination algorithmique à l'embauche.

En second lieu, aussi jeune et carencé que soit l'encadrement juridique de l'IA de recrutement, le droit des obligations offre également des mécanismes transférables à la gestion du risque de discrimination dans les processus de recrutement augmentés.

D'abord, il faut rappeler que la gestion du risque informationnel requiert l'identification des vulnérabilités d'un système. Par analogie, la base de données des IA de recrutement constitue donc la principale vulnérabilité. En effet, c'est sur la base des informations agglomérées dans la base de données que l'IA de recrutement s'entraînera, avant d'analyser des candidatures dans des situations réelles.

Nous l'avons vu au chapitre 2 à travers le cas d'Amazon : les décisions discriminatoires d'une IA de recrutement proviennent essentiellement de biais appris au cours de l'entraînement de l'algorithme. Aussi, la mitigation de ce risque de transmission de biais nécessite l'application d'un principe de précaution tout au long de la constitution base de données de l'IA de recrutement. À l'instar de l'obligation de sécurité informationnelle, l'obligation de moyens renforcée peut donc devenir le véhicule de la prévention des biais algorithmiques des IA de recrutement.

Obligation centrale aussi bien en gestion des données personnelles qu'en sécurité informationnelle, l'obligation de moyens renforcée permet en effet un rééquilibre de la relation de pouvoir entre les candidats et les employeurs, en inversant le fardeau de la preuve au bénéfice des candidats. De même que la diligence attendue des gestionnaires de données sensibles, l'obligation de moyens renforcée dans un contexte d'IA de recrutement forcerait une proactivité des employeurs — autant que des programmeurs d'IA de recrutement — afin d'éviter, autant que faire se peut, l'exposition des postulants à des discriminations à l'embauche. Ce type d'obligation à la charge des employeurs et des concepteurs d'IA assurerait ainsi une mitigation du risque de discrimination, sans que cela nécessite la création de nouveaux mécanismes juridiques. En revisitant certains mécanismes classiques du droit, il est donc possible d'y trouver des réponses

aux défis des technologies de l'information. L'existant permet ainsi de se repérer, tout en constituant un point de départ quant à la compréhension des enjeux de l'IA de recrutement, et de la gestion de risque subséquente.

Cependant, et sans préjudice pour ces constats précédents rassurants, les réflexions ne pourraient tendre vers la complétude sans s'attarder sur la question du contentieux de la discrimination algorithmique, et plus spécifiquement, sur l'accès difficile à la justice pour des candidats discriminés par IA.

Alors que la nécessité d'une « explication valable » des décisions algorithmiques fait consensus auprès des parties prenantes de l'IA, la transparence nécessaire à la détection des biais discriminatoires soulève des difficultés que les connaissances techniques ne peuvent pallier à elles seules.

Qu'il s'agisse de la méthode de l'explication locale ou de l'équité contrefactuelle, ces réponses techniques tendent à tracer la voie vers la détection des biais affectant une décision algorithmique de recrutement. Toutefois, en dépit de leurs résultats encourageants, les imperfections de ces méthodes aussi bien que leur technicité entravent leur accès pour des candidats discriminés.

D'autres méthodes visent donc à contourner le besoin d'expliquer une décision X donnée : les garanties théoriques et la preuve statistique, quant à elles, évaluent la fiabilité globale d'une IA. Elles permettraient par exemple de tester les IA de recrutement de la même façon que le testing de CV s'effectue auprès de recruteurs humains. Bien que ces deux autres méthodes facilitent l'accès à la preuve de biais discriminatoires affectant une IA de recrutement, ici également les candidats se heurtent à des obstacles pratiques. D'une part, la preuve statistique permet l'évaluation de la fiabilité de l'IA de recrutement, mais ne concèderait pas la preuve qu'une décision portant sur une personne X est viciée. D'autre part, les garanties théoriques semblent applicables uniquement dans un contexte préservé du monde réel.

De plus, comment une personne candidate pourrait-elle se procurer de telles preuves sans bénéficier d'une expertise particulière en ce domaine ? Plus encore, comment un individu pourrait-il rechercher des preuves de discrimination à l'embauche sans savoir préalablement qu'une IA traitera sa candidature en tout ou partie ? L'information des candidats conditionne donc drastiquement l'accès à la justice des candidats lésés. Les méthodes d'accès à la preuve de biais discriminatoire dans un processus de recrutement augmenté ne peuvent donc produire leurs effets sans un cadre légal adapté.

À ce titre, les obligations d'information et de sûreté appuient et garantissent un droit effectif au recours pour les personnes candidates soumises à une discrimination par IA. Si ces obligations ne présentent pas de caractère innovant en soi, il convient de remarquer qu'elles s'intègrent à un paradigme de protection des droits fondamentaux qui reposerait sur le seul consentement des individus. Ces derniers contrôleraient leurs données, consentiraient au traitement algorithmique de leurs candidatures.

Or, l'angle mort principal de ce paradigme repose sur le constat que même en cas d'information adéquate des candidats, même si l'on procède au recueil de leur consentement, le consentement n'aurait pas de valeur. Le contexte d'embauche génère, en effet, un déséquilibre contractuel entre le candidat, vulnérable, et l'employeur en position de force.

En outre, au-delà de la vulnérabilité des candidats, l'inégal accès à l'information quant au fonctionnement et à l'implication d'un traitement algorithmique des données d'un individu, ne saurait fonder une protection dont le citoyen néophyte conserverait le contrôle. Ce contrôle théorique résiste difficilement aux situations concrètes.

En effet, qu'il s'agisse d'IA de recrutement ou d'IA appliquées à d'autres activités, il serait judicieux de constater sans équivoque la désuétude du consentement. Ce mécanisme, dans un contexte de protection des droits fondamentaux, signifie que les individus se retrouvent constamment dans un rapport à l'information qui leur est aussi défavorable qu'un consommateur face aux vendeurs professionnels. L'entre-deux dans lequel se sont engouffrés les législateurs québécois et fédéral entrave un changement de paradigme. L'abandon du consentement comme pilier de la protection des droits fondamentaux permettrait ainsi l'arrêt d'une érosion croissante

desdits droits fondamentaux. Nous l'avons vu, il devient crucial de rappeler la prévalence des droits fondamentaux sur les intérêts commerciaux dont la légitimité serait discutable. Or, les législateurs fédéral et québécois se dirigent vers une sorte de dénaturation croissante du consentement, et ce, notamment par la reconnaissance d'exceptions au consentement aussi floue et dangereuse que l'intérêt commercial légitime.

Assumer un changement de paradigme permettrait l'engagement de réflexions approfondies sur d'autres mécanismes plus adaptés à la réalité des IA. À ce titre, l'idée d'une fiducie des données représente une alternative au paradigme du consentement[420]. Reposant sur des tiers experts ou institutionnels, la fiducie de données conduirait à décharger les candidats lésés des vérifications et recours nécessaires à la protection de leurs données personnelles, y compris en contexte de recrutement. La délégation de la protection des données à des tiers experts assurerait ainsi un rééquilibrage entre le candidat et l'employeur. Les candidats échapperaient à un consentement contraint et susceptible de les exposer à des discriminations algorithmiques à l'embauche.

De même, l'EFVP (évaluation des facteurs de la vie privée), mise au service d'un régime d'autorisation des IA de recrutement, créerait un mécanisme d'évaluation et donc de neutralisation des biais discriminatoires des IA de recrutement, avant leur déploiement. Aussi, l'exploration de ces mécanismes de protection du droit à l'égalité ne sera possible qu'à la condition d'un changement de paradigme de nos législations. Certains canons des législations fédérales et québécoises commandent donc des transformations normatives plus audacieuses, a fortiori dans des secteurs aussi inégalitaires que le milieu du recrutement.

---

[420] A.-S. HULIN, préc., note 342; George ZARKADAKIS, « "Data Trusts" Could Be the Key to Better AI », *Harvard Business Review* (10 novembre 2020), en ligne : <https://hbr.org/2020/11/data-trusts-could-be-the-key-to-better-ai> (consulté le 26 novembre 2020); « What is a data trust? – The ODI », en ligne : <https://theodi.org/article/what-is-a-data-trust/> (consulté le 26 novembre 2020); Pierre TRUDEL et Anne-Sophie HULIN, *Web-discussion | La fiducie de données : le bon véhicule juridique pour encadrer le développement de l'intelligence artificielle?*, 24 novembre 2020, en ligne : <https://www.cyberjustice.ca/2020/10/29/web-discussion-la-fiducie-de-donnees%e2%80%af-le-bon-vehicule-juridique-pour-encadrer-le-developpement-de-lintelligence-artificielle/> (consulté le 27 novembre 2020). « La fiducie est un patrimoine d'affectation, soit une universalité de droits et d'obligations affectée à un but déterminé (art. 1260 C.c.Q). Elle peut être constituée dans le but de favoriser des personnes désignées. », la fiducie des données implique donc de considérer les données personnelles comme un patrimoine dont on peut déléguer la gestion et la protection.

Enfin, si les évolutions et réflexions amorcées reflètent une compréhension progressive des impacts de l'IA sur la société[421], *a fortiori* pour une IA de recrutement, force est de constater le manque d'anticipation de situations conflictuelles que généreront les IA de recrutement. Alors que ces dernières pénètrent la pratique des services de ressources humaines de manière croissante, le rapprochement des perspectives fédérale et québécoise de régulation de l'IA reste pour le moment un échec. L'examen conjoint de la réforme québécoise et du projet de réforme fédéral illustre des incohérences et un manque d'harmonisation sur la régulation de l'IA et le marché de la donnée numérique[422].

En comparaison, le livre blanc de l'UE sur la réglementation de l'IA dessine d'ores et déjà la structure juridique et organique de la régulation de l'IA en Europe. D'une part, l'Europe s'appuie sur une distinction des IA à haut risque et des IA à risque modéré afin d'établir une gradation des exigences et protections requises pour les individus[423]. D'autre part, la régulation de l'IA s'exercera par le biais d'un écosystème d'autorités régulatrices sectorielles[424], coordonné par une structure de gouvernance européenne. Au Québec et au niveau fédéral, cependant, si l'étirement des lois et mécanismes existants offre un gain de temps maigre, il serait regrettable que le contentieux de la discrimination algorithmique surprenne les législateurs. Par exemple, les justiciables exposés aux IA de recrutement doivent connaître leur interlocuteur : est-ce l'autorité de protection de la vie privée, l'autorité relative aux normes du travail, ou celle dédiée à la protection des droits de la personne ? Compte tenu du modèle juridictionnel en place, le risque de concurrence entre ces organes et les tribunaux de droit commun pourrait conduire à des doublons d'enquêtes, voire à des dénis de justice. Il serait intéressant de penser un arrimage entre ces différentes autorités essentielles à la protection des droits fondamentaux des individus. À cet égard, le modèle envisagé par l'Europe pourrait inspirer le Québec ou le gouvernement fédéral, sans toutefois restreindre la créativité juridique des législateurs.

---

[421] S. DU PERRON, préc., note 284.

[422] *Id.*

[423] préc., note 148, p. 20 à 24.

[424] *Id.*, p. 28-29.

Enfin, alors que l'UE classe l'IA de recrutement dans la catégorie des IA à haut risque, le Québec et le gouvernement fédéral ne disposent pas encore d'une grille de classement des IA en fonction des risques auxquels la technologie expose la population. Or, c'est sur le fondement d'une évaluation des risques que l'on pourra définir des niveaux d'exigences et de responsabilité des gestionnaires d'IA, ou ici, des gestionnaires du recrutement augmenté.

Pour conclure, l'enjeu d'une actualisation du cadre juridique des technologies de l'information ne se résume pas qu'à la création d'un droit innovant et résilient face au développement exponentiel de l'IA. Sous l'influence monopolistique des GAFAM, ces technologies progressent dans un contexte de contractualisation[425] croissante du droit. Par conséquent, la problématique de la régulation de l'IA oblige les questionnements sur nos appareils judiciaires et législatifs. Plus encore, cet enjeu rappelle qu'en période de grands bouleversements, le point de repère de la société demeure la loi.

La pacification des interactions sociales à laquelle aspire le droit requiert du législateur une réaffirmation de sa fonction : la recherche d'équilibre entre les intérêts particuliers et collectifs, l'expression de la *vox populi*. Le rôle du législateur ne se limite donc pas à la définition d'un cadre de régulation, face aux impacts de l'IA, il s'agit d'inscrire nos choix de société dans le droit.

De l'ensemble de ces réflexions demeure une question essentielle : notre liberté survivra-t-elle à une érosion constante de nos droits fondamentaux ? Les juristes peuvent déployer une créativité dont les fruits guideront un monde des possibles qui, plutôt que de la concurrencer, réaffirmera notre liberté. Jean-Paul Sartre rappelait d'ailleurs :

> « Nous sommes une liberté qui choisit, mais nous ne choisissons pas d'être libres : nous sommes condamnés à la liberté ».

[425] Alain SUPIOT, « Les deux visages de la contractualisation : déconstruction du Droit et renaissance féodale », dans Sandrine. CHASSAGNARD-PINET, David. HIEZ, et RÉSEAU EUROPÉEN DROIT ET SOCIÉTÉ. (dir.), *Approche critique de la contractualisation*, coll. Droit et société. Recherches et travaux ; 16, Paris, LGDJ, 2007.

TABLE DE LA LÉGISLATION

Textes constitutionnels

Charte Canadienne des Droits et Libertés, (1982) Loi constitutionnelle de 1982, partie I, c. 11

Loi Canadienne sur les droits de la personne, (1985), ch. H-6, L.R.C.

Loi constitutionnelle de 1867, 1867

Textes fédéraux

Loi sur l'équité en matière d'emploi, (1995), L.C. 1995, ch. 44

Loi sur la protection de la vie privée des consommateurs, projet de loi n° C-11, en ligne : https://parl.ca/DocumentViewer/fr/43-2/projet-loi/C-11/premiere-lecture#ID2E0IB0BA

Textes québécois

Charte des Droits et Libertés de la Personne, (1975), RLRQ c C-12

Code civil du Québec, L. Q. 1991, c. 64.

Loi concernant le cadre juridique des technologies de l'information, (2001) RLRQ c. C -1.1

Loi modernisant des dispositions législatives en matière de protection des renseignements personnels, RLRQ, 64, en ligne : http://www.assnat.qc.ca/fr/travaux-parlementaires/projets-loi/projet-loi-64-42-1.html

Loi sur la protection des renseignements personnels dans le secteur privé, (1993) RLRQ c P-39.1,

Loi sur l'équité salariale, (1996), RLRQ c E-12.001

Loi sur les commissions d'enquête, (1964), RLRQ c C-37

Textes internationaux

Convention internationale sur l'élimination de toutes les formes de discrimination raciale, (1966) Recueil des Traités, résolution 2106 (XX)

Convention sur l'élimination de toutes les formes de discrimination à l'égard des femmes, (1979) Recueil des Traités, résolution 34/180

Déclaration des Nations Unies sur les droits des peuples autochtones, (2007), Résolution 61/295.

Déclaration Universelle des Droits de l'Homme, (1948), résolution 217 A (III)

Pacte International Relatif aux Droits Civils et Politiques, (1966) Recueil des Traités

Règlements européens

Règlement Général sur la Protection des Données personnelles, 88 (2016), 2016/679

TABLE DE LA JURISPRUDENCE

Jurisprudence canadienne

Andrews c. Law Society of British Columbia, [1989] 1 RCS 143 (CSC)

Canada (Human Rights Commission) c. Canada (Department of National Health and Welfare), 1998 Cour fédérale

Canada (Procureur Général) c. Walden, 2010 Cour fédérale

Khiamal c. Canada (Commission des droits de la personne), 2009 Cour fédérale

Rapport d'enquête sur la sécurité, la collecte et la conservation des renseignements personnels TJX Companies Inc./Winners Merchant International L.P., 2007 Commissaire à la protection de la vie privée du Canada

Singh v. Minister of Employment and Immigration, [1985] 1 SCR 177 (SCC)

Sharon Turpin and Latif Siddiqui c. Her Majesty the Queen and the Attorney General of Canada et al. Case, [1989] 1 R.C.S 1296-1336 (CSC).

Syndicat des travailleurs de l'industrie de la fibre de Chambly Inc. et Bennett Fleet Inc., (T.A., 1990-04-30) SOQUIJ AZ-90141114, D.T.E. 90T-799, [1990] T.A. 470

University of Alberta c. Alberta (Human Rights Comm.), [1992] 2 R.C.S 1103-1198 (CSC).

Jurisprudence étatsunienne

T.J Hooper c. Nothern Barge, [1932] 60 F. 2d 737 (2 d Cir.)

BIBLIOGRAPHIE

Monographies et ouvrages collectifs

BEAUREGARD, J.-P., *Les frontières invisibles de l'embauche des Québécois minoritaires : hiérarchie ethnique, effet modérateur du genre féminin et discrimination systémique : dévoiler la barrière à l'emploi par un testing à Québec*, Thèse de doctorat, Université de Laval, 2020

BENSAMOUN, A. et G. LOISEAU, *Droit de l'intelligence artificielle*, coll. Les Intégrales, 2110-9680, 15, Issy-les-Moulineaux, LGDJ, 2019.

BENYEKHLEF, K., *Une possible histoire de la norme : les normativités émergentes de la mondialisation*, 2e édition, Montréal, Éditions Thémis, 2015.

CRÉPEAU, P.-A. et CENTRE DE RECHERCHE EN DROIT PRIVÉ ET COMPARÉ DU QUÉBEC, *L'intensité de l'obligation juridique ou des obligations de diligence, de résultat et de garantie*, Cowansville (Québec), Y. Blais, 1989.

JULIA, L., O. KHAYAT et J.-L. GASSÉE, *L'intelligence artificielle n'existe pas*, Paris, First éditions, 2019.

LAJOIE, A., *Pouvoir disciplinaire et tests de dépistage de drogues en milieu de travail : illégalité ou pluralisme*, coll. Collection Relations industrielles ; 27, Cowansville, Québec, Éditions Y. Blais, 1995.

O'NEIL, C., *Weapons of Math Destruction : How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016

Roussy, D., C. Bernier et C. Malone, Paul-André Crépeau, L'intensité de l'obligation juridique ou des obligations de diligence, de résultat et de garantie, Montréal, Les Éditions Yvon Blais Inc., 1989, 232 pages, ISBN 2-89073-726-8.

Articles de revue d'études d'ouvrages collectifs

AMBLARD, M., « Idée reçue : Les algorithmes prennent-ils des décisions ? », *Interstices Info* 2018,

BENGIO, Y., « La révolution de l'apprentissage profond - Interstices », 2019

BENSAMOUN, A. et G. LOISEAU, « La gestion des risques de l'intelligence artificielle . - De l'éthique à la responsabilité », (2017) 46-1203 *JCP G*.

BENYEKHLEF, K., « Droit global : un défi pour la démocratie », *Revue Projet : comprendre pour agir* À l'heure des multinationales, le retard du droit ? (24 juin 2016)

———, « L'IA et nos principes de justice fondamentale », *Policy Options* Ethical and Social Dimensions of AI (15 février 2018)

BERTAIL, P., D. BOUNIE, P. WAELBROECK et S. CLÉMENÇON, *Algorithmes : biais, discrimination et équité*, Télécom ParisTech & Fondation ABEONA, 2019

BESSON, S., « Les obligations positives de protection des droits fondamentaux », (2003) 1 *Revue de droit suisse* 49-96.

CASILLI, A., P. TUBARO, C. LE LUDEC, M. COVILLE, M. BESENVAL, T. MOUHTARE et E. WAHAL, *Le Micro-Travail en France*, DiPLab, 2019

CRDP. *Entre propriété et liberté : 30 ans de protection des données personnelles. Regards croisés Europe/Amérique*, Colloque, Montréal, 6 novembre 2018.

CREPEAU, P.-A., « Le Contenu Obligationnel d'un Contrat », (1965) 43-1 *Can. Bar. Rev.* 1-48

DOSHI-VELEZ, F., M. KORTZ, R. BUDISH, C. BAVITZ, S. J. GERSHMAN, D. O'BRIEN, S. SHIEBER, J. WALDO, D. WEINBERGER et A. WOOD, « Accountability of AI Under the Law: The Role of Explanation », *SSRN Electronic Journal* 2017, DOI : 10.2139/ssrn.3064761.

DU MARAIS, B., « Analyses et propositions pour une régulation de l'Internet », (2002) 7-2 *Lex Electronica*

DU PERRON, S., « Projet de loi 64 : une réforme à l'Européenne du droit à la protection des renseignements personnels », *Laboratoire de cyberjustice* (17 juin 2020)

———, « Projet de loi n° 64 : échos de commission parlementaire », *Laboratoire de cyberjustice* (15 octobre 2020)

———, « Le temps des réformes : cinq comparaisons entre le projet de loi n° 64 et le projet de loi C-11 », *Projet AJC | ACT Project* (24 novembre 2020)

EASTERBROOK, F. H., « Cyberspace and the Law of the Horse », *U Chi Legal F 207* 1996.

EID, P., « Les inégalités ethnoraciales » dans l'accès à l'emploi à Montréal : le poids de la discrimination », (2012) 53-2 *Rs* 415-450, DOI : 10,720 2/1012407ar.

GAUTRAIS, V. et C. PAPINEAU, « L'explosion documentaire : la normativité individuelle, nouveau centre de gravité normatif », dans *Lex electronica*, coll. Hors-série « Normes énormes », n°23, Montréal, CRDP, Université de Montréal, 2018, p. 143-172

HOUSER, K., *Can AI Solve the Diversity Problem in the Tech Industry? Mitigating Noise and Bias in Employment Decision-Making*, SSRN Scholarly Paper, ID 3344751, Rochester, NY, Social Science Research Network, 2019

KUSNER, M., J. LOFTUS, C. RUSSELL et R. SILVA, « Counterfactual Fairness », 2018, en ligne: https://arxiv.org/abs/1703.06856

LARRÈRE, C., « Montesquieu et le « doux commerce » : un paradigme du libéralisme », *Cahiers d'histoire. Revue d'histoire critique* 2014.123.21-38

LEDUC, A., « L'émergence d'une nouvelle lex mercatoria à l'enseigne des principes d'UNIDROIT relatifs aux contrats du commerce international : thèse et antithèse », 2001.23.

LESSIG, L., « The Law of the Horse: What Cyberlaw Might Teach », (1999) 113-501 *Harvard Law Review*

MOREL, A., « La Charte Quebecoise : Un Document Unique dans l'Histoire Legislative Canadienne », (1987) 21-1 *R.J.T. n.s.* 1-24

PAVAGEAU, S., « Les obligations positives dans les jurisprudences des cours européenne et interaméricaine des droits de l'homme », *International Law : Revista colombiana de derecho internacional* 2005.6.201-246.

POULLET, Y., « Technologies de l'information et de la communication et "co-régulation" : une nouvelle approche ? », dans *Liber amicorum Michel Coipel*, Bruxelles, Kluwer, 2004, p. 167

ROUSSY, D., C. BERNIER et C. MALONE, « Paul-André Crépeau, L'intensité de l'obligation juridique ou des obligations de diligence, de résultat et de garantie », (1991) 22-1 RGD 249-254

ROYER, J-C., « PAUL-ANDRÉ CRÉPEAU, L'intensité de l'obligation juridique ou Des obligations de diligence, de résultat et de garantie, Montréal, Édition Yvon Blais, 1989, 232 p., ISBN 2-89073-726-8. », (1991) 32-1 cd1 240-240

SNOW, J., « Google Photos Still Has a Problem with Gorillas », *MIT Technology Review* 2020

SUPIOT, A., « Les deux visages de la contractualisation : déconstruction du Droit et renaissance féodale », dans Sandrine. CHASSAGNARD-PINET, David. HIEZ, et RÉSEAU EUROPÉEN DROIT ET SOCIÉTÉ. (dir.), *Approche critique de la contractualisation*, coll. Droit et société. Recherches et travaux ; 16, Paris, LGDJ, 2007.

TODOLÍ-SIGNES, A., « Algorithms, artificial intelligence and automated decisions concerning workers and the risks of discrimination: the necessary collective governance of data protection », (2019) 25-4 *Transfer: European Review of Labour and Research* 465-481, DOI : 10.1177/1024258919876416.

TRUDEL, P., « Les Effets Juridiques de l'Autoreglementation », (1988) 19-2 *R.D.U.S.* 247-286.

VINING, A. R., D. C. MCPHILLIPS et A. E. BOARDMAN, « Use of Statistical Evidence in Employment Discrimination Litigation », (1986) 64-4 *Can. Bar. Rev.* 660-702

ZHENG, N., Z. LIU, P. REN, Y. MA, S. CHEN, S. YU, J. XUE, B. CHEN et F. WANG, « Hybrid-augmented intelligence: collaboration and cognition », (2017) 18-2 *Frontiers Inf Technol Electronic Eng* 153-179, DOI : 10,163 1/FITEE.1700053.

Documents institutionnels et gouvernementaux

BERNARD, C. et CDPDJ, *La position de la Commission des droits de la personne et des droits de la jeunesse du Québec face aux tests de dépistage de drogue en milieu de travail*, Montréal, CDPDJ, 2002.

BORGESIUS, F. Z., *Discrimination, intelligence artificielle et décisions algorithmiques*, Strasbourg, Conseil de l'Europe, 2018.

CAI. *Évaluation des facteurs relatifs à la vie privée : Savoir détecter et atténuer les risques d'atteinte aux renseignements personnels*, janvier 2018.

COMMISSARIAT DE LA PROTECTION DE A VIE PRIVÉE DU CANADA, « Nos attentes : Guide du Commissariat au sujet du processus d'évaluation des facteurs relatifs à la vie privée » (13 octobre 2011).

CDPDJ, « Discrimination : pratique interdite », *Pratiques interdites*.

———, « Origine et mission de la Commission des droits de la personne et des droits de la jeunesse », *La Commission*.

CNIL, « Profilage et décision entièrement automatisée ».

COMMISSARIAT À LA PROTECTION DE LA VIE PRIVÉE DU CANADA, *Consultation sur les propositions du Commissariat visant à assurer une réglementation adéquate de l'intelligence artificielle*, coll. Consultations, Commissariat à la protection de la vie privée du Canada, 2018.

———, « Pour une législation efficace sur la protection des renseignements personnels et l'accès à l'information dans une société guidée par les données » (6 novembre 2019).

———, *Un cadre réglementaire pour l'IA : recommandations pour la réforme de la LPRPDE*, Gatineau, Commissariat à la protection de la vie privée du Canada, 2020.

CNESST et CDPDJ. *Entente de collaboration entre la CDPDJ et la CNESST concernant leurs interventions en matière de harcèlement | Publications | CDPDJ*, 10 avril 2019.

COUTU, M. et P. BOSSET, *Etude 6 : La dynamique juridique de la Charte*, 6, coll. Études, vol.2, CDPDJ, 2003.

GROTHER, P., M. NGAN et K. HANAOKA, *Face recognition vendor test part 3 : demographic effects*, NIST IR 8280, Gaithersburg, MD, NIST, 2019, DOI : 10,602 8/NIST.IR.8280.

GROUPE DE TRAVAIL « ARTICLE 29 », *Lignes directrices relatives à la prise de décision individuelle automatisée et au profilage aux fins du règlement (UE) 2016/679*, 17/FR WP251rev.01, Bruxelles, Commission Européenne, 2018.

———, *Lignes directrices sur le consentement au sens du règlement 2016679.pdf*, 17/FR WP259 rév.01, Bruxelles, Commission Européenne, 2018.

GROUPE D'EXPERTS DE HAUT NIVEAU SUR L'INTELLIGENCE ARTIFICIELLE, *Lignes directrices en matière d'éthique pour une IA digne de confiance*, Commission Européenne, 2019.

MINISTÈRE DE LA JUSTICE, « Modernisation de la Loi sur la protection des renseignements personnels du Canada », *Site Internet du ministère de la Justice Canada* (4 juin 2020), en ligne : <https://www.justice.gc.ca/fra/sjc-csj/lprp-pa/modern.html#s2> (consulté le 18 juin 2020).

MULLER, C., *Intelligence artificielle - Une approche européenne axée sur l'excellence et la confiance*, Livre Blanc, INT/894-EESC-2020, 2020, en ligne : <https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/livre-blanc-sur-lintelligence-artificielle> (consulté le 23 juin 2020).

NIST, « NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software », *NIST* (19 décembre 2019), en ligne : <https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software> (consulté le 13 juillet 2020).

OFFICE OF THE AUSTRALIAN INFORMATION COMMISSIONER, « De-identification and the Privacy Act » (mars 2018).

SECRÉTARIAT DU CONSEIL DU TRÉSOR DU CANADA, *Directive sur la prise de décision automatisée*, 5 février 2019

SECRÉTARIAT DU CONSEIL DU TRÉSOR DU CANADA, « Évaluation de l'incidence algorithmique (EIA) », *aem* (31 mai 2019).

SECRÉTARIAT DU CONSEIL DU TRÉSOR DU CANADA*, L'Outil d'évaluation de l'impact algorithmique* (27 mai 2019).

THERRIEN, D., « Modernizing federal privacy laws to better protect Canadians », *Commissariat à la Protection de la Vie Privée du Canada* (1 juin 2020), en ligne : <https://priv.gc.ca/en/opc-news/speeches/2020/sp-d_20200309/> (consulté le 18 juin 2020).

TRUDEL, J.-F., *Mémoire à la Commission d'Accès à l'Information sur le document de consultation « Intelligence artificielle »*, Cat. 2.412.133, CDPDJ, 2020

Documents et publications non gouvernementales

ABRASSART, C., Y. BENGIO, G. CHICOISNE & al., *La Déclaration de Montréal pour un développement responsable de l'Intelligence Artificielle*, Montréal, en ligne : <https://www.declarationmontreal-iaresponsable.com/la-declaration>.

DÉZIEL, P.-L., E. GAUMOND et K. BENYEKHLEF. *Repenser la protection des renseignements personnels à la lumière des défis soulevés par l'IA*, avril 2020.

DIAMOND, S.-P. *Les fiducies de données - Un véhicule juridique pour rétablir l'équilibre entre la protection des données personnelles et l'essor de l'intelligence artificielle*, 28 octobre 2019.

ELEMENT AI et NESTA, « Fiducies de Données : Un nouvel outil pour la gouvernance des données », dans *Livre Blanc*, 2018, en ligne : <https://hello.elementai.com/rs/024-OAQ-547/images/Fiducies_de_Donnees_FR_201914.pdf> (consulté le 26 novembre 2020).

ISO/IEC 27000, 2018.

WHITTAKER, M., K. CRAWFORD, R. DOBBE, G. FRIED, E. KAZIUNAS, V. MATHUR, S. WEST MYERS, R. RICHARDSON, J. SCHULTZ et O. SCHWARTZ, *Discriminating Systems: Gender, Race, and Power in AI*, coll. Publications, New York, AI Now Institute, 2018, en ligne : <https://ainowinstitute.org/discriminatingsystems.html> (consulté le 29 novembre 2019).

« Counterfactual Fairness », *Microsoft Research*, en ligne : <https://www.microsoft.com/en-us/research/video/counterfactual-fairness/> (consulté le 3 novembre 2020).

Lois et codes annotés

ROY, A., B. MOORE, É. M. CHARPENTIER et S. LANCTÔT, *Code civil du Québec : annotations, commentaires*, (1991), L. Q.

Articles de journaux

AGENCE FRANCE PRESSE, « La victoire d'Alphago contre le champion du jeu de go restera dans l'Histoire », *Les Affaires*, sect. Technologies de l'information (14 mars 2016), en ligne : <https://www.lesaffaires.com/techno/technologie-de-l-information/la-victoire-d-alphago-contre-le-champion-du-jeu-de-go-restera-dans-l-histoire/585999> (consulté le 23 novembre 2020).

———, « Google | Une deuxième chercheuse en éthique renvoyée par le géant américain », *La Presse*, sect. Entreprises (19 février 2021), en ligne : <https://www.lapresse.ca/affaires/entreprises/2021-02-19/google/une-deuxieme-chercheuse-en-ethique-renvoyee-par-le-geant-americain.php> (consulté le 3 août 2021).

AMER-YAHIA, S., A. FAVREAU et J. SÉNÉCHAL, « D'où vient le risque ? Des données et des algorithmes », *Le Monde*, sect. Binaire (5 février 2020), en ligne : <https://www.lemonde.fr/blog/binaire/2020/02/05/les-plateformes-numeriques-un-foyer-pour-les-risques-donnees-et-algorithmes/> (consulté le 13 juillet 2020).

ANGWIN, J. et T. P. JR, « Facebook Lets Advertisers Exclude Users by Race », *ProPublica*, sect. Machine Bias (28 octobre 2016), en ligne : <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race?token=lZ_nPrh6oVJEnMzcTH1Jr59Ibe3K8XZC> (consulté le 7 juillet 2020).

ANGWIN, J., J. LARSON, S. MATTU, L. KIRCHNER, et PROPUBLICA, « Machine Bias », *ProPublica*, éd. ProPublica (2016 2016), en ligne : <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing?token=sYBNO6t1202JOb6ILFkA_eTWzPmpol3N> (consulté le 14 juillet 2020).

ANGWIN, J., A. TOBIN et M. VARNER, « Facebook (Still) Letting Housing Advertisers Exclude Users by Race », *ProPublica*, sect. Machine Bias (21 novembre 2017), en ligne : <https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin> (consulté le 7 juillet 2020).

AUDUREAU, W. et F. REYNAUD, « Jeu de go : victoire décisive de l'intelligence artificielle contre Lee Sedol », *Le Monde*, sect. Pixels (12 mars 2016), en ligne :

<https://www.lemonde.fr/pixels/article/2016/03/12/jeu-de-go-victoire-decisive-de-l-intelligence-artificielle-contre-lee-sedol_4881624_4408996.html>.

BERGER, A., « Fraude fiscale : explosion des recouvrements grâce à l'intelligence artificielle », *Capital.fr*, sect. Economie-et-politique (17 février 2020), en ligne : <https://www.capital.fr/economie-politique/fraude-fiscale-explosion-des-recouvrements-grace-a-lintelligence-artificielle-1362432> (consulté le 28 novembre 2020).

BUCCO, A., « Vera, le robot recruteur… licencié par ses employeurs », *FIGARO*, sect. Vie de Bureau (30 mai 2018), en ligne : <http://www.lefigaro.fr/vie-bureau/2018/05/30/09008-20180530ARTFIG00003-vera-le-robot-recruteur-licencie-par-ses-employeurs.php> (consulté le 7 août 2019).

———, « Le logiciel de recrutement d'Amazon n'aimait pas les femmes », *FIGARO* (11 octobre 2018), en ligne : <http://www.lefigaro.fr/social/2018/10/11/20011-20181011ARTFIG00096-le-logiciel-de-recrutement-d-amazon-n-aimait-pas-les-femmes.php> (consulté le 7 août 2019).

CELLAN-JONES, R., « Amazon scrapped "sexist AI" tool », sect. Technology (10 octobre 2018), en ligne : <https://www.bbc.com/news/technology-45809919>.

CHAMORRO-PREMUZIC, T., « Digital Staffing : The Future of Recruitment-by-Algorithm », *Harvard Business Review* (26 octobre 2012), en ligne : <https://hbr.org/2012/10/digital-staffing-the-future-of> (consulté le 3 novembre 2020).

CHINOY, S., « The Racist History Behind Facial Recognition », *The New York Times*, sect. Opinion (10 juillet 2019), en ligne : <https://www.nytimes.com/2019/07/10/opinion/facial-recognition-race.html> (consulté le 28 août 2019).

CURTIS, S., « Google Photos labels black people as "gorillas" », *The Telegraph*, sect. Technology (1 juillet 2015), en ligne : <https://www.telegraph.co.uk/technology/google/11710136/Google-Photos-assigns-gorilla-tag-to-photos-of-black-people.html>.

DASTIN, J., « Amazon scraps secret AI recruiting tool that showed bias against women », *Reuters* (9 octobre 2018), en ligne : <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> (consulté le 7 août 2019).

DEMICHELIS, R., « Le recrutement dans l'œil de l'intelligence artificielle », *Les Echos*, sect. Tech-Médias (19 février 2019), en ligne : <https://www.lesechos.fr/tech-medias/intelligence-artificielle/le-recrutement-dans-loeil-de-lintelligence-artificielle-991588> (consulté le 30 avril 2020).

DEVILLARD, A., « Voiture autonome : quand le passager et l'intelligence artificielle collaborent », *Sciences et Avenir* (7 novembre 2019), en ligne : <https://www.sciencesetavenir.fr/high-tech/intelligence-artificielle/conducteur-et-ia-coequipiers-en-vehicule-autonome_138822> (consulté le 28 novembre 2020).

———, « Les voitures autonomes et le casse-tête de la fluidité du trafic », *Sciences et Avenir* (14 novembre 2020), en ligne : <https://www.sciencesetavenir.fr/high-tech/transports/les-voitures-autonomes-et-le-casse-tete-de-la-fluidite-du-trafic_148648> (consulté le 28 novembre 2020).

DIMANCHE, L. et E. JONCHÈRES, « Les travailleurs du clic et l'intelligence artificielle », *Laboratoire de cyberjustice* (21 juin 2019), en ligne : <https://www.cyberjustice.ca/2019/06/21/les-travailleurs-du-clic-et-lintelligence-artificielle/> (consulté le 10 août 2020).

FRENKEL, S. et M. ROSENBERG, « Facebook Sued by District of Columbia Over Cambridge Analytica », *The New York Times*, sect. Technology (20 décembre 2018), en ligne : <https://www.nytimes.com/2018/12/19/technology/dc-sues-facebook-cambridge-analytica.html> (consulté le 20 décembre 2018).

GEORGES, B., « Les boîtes noires du « deep learning » », *Les Echos* (27 août 2018), en ligne : <https://www.lesechos.fr/tech-medias/intelligence-artificielle/les-boites-noires-du-deep-learning-137363>.

GRANVILLE, K., « Facebook and Cambridge Analytica : What You Need to Know as Fallout Widens », *The New York Times*, sect. Technology (19 mars 2018), en ligne : <https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html>.

HARWELL, D., « A face-scanning algorithm increasingly decides whether you deserve the job », *Washington Post*, sect. Technology (6 novembre 2019), en ligne : <https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/> (consulté le 1 juillet 2020).

———, « EPIC's FTC complaint about HireVue », *Washington Post*, sect. Technology (6 novembre 2019), en ligne : <https://www.washingtonpost.com/context/epic-s-ftc-complaint-about-hirevue/9797b738-e36a-4b7a-8936-667cf8748907/> (consulté le 3 novembre 2020).

———, « Rights group files federal complaint against AI-hiring firm HireVue, citing 'unfair and deceptive' practices », *Washington Post*, sect. Technology (6 novembre 2019), en ligne : <https://www.washingtonpost.com/technology/2019/11/06/prominent-rights-group-files-federal-complaint-against-ai-hiring-firm-hirevue-citing-unfair-deceptive-practices/> (consulté le 1 juillet 2020).

HERN, A., « "Partnership on AI" formed by Google, Facebook, Amazon, IBM and Microsoft », *the Guardian*, sect. Artificial intelligence (28 septembre 2016), en ligne : <http://www.theguardian.com/technology/2016/sep/28/google-facebook-amazon-ibm-microsoft-partnership-on-ai-tech-firms> (consulté le 15 août 2021).

HIREVUE, « HireVue's Industry Solutions From Technical Hiring to Campus Recruiting », *HireVue*, en ligne : <https://www.hirevue.com/why-hirevue/solutions> (consulté le 1 juillet 2020).

HULIN, A.-S., « Introduction à la fiducie québécoise de données », *Laboratoire de cyberjustice* (26 novembre 2020), en ligne : <https://www.cyberjustice.ca/2020/11/26/introduction-a-la-fiducie-quebecoise-de-donnees/> (consulté le 27 novembre 2020).

HUNT, E., « Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter », *The Guardian*, sect. Technology (24 mars 2016), en ligne : <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>.

ICI.RADIO-CANADA.CA, « Une IA détecterait la toux de personnes asymptomatiques atteintes de COVID-19 », *Radio-Canada.ca*, sect. Zone Techno (2 novembre 2020), en ligne : <https://ici.radio-canada.ca/nouvelle/1746457/mit-covid-19-intelligence-artificielle-toux-ia> (consulté le 28 novembre 2020).

ICI.RADIO-CANADA.CA, Z. T.-, « Pétition d'employés de Google contre une collaboration avec le Pentagone », *Radio-Canada.ca*, en ligne : <https://ici.radio-canada.ca/nouvelle/1101433/employes-google-petition-contre-collaboration-pentagone>.

JALINIÈRE, H., « Dépistage : l'IA du MIT qui prédit l'apparition du cancer du sein », *Sciences et Avenir* (30 mai 2019), en ligne : <https://www.sciencesetavenir.fr/sante/cancer/une-intelligence-artificielle-pour-predire-le-cancer-du-sein_134056> (consulté le 28 novembre 2020).

LARSON, J., S. MATTU, L. KIRCHNER et J. ANGWIN, « How We Analyzed the COMPAS Recidivism Algorithm », *ProPublica* (23 mai 2016), en ligne : <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm?token=sYBNO6t1202JOb6ILFkA_eTWzPmpol3N> (consulté le 14 juillet 2020).

LAURÈNE CHAMPALLE, « En Inde, un logiciel de reconnaissance faciale pour retrouver les enfants disparus », *leparisien.fr* (19 mai 2018), en ligne : <http://www.leparisien.fr/international/en-inde-un-logiciel-de-reconnaissance-faciale-pour-retrouver-les-enfants-disparus-16-05-2018-7719015.php>.

LIBÉRATION, « Microsoft muselle son robot "Tay", devenu nazi en 24 heures », *Libération*, sect. Futurs (25 mars 2016), en ligne : <https://www.liberation.fr/futurs/2016/03/25/microsoft-muselle-son-robot-tay-devenu-nazi-en-24-heures_1441963> (consulté le 14 juillet 2020).

LOUBIÈRE, P., « Pourquoi l'IA et la voiture autonome sont des mythes », *Challenges* (26 septembre 2019), en ligne : <https://www.challenges.fr/sommet-start-up/sommet-des-startup-de-challenges-pourquoi-l-ia-et-la-voiture-autonome-sont-des-mythes_676606> (consulté le 28 novembre 2020).

MERCURE, P., « L'intelligence artificielle contre le cancer », *La Presse*, sect. Sciences (1 mars 2020), en ligne : <https://www.lapresse.ca/actualites/sciences/2020-03-01/l-intelligence-artificielle-contre-le-cancer> (consulté le 28 novembre 2020).

———, « La génomique et l'intelligence artificielle contre la COVID-19 », *La Presse*, sect. Sciences (17 juin 2020), en ligne : <https://www.lapresse.ca/actualites/sciences/2020-06-17/la-

genomique-et-l-intelligence-artificielle-contre-la-covid-19> (consulté le 28 novembre 2020).

NOISETTE, T., « Facebook permet toujours des publicités filtrées par couleur de peau ou religion », *L'Obs*, sect. Tech (23 novembre 2017), en ligne : <https://www.nouvelobs.com/tech/20171123.OBS7749/facebook-permet-toujours-des-publicites-filtrees-par-couleur-de-peau-ou-religion.html> (consulté le 7 juillet 2020).

———, « Facebook retire 5.000 filtres publicitaires permettant des discriminations », *L'Obs*, sect. Les internets (22 août 2018), en ligne : <https://www.nouvelobs.com/les-internets/20180822.OBS1184/facebook-retire-5-000-filtres-publicitaires-permettant-des-discriminations.html> (consulté le 10 décembre 2018).

PÉRINEL, Q., « Vera, le robot qui recrute dans les grands groupes », *FIGARO* (4 avril 2018), en ligne : <http://www.lefigaro.fr/vie-bureau/2018/04/04/09008-20180404ARTFIG00111-vera-le-robot-qui-recrute-dans-les-grands-groupes.php> (consulté le 7 août 2019).

PIERRE, M., M. POIRIER, J.-P. BEAUREGARD et P. EID. *Conférence - Discrimination à l'embauche : que nous disent les recherches sur le testing de CV ?*, 11 février 2021.

PRIEST, D., C. TIMBERG et S. MEKHENNET, « Private Israeli spyware used to hack cellphones of journalists, activists worldwide », *Washington Post* (18 juillet 2021), en ligne : <https://www.washingtonpost.com/investigations/interactive/2021/nso-spyware-pegasus-cellphones/> (consulté le 3 août 2021).

RADIO-CANADA, « Google et Microsoft en désaccord sur l'interdiction de la reconnaissance faciale », *Radio-Canada.ca*, sect. Zone techno (21 janvier 2020), en ligne : <https://ici.radio-canada.ca/nouvelle/1483573/reconnaissance-faciale-moratoire-interdiction-banissement-temporaire-5-ans-europe-sundar-pichai-brad-smith> (consulté le 14 août 2021).

REYNOLDS, C., « Réglementation de l'intelligence artificielle : le Canada à la traîne », *La Presse*, sect. Techno (19 mai 2019), en ligne : <https://www.lapresse.ca/affaires/techno/201905/19/01-5226739-reglementation-de-lintelligence-artificielle-le-canada-a-la-traine.php> (consulté le 30 avril 2020).

RIDLEY, M., *Explainable AI: Explanations, Expectations, and Options*, , Montréal, 2019, en ligne : <https://www.cyberjustice.ca/2019/01/22/graduate-law-and-artificial-intelligence-conference-fostering-empowerment-through-artificial-intelligence-where-do-we-go-from-here/> (consulté le 20 novembre 2020).

SCHIFFER, Z., « Timnit Gebru was fired from Google — then the harassers arrived », *The Verge* (5 mars 2021), en ligne : <https://www.theverge.com/22309962/timnit-gebru-google-harassment-campaign-jeff-dean> (consulté le 3 août 2021).

SIMONITE, T., « When It Comes to Gorillas, Google Photos Remains Blind », *Wired* (1 novembre 2018), en ligne : <https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/> (consulté le 13 octobre 2020).

SMITH, B., « Facial recognition : It's time for action », *Facial recognition: It's time for action*, en ligne : <https://blogs.microsoft.com/on-the-issues/2018/12/06/facial-recognition-its-time-for-action/>.

———, « Facial recognition technology: The need for public regulation and corporate responsibility », *Microsoft on the Issues*, en ligne : <https://blogs.microsoft.com/on-the-issues/2018/07/13/facial-recognition-technology-the-need-for-public-regulation-and-corporate-responsibility/>.

SU, J. B., « SmartRecruiters Unveils Artificial Intelligence Recruiting Assistant, Hiring Scorecard », *Forbes*, en ligne : <https://www.forbes.com/sites/jeanbaptiste/2018/03/14/smartrecruiters-unveils-artificial-intelligence-recruiting-assistant-hiring-scorecard/> (consulté le 7 août 2019).

TRUDEL, P. et A.-S. HULIN. *Web-discussion | La fiducie de données : le bon véhicule juridique pour encadrer le développement de l'intelligence artificielle ?*, 24 novembre 2020.

VENNE, J.-F., « L'intelligence artificielle en bref », *Le Devoir* (28 novembre 2020), en ligne : <https://www.ledevoir.com/societe/science/590344/en-bref> (consulté le 28 novembre 2020).

ZARKADAKIS, G., « "Data Trusts" Could Be the Key to Better AI », *Harvard Business Review* (10 novembre 2020), en ligne : <https://hbr.org/2020/11/data-trusts-could-be-the-key-to-better-ai> (consulté le 26 novembre 2020).

« Amazon exhorté à ne plus fournir son outil de reconnaissance faciale à la police », *La Presse* (22 mai 2018), en ligne : <https://www.lapresse.ca/techno/actualites/201805/22/01-5182761-amazon-exhorte-a-ne-plus-fournir-son-outil-de-reconnaissance-faciale-a-la-police.php> (consulté le 19 décembre 2018).

« Ciblage publicitaire. Facebook est accusé de favoriser la discrimination à l'égard des femmes », *Courrier international* (19 septembre 2018), en ligne : <https://www.courrierinternational.com/revue-de-presse/ciblage-publicitaire-facebook-est-accuse-de-favoriser-la-discrimination-legard-des> (consulté le 10 décembre 2018).

« IA : les fermes à clics remplacent les sweatshops », *Les affaires* (21 mars 2019), en ligne : <https://www.lesaffaires.com/blogues/diane-berard/ia-les-fermes-a-clics-remplacent-les-sweatshops/609039> (consulté le 14 juillet 2020).

« Fraude fiscale : la manne de l'intelligence artificielle », *Le Point* (17 février 2020), en ligne : <https://www.lepoint.fr/economie/fraude-fiscale-la-manne-de-l-intelligence-artificielle-17-02-2020-2363033_28.php> (consulté le 28 novembre 2020).

« Nombre d'utilisateurs de Facebook dans le monde », *Journal du Net* (31 juillet 2020), en ligne : <https://www.journaldunet.com/ebusiness/le-net/1125265-nombre-d-utilisateurs-de-facebook-dans-le-monde/> (consulté le 7 octobre 2020).

« « Projet Pegasus » : révélations sur un système mondial d'espionnage de téléphones », *Le Monde.fr* (18 juillet 2021), en ligne : <https://www.lemonde.fr/projet-pegasus/article/2021/07/18/projet-pegasus-revelations-sur-un-systeme-mondial-d-espionnage-de-telephones_6088652_6088648.html> (consulté le 3 août 2021).

« Toucher de nouvelles audiences », *Pages d'aide de Facebook Business*, en ligne : <https://fr-fr.facebook.com/business/help/717368264947302> (consulté le 31 mai 2020).

« What is a data trust? – The ODI », en ligne : <https://theodi.org/article/what-is-a-data-trust/> (consulté le 26 novembre 2020).

Dictionnaires et ouvrages de références

Dictionnaire de la comptabilité et de la gestion financière, 2006

Dictionnaire Larousse, Larousse

Dictionnaire encyclopédique du Droit québécois, Gaudet Éditeur ltée

Le Petit Robert, Le Robert

OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE, *Grand dictionnaire terminologique*

REID, H. et S. REID, *Dictionnaire de droit québécois et canadien*, 2016

Sites internet

https://www.cdpdj.qc.ca/fr

https://www.chrc-ccdp.gc.ca/fr

https://www.chrt-tcdp.gc.ca/index-fr.html

https://www.cyberjustice.ca/

https://foldoc.org/

https://www.justice.gc.ca/fra/

https://www.ssrn.com/index.cfm/en/

http://www.tribunaux.qc.ca/TDP/index-tdp.html