



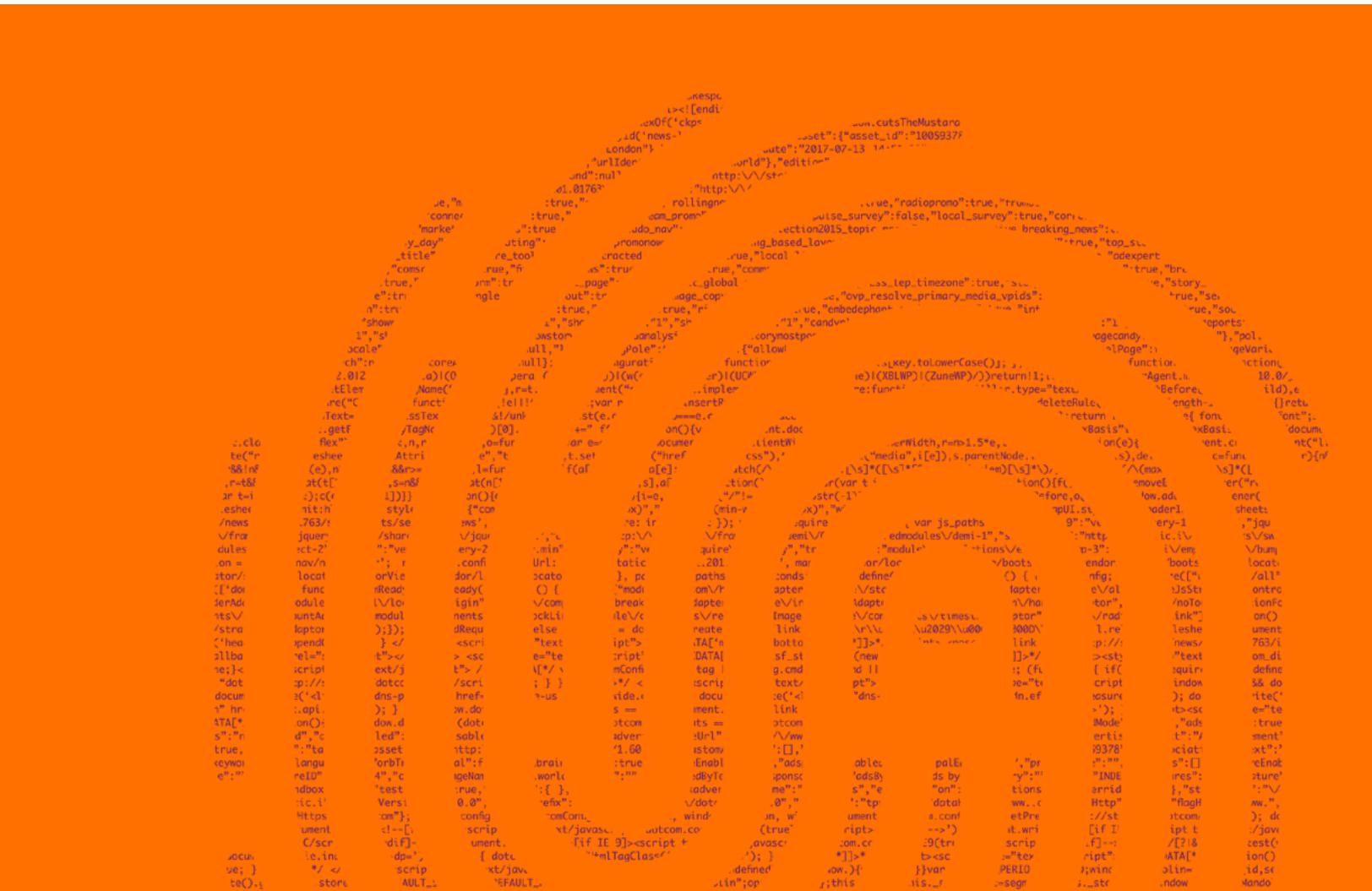
< >

# Déclaration de Montréal IA responsable\_

</ >

## PARTIE 4

# LA COCONSTRUCTION DE L'AUTOMNE 2018 : LES ACTIVITÉS CLÉS



# TABLE DES MATIÈRES

Ce document est une partie du  
**RAPPORT DE LA DÉCLARATION DE MONTRÉAL  
 POUR UN DÉVELOPPEMENT RESPONSABLE  
 DE L'INTELLIGENCE ARTIFICIELLE 2018.**  
 Vous retrouverez le rapport complet [ICI](#).

<b>1. INTRODUCTION</b>	<b>158</b>
<b>2. JOURNÉE DE COCONSTRUCTION HORS QUÉBEC (PARIS, FRANCE)</b>	<b>159</b>
2.1 Aborder les enjeux démocratiques à travers les fausses nouvelles	162
2.2 Aborder les enjeux liés à l'environnement	166
2.3 Aborder les enjeux de la transition numérique dans le monde du travail	170
<b>3. DISCUSSION AUTOUR DU THÈME DE LA CULTURE avec les membres de la Coalition de la diversité des expressions culturelles (CDEC)</b>	<b>177</b>
3.1 Trois thématiques proposées par l'équipe de la Déclaration pour aborder les enjeux du développement de l'IA dans le domaine de la culture	177
3.2 Les enjeux de la promotion de la diversité culturelle à l'heure de l'IA	178
3.3 Les principes éthiques de la CDEC	180
3.4 Quelques-unes des recommandations formulées	182
<b>4. FAIRE LE PONT ENTRE LES DÉLIBÉRATIONS CITOYENNES ET LA RELÈVE EN RECHERCHE : Simulation de la rédaction de brèves politiques</b>	<b>183</b>
4.1 Description de l'activité	183
4.2 Les problématiques issues des préoccupations citoyennes	185
Problématique 1 : Sécurité publique et intégrité des systèmes	185
Problématique 2 : L'IA, les médias et la manipulation de l'information	186
Problématique 3 : Gouvernance publique, privée ou participative : les communs numériques	187
4.3 Les recommandations de la relève en recherche	188

<b>5. CONCLUSION</b>	<b>190</b>
<b>ANNEXE 1 – LES SCÉNARIOS</b>	<b>191</b>
Démocratie	191
Environnement	192
Monde du travail	193
<b>ANNEXE 2 – LES BRÈVES ÉTUDIANTES</b>	<b>194</b>
CRÉDITS	I
PARTENAIRES	II
<b>GRAPHIQUES, SCHÉMAS ET TABLEAUX</b>	
Graphiques 1, 2, 3 et 4 : Profils des participants à la journée de coconstruction de Paris	159
Schéma 1 : Enjeux soulevés et pistes de solution proposées lors de la journée de coconstruction de Paris	161
Tableau 1 : Démocratie, Premier moment délibératif : formulation d'enjeux éthiques en 2022	164
Tableau 2 : Démocratie, Deuxième moment délibératif : propositions d'encadrement de l'IA pour 2018-2020	164
Tableau 3 : Environnement, Premier moment délibératif : formulation d'enjeux éthiques en 2025	168
Tableau 4 : Environnement, Deuxième moment délibératif : propositions d'encadrement pour 2018-2020	169
Tableau 5 : Les enjeux prioritaires	173
Tableau 6 : Les propositions retenues	175
Graphique 5 : Profils des étudiants participants à l'activité en fonction de leur domaine d'étude (selon les secteurs des trois Fonds)	184

# 1. INTRODUCTION

Les principales activités de coconstruction de la Déclaration de Montréal se sont déroulées du 3 novembre 2017 au 31 avril 2018. Toutefois, l'équipe de la Déclaration a poursuivi ses activités lors de l'automne 2018 en organisant trois activités clés afin de mobiliser les connaissances d'un plus grand nombre d'acteurs sur différents thèmes essentiels lorsqu'il est question du développement responsable de l'intelligence artificielle (IA). Ainsi, une journée de coconstruction à Paris a été organisée selon le modèle des coconstructions de l'hiver 2018. Afin de mobiliser les connaissances de parties prenantes du secteur culturel, un groupe de discussion sur les enjeux liés à l'avènement de l'IA dans les domaines relatifs à l'art et la culture a également discuté de ces questions. Enfin, afin de faire le pont entre les délibérations citoyennes de l'hiver et la relève en recherche, une activité de simulation de rédaction de brèves politiques avec des étudiants des cycles supérieurs a été réalisée en partenariat avec le Comité intersectoriel étudiant (CIÉ) des Fonds de recherche du Québec (FRQ).

Ces différentes activités sont venues alimenter l'analyse et la rédaction des recommandations formulées en vue de politiques publiques (cf. rapport *Les chantiers prioritaires et recommandations pour un développement responsable de l'IA*). Les sections suivantes récapitulent les enjeux identifiés et les principales pistes de solutions formulées par les personnes ayant participé aux activités. Certaines reprennent les mécanismes proposés lors de la coconstruction de l'hiver et viennent d'autant plus appuyer la nécessité de leur mise en place, tandis que d'autres ont un caractère inédit.

## RÉDACTION

**NATHALIE VOARINO**, coordonnatrice scientifique de l'équipe de la Déclaration, candidate au doctorat en bioéthique, Université de Montréal

**CHRISTOPHE ABRASSART**, professeur à l'École de design, Université de Montréal

**CAMILLE VÉZY**, candidate au doctorat en communication, Université de Montréal

## COLLABORATION

**LOUBNA MEKKI BERRADA**, doctorante en neuropsychologie clinique, Université de Montréal

**VINCENT MAI**, doctorant en robotique, Université de Montréal

Dans ce document, l'utilisation du genre masculin a été adoptée afin de faciliter la lecture et n'a aucune intention discriminatoire.

## 2. JOURNÉE DE COCONSTRUCTION HORS QUÉBEC (PARIS, FRANCE)

Cette section présente les résultats de la journée de coconstruction réalisée à Paris le 9 octobre 2018, organisée en partenariat avec l'Ambassade du Canada à Paris, le Centre culturel canadien et la Maison des étudiants canadiens. Lors de cette journée, 26 personnes aux profils variés

ont été mobilisées afin de se pencher sur les enjeux du développement responsables de l'IA. Les participants ont été regroupés autour de trois tables de coconstruction, chacune abordant un thème clé du développement de l'IA : la démocratie, l'environnement et le monde du travail.

Figure 1 : Profils des participants à la journée de coconstruction de Paris.

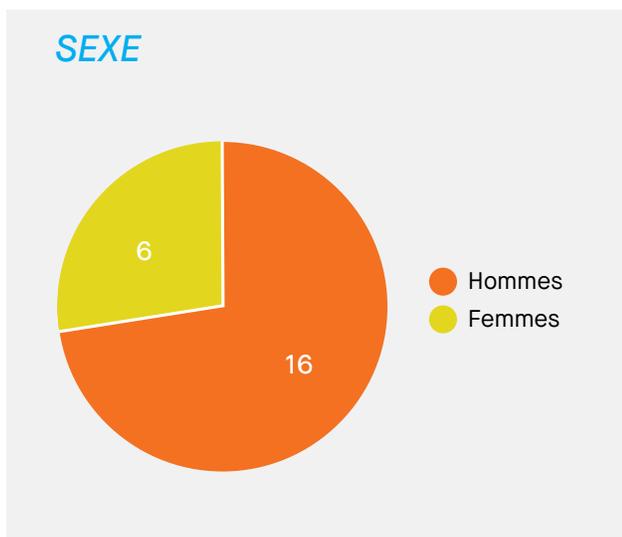


Figure 2 : Profils des participants à la journée de coconstruction de Paris.

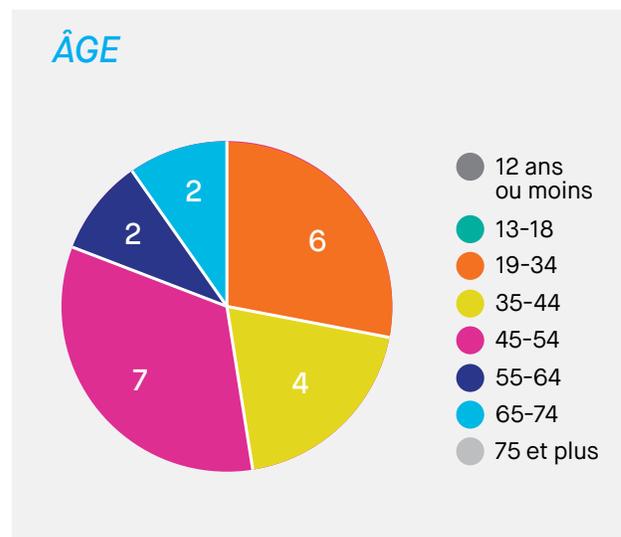


Figure 3 : Profils des participants à la journée de coconstruction de Paris.

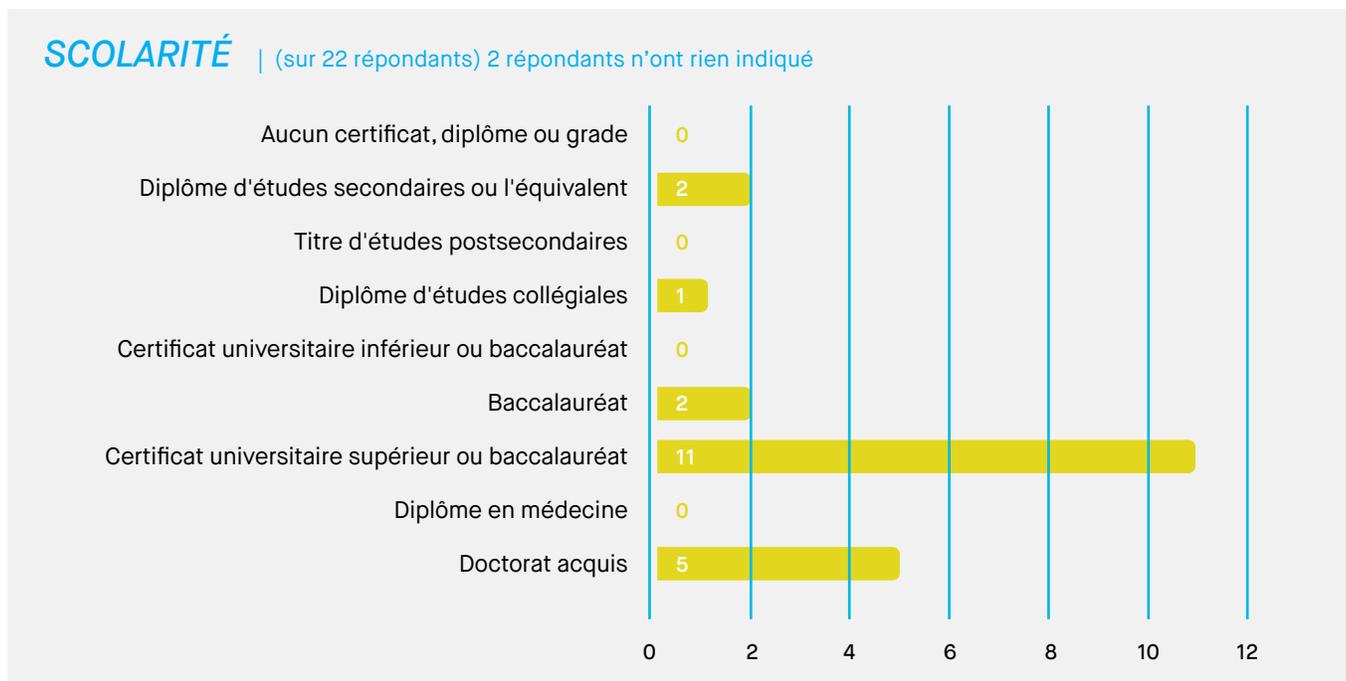
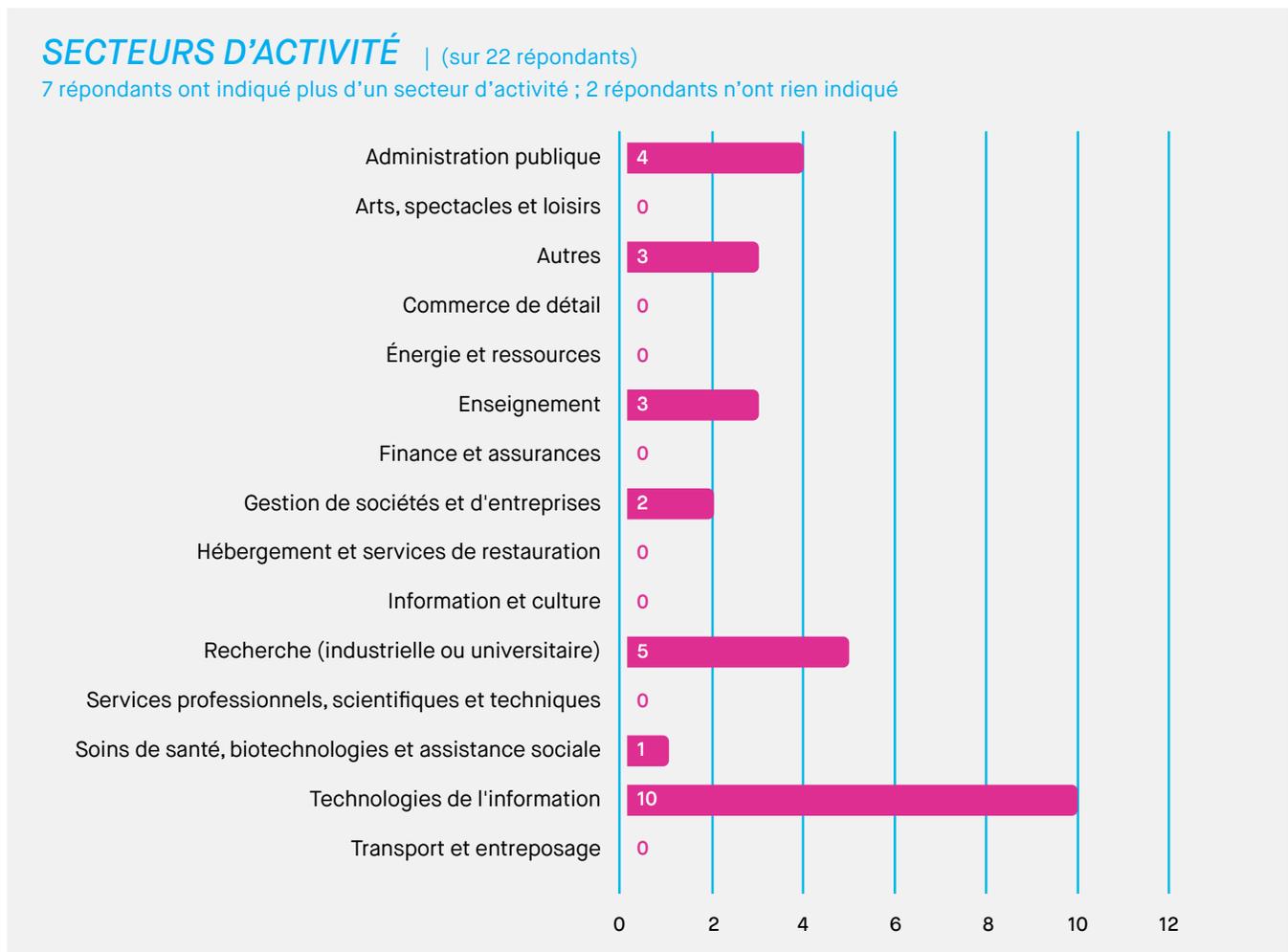


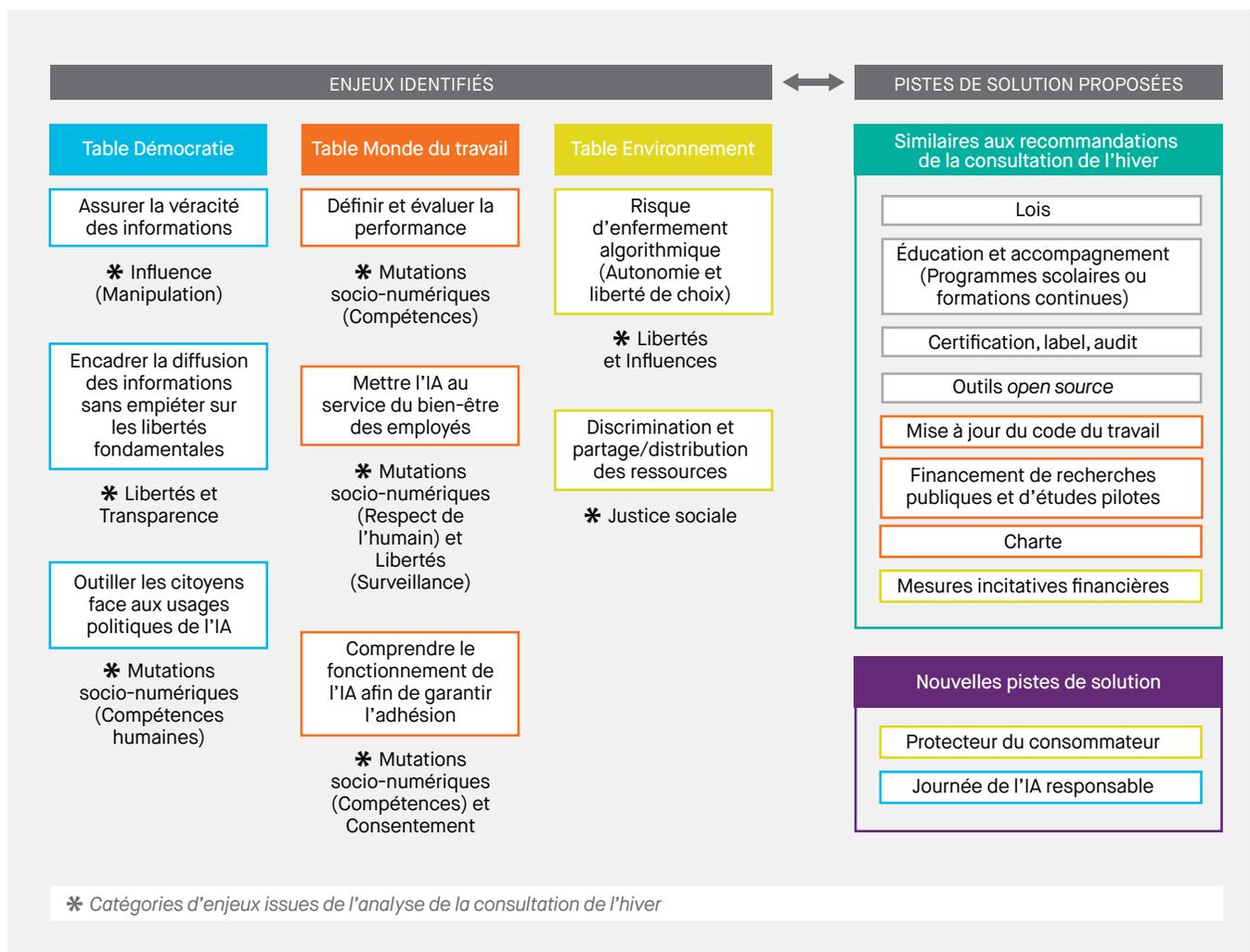
Figure 4 : Profils des participants à la journée de coconstruction de Paris.



Les discussions se sont organisées autour de trois moments délibératifs (identifications des enjeux, formulation de recommandations et écriture de la « Une » d'un journal), suivant le modèle de coconstruction développé pour les activités de l'hiver. Trois scénarios déclencheurs ont ici été utilisés (cf. Annexe 1) mettant en scène l'utilisation de l'IA dans le cadre de l'incitation aux comportements écologiques, de la gestion de personnel dans l'entreprise, et de l'utilisation de l'IA pour la création de fausses nouvelles dans le cadre d'une campagne électorale. Ces trois scénarios nous ont permis d'explorer les enjeux relatifs à la démocratie, à l'environnement et au monde du travail sous un angle nouveau, n'ayant jamais été utilisé sous cette forme précédemment.

Les sections suivantes retracent le parcours de chacune des discussions ayant eu lieu autour des trois tables<sup>1</sup>. Ces discussions ont mis en évidence l'apparition d'enjeux déjà identifiés lors de la coconstruction de l'hiver, mais selon les spécificités des contextes donnés. Également, différentes pistes de solution ou mécanismes d'encadrement pour un développement responsable de l'IA ont été proposés. Certains sont similaires à ceux formulés lors des discussions de l'hiver (ex. Lois et Formations) et d'autres sont nouveaux (ex. Journée de l'IA responsable).

**Schéma 1 : Enjeux soulevés et pistes de solution proposées lors de la journée de coconstruction de Paris**



<sup>1</sup> Les citations sont tirées de notes rédigées par des participants.

## 2.1.

### ABORDER LES ENJEUX DÉMOCRATIQUES À TRAVERS LES FAUSSES NOUVELLES

#### Résumé du scénario de départ

##### Fausse nouvelle dans la campagne électorale

À deux semaines des élections présidentielles, une réunion de crise a lieu au sein de l'Agence sur l'intégrité de l'information (All) mise en place dans le cadre de la Loi contre la manipulation de l'information. Une vidéo diffusant des propos compromettants du président sortant sur des travailleurs immigrés est devenue virale. Le porte-parole de la présidence a fait savoir qu'il s'agissait d'un faux créé par une agence étrangère qui tentait d'interférer dans les élections en tirant parti des algorithmes GAN (*generative adversarial networks*). Bien que la diffusion de la vidéo soit interdite, celle-ci continue à circuler via différents sites étrangers. À un mois du premier tour des élections présidentielles, l'All doit présenter un plan pour enrayer les effets dévastateurs de cette fausse information et rétablir les conditions d'une campagne électorale saine.

L'objectif de ce scénario était d'ouvrir une discussion sur les enjeux éthiques liés à la manipulation de l'information pouvant nuire aux démocraties par leur diffusion virale. Et ce, en particulier quand des techniques d'intelligence artificielle permettent d'imiter des personnes et modifier des contenus en conservant un très haut degré de réalisme rendant particulièrement difficile la détection du faux.

Le parcours délibératif présenté est issu d'une journée complète de discussion regroupant sept chercheurs, experts et étudiants travaillant en

éthique, en développement organisationnel, en machine learning, en web social et en sciences politiques. Partant de ce scénario en 2022, la discussion a conduit à la formulation d'une « Une » du Journal de l'IA responsable du 9 octobre 2022 : « Première journée citoyenne de l'IA responsable ».

#### Premier moment délibératif : Formulation d'enjeux éthiques en 2022

##### DÉMOCRATIE

Les participants sont d'avis que l'IA en elle-même n'est pas la cause d'atteintes à la démocratie : ces problèmes existent déjà, ils sont cependant nuancés ou amplifiés par les possibilités permises par l'IA. Il faut donc se préparer et s'adapter à cette nouvelle réalité. « La santé de la démocratie » a alors été avancée comme un enjeu particulièrement important dans un contexte où les « choix basés sur une information erronée » sont problématiques. L'encadrement par le droit et la formation de l'esprit critique ont été abordés comme deux pistes de réflexion pour se prémunir contre les effets de la manipulation de l'information et ainsi veiller à la santé de la démocratie.

##### ENCADREMENT PAR LE DROIT

Les participants se sont demandé comment les pratiques liées à la manipulation de l'information pourraient être encadrées :

#### « Le droit peut-il encadrer l'entière des pratiques de l'IA (manipulation de l'information) ? »

##### Un participant

Dans cette perspective, la question du « comment faire justice » quand la réputation est entachée par la propagation de fausses nouvelles a été mentionnée plusieurs fois. Cela se heurte cependant à la difficulté d'encadrer la manipulation de l'information sans entraver la liberté d'expression et les autres libertés fondamentales.

## CONNAISSANCE ET ESPRIT CRITIQUE

« [L'] importance de la corrélation démocratie/éducation (notamment l'éducation à la pensée critique) » a été soulignée, marquant un appel au développement de l'esprit critique pour la « participation éclairée à la vie publique ». Un participant précise que plutôt que de se prémunir contre la propagande elle-même, le développement de l'esprit critique permet de se prémunir contre les effets de la propagande (endoctrinement, modifications de choix et comportements, polarisation vers les extrêmes, etc.) En effet, empêcher toute propagande, c'est-à-dire toute action exercée sur l'opinion, risquerait de mener à la censure ou d'entraver la liberté d'expression. Il faudrait donc plutôt miser sur l'éducation des citoyens pour développer la « pensée critique » ainsi que la « littératie médiatique et statistique ».

L'enjeu de démocratisation de l'accès à l'information et à la connaissance a par la suite été soulevé. Plusieurs conditions sont ici nécessaires : que la neutralité d'internet soit assurée afin de garantir le libre accès à toute information, et que tous aient accès à des outils technologiques permettant de se renseigner et de s'exprimer. Chacun, sans discrimination, devrait ainsi pouvoir être sensibilisé au fonctionnement de l'IA et à ses enjeux.

## AUTHENTIFICATION ET CAPACITÉ DE L'IA POUR IDENTIFIER LE FAUX

En apprenant certaines techniques d'IA, plusieurs pourraient par exemple devenir capables de développer des outils de détection et de correction de fausses nouvelles. Cela dépend cependant de la possibilité et de la nécessité d'utiliser l'IA pour identifier le faux du vrai dans un contexte où il est par exemple impossible de distinguer à l'œil nu la différence entre une vidéo créée par IA et une vidéo originale. Les participants se sont ainsi interrogés sur les « mécanismes d'authentification des nouvelles/informations » ou encore sur la « capacité de la technologie à comprendre les sarcasmes ». La « traçabilité des sources sur tous les outils de diffusion (ex : WhatsApp) » a été par exemple proposée comme piste de mécanisme d'authentification. Il a cependant été précisé

qu'une IA qui censurerait automatiquement les publications considérées comme fausses, malveillantes ou non fiables aurait un effet négatif sur la liberté d'expression.

## RESPONSABILITÉ

Face à l'impression de « perte de contrôle de l'information » dans un contexte de propagation très rapide, voire virale des fausses nouvelles, l'enjeu de la responsabilité a été abordé à plusieurs étapes : création, diffusion, partage, lecture des informations. Les participants se sont demandé à quel moment la véracité des nouvelles devrait être évaluée (avant sa création, avant sa publication, avant son partage, pendant sa lecture, etc.), par quelle entité (organisation internationale, État, journalistes, lecteurs, plateforme de diffusion qui permet la publication et/ou le partage tel que WhatsApp ou Facebook), et de quelle manière (développement de l'esprit critique du lecteur, création d'un label indiquant le niveau de véracité de l'information et de fiabilité de la source).

La création et la diffusion de fausses nouvelles font ainsi appel à une combinaison des rôles de plusieurs acteurs devant assumer des responsabilités quant au fait de créer des informations, les diffuser, juger de leur crédibilité, vérifier leur véracité et la fiabilité des sources qu'ils soient journalistes, lecteurs ou autres. Cette entité devrait évaluer la véracité des nouvelles, la plateforme de diffusion d'information ainsi que tout individu qui crée ou diffuse des informations à titre personnel.

Les participants ont ainsi discuté de la « déontologie des médias (captation et diffusion d'informations) » et de la « crédibilité du journalisme ». Un nouveau rôle des journalistes et médias pourrait être de juger l'information publiée et partagée en vérifiant et déclarant si celle-ci provient ou non d'une source fiable.

Tableau 1 : Démocratie, Premier moment délibératif : formulation d'enjeux éthiques en 2022

Enjeux éthiques 2022	1	2	3
Description	Assurer la véracité des informations	Encadrer la diffusion des informations sans empiéter sur les libertés fondamentales	Outiller les citoyens face aux usages politiques de l'IA
Principes associés	Démocratie, responsabilité	Démocratie, autonomie	Démocratie, connaissance, autonomie

À la suite de cette discussion, les participants ont formulé ces trois enjeux prioritaires en proposant des pistes d'encadrement afin de veiller à la santé de la démocratie face aux effets de la propagande via la création et la diffusion automatisées de fausses informations :

1. Assurer la véracité et la fiabilité des informations, notamment pour préserver la santé des échanges démocratiques.
2. Encadrer la diffusion des informations sans empiéter sur les libertés fondamentales, dont la liberté d'expression, en particulier par le développement de normes journalistiques et technologiques en matière de diffusion de l'information.

3. Outiller les citoyens quant aux usages politiques de l'IA, pour que ceux-ci puissent se renseigner et développer leur propre opinion librement.

## Deuxième moment délibératif : Propositions d'encadrement de l'IA

Pour répondre à ces enjeux, l'équipe a poursuivi ses discussions pour formuler ensemble cinq propositions d'encadrement de l'IA :

Tableau 2 : Démocratie, Deuxième moment délibératif : propositions d'encadrement de l'IA pour 2018-2020

Propositions d'encadrement	1	2	3	4	5
Description	Création d'une autorité de certification et de normes journalistiques via un label (signal rouge, jaune, vert)	Création d'outils <i>open source</i> pour détecter/ distinguer le faux du vrai. (ex : application sur téléphone mobile)	Projet de loi garantissant la neutralité du web.	Mise en place de programmes scolaires + formation continue (par exemple MOOC) pour le développement de l'esprit critique	Mise en place d'une journée de l'IA citoyenne responsable
Enjeu associé	Assurer la véracité des informations	Encadrer la diffusion des informations sans empiéter sur les libertés fondamentales		Outiller les citoyens face aux usages politiques de l'IA	

Dans la lignée des réflexions précédentes, les discussions menant à ces propositions ont porté sur les dimensions techniques, politiques et éducatives du problème de la manipulation de l'information. Les participants ont ainsi proposé un ensemble de mesures visant à encadrer la vérification des nouvelles (tables 1 et 2) et à éduquer les citoyens pour participer à la vie démocratique de façon libre et éclairée (tables 3, 4 et 5). Cela passe par la normalisation des pratiques de création et diffusion de l'information (certification) et l'aide à la détection du faux (outil), l'accès égal de tous à l'information via un réseau internet neutre (projet de loi), l'éducation pour tous, à tous les âges de la vie pour le développement de l'esprit critique (programmes scolaires et formations continues) et la sensibilisation aux usages politiques de l'IA (journée de l'IA responsable).

La création d'une autorité de certification établissant des normes journalistiques via un système d'indicateur de fiabilité des informations a émergé de la réflexion autour de la responsabilité des journalistes pour établir la fiabilité des sources et la crédibilité des informations. Les participants ont été d'avis que cet organisme de certification devait être indépendant du gouvernement.

Plutôt qu'un encadrement juridique, une approche par les normes professionnelles et l'établissement de lignes directrices à la création, la vérification et la diffusion des informations a été préférée. Un participant a souligné qu'il était plus raisonnable d'adopter une approche minimaliste dans la conception de ce système indicatif de la fiabilité d'une information. En effet, il faudrait éviter qu'un tel système influence davantage les processus électoraux et autres situations de compétition où les nouvelles jouent un rôle crucial. Ce serait le cas par exemple si le système indicatif de fiabilité pouvait donner avantage à un parti ou être utilisé comme une stratégie contre l'adversaire dans une perspective de manipulation ou de lobbyisme. En même temps, en optant pour une approche minimaliste, on pourra mieux éviter des limitations excessives risquant de porter atteinte aux libertés fondamentales.

En optant pour cette solution préventive, les participants n'ont pas proposé de solutions réactives en cas de diffusion d'une fausse nouvelle.

Ils ont toutefois mentionné la nécessité dans ce cas d'interrompre le plus rapidement possible la propagation et de contrecarrer les fausses nouvelles par leur version vérifiée.

Les pratiques journalistiques de vérification de nouvelles instaurées par cette autorité de certification pourraient avoir recours à un outil *open source* dont la technologie permettrait de distinguer le faux même quand celui-ci est indétectable à l'œil nu, en particulier quand ce faux est créé par des techniques d'IA. Ces pratiques devraient aussi permettre de vérifier que certains renseignements véridiques ne soient pas validés comme étant des fausses nouvelles.

Soulignant également la part de responsabilité importante des plateformes de partage d'information comme les médias sociaux (Facebook, Twitter, Youtube, Snapshat, etc.) et les messageries instantanées (ex. : WhatsApp), les participants se sont interrogés sur la nécessité d'obliger ces plateformes d'instaurer un outil faisant appel à l'IA pour identifier les fausses informations. Des participants ne semblaient pas faire confiance aux entreprises privées pour instaurer un système d'identification et de blocage de fausses nouvelles qui soit juste, étant donné que pour l'instant ce type de plateforme laisse en place des vidéos contenant de faux renseignements qui influent sur l'opinion des différents publics. Un système indiquant le niveau de fiabilité d'une nouvelle serait cependant intéressant à voir apparaître sur ces plateformes qui doivent être absolument responsabilisées par rapport au rôle qu'elles jouent dans la propagation des informations manipulées qui affectent les démocraties.

La formation et la sensibilisation des citoyens font l'objet d'un deuxième ensemble de mesures visant le développement de l'esprit critique des citoyens, en particulier par la littératie médiatique qui les outillerait pour naviguer de façon libre et éclairée dans l'univers de l'information. Cela peut passer à la fois par les curriculums scolaires et par les espaces publics ouverts à tous, via des universités populaires dans des bibliothèques ou des cafés par exemple, ou bien des campagnes créatives de sensibilisation qui s'immisceraient dans la vie quotidienne pour entretenir la vigilance quant aux pratiques de manipulation de l'information.

## Troisième moment délibératif : Écriture de la « Une » du Journal en 2022

Ces propositions ont ensuite été mises en récit, à la « Une » du Journal de l'IA responsable du 9 octobre 2022 formulée ainsi :

### Première journée citoyenne de l'IA responsable

*Dans le cadre de la première journée internationale de l'IA responsable, plusieurs événements ont été organisés simultanément au Canada et en France, notamment afin d'informer et outiller les citoyens face aux nouvelles possibilités offertes par l'IA et à l'importance de la pensée critique. À cet effet, notre journal est très fier de vous informer qu'il a reçu son accréditation de l'Ordre des journaux IA responsable.*

En projetant la lumière sur la « Journée citoyenne de l'IA responsable », les participants soulèvent ainsi l'importance de la sensibilisation et l'appropriation des enjeux de l'IA par l'ensemble des citoyens pour pouvoir participer à la vie démocratique. La mention de l'obtention de l'accréditation de la publication à titre de « journal IA responsable » traduit également un souci de renforcement du rôle des médias et du journalisme à titre d'acteurs importants de la santé démocratique.

## 2.2. ABORDER LES ENJEUX LIÉS À L'ENVIRONNEMENT

### Résumé du scénario 2025 de départ

Les températures continuent à battre des records de chaleur dans le monde entier. Pour faire face à cette crise du changement climatique, des villes proposent à leurs habitants un système de permis carbone individuel fortement incitatif, le système EcoFit, qui est connecté à leur compte bancaire et aux différentes applications d'achat en ligne : dans ces villes, le prix des biens et services est affiché en euros et en carbone, et chaque citoyen doit viser 4 tonnes d'émission de carbone par an pour l'ensemble de sa consommation. Cette cote leur donne un accès à de multiples services écoresponsables en transports, éducation, formation et culture. Au cours de l'année 2025, Ivo et Charles réussissent à ajuster progressivement leur consommation sur cette cible, et même à faire mieux. Et comme ils ont moins dépensé, ils ont réalisé une épargne inattendue. Ils imaginent alors un projet de séjour à Cuba pour Noël et commencent à consulter les sites de voyages. Un message leur parvient alors sur leur téléphone : « Attention à l'effet rebond : dépenser vos économies dans un voyage en avion annulerait tous vos efforts ! Pensez à voyager local ! »

L'objectif de ce scénario était d'ouvrir une discussion sur les possibilités autant que sur les enjeux éthiques d'une gestion prédictive par SIA des effets rebonds environnementaux sur les marchés de biens de consommation et d'équipements. Les effets rebonds peuvent s'expliquer de la manière suivante. Alors que l'efficacité énergétique des équipements

s'améliore, alors que l'empreinte environnementale des biens de consommation diminue grâce à l'écoconception, plutôt que de verrouiller ces acquis, on consomme proportionnellement plus d'équipements, de biens et de services : par exemple la taille des écrans augmente, le taux d'équipement des ménages augmente, on parcourt plus de km en automobile, on voyage en avion, etc. Il en résulte une augmentation des émissions de GES et une pression accrue sur les ressources et la biodiversité. Avec ces effets rebonds, il n'y a donc pas de découplage entre le développement économique d'une part, et sa matérialité et son empreinte écologique d'autre part.

Le dispositif algorithmique imaginé ici se situe dans une approche d'utilisation des SIA pour la planète (« AI for Earth »). La possibilité d'une gestion prédictive et personnalisée des effets rebonds, par apprentissage supervisé sur des historiques de consommation (ex. des données de transactions bancaires), associée à des dispositifs incitatifs (nudges) a ainsi été mise en scène dans ce scénario exploratoire.

Le parcours délibératif présenté est issu d'une table tenue en 3 h avec huit citoyens intéressés par les nouvelles technologies et les enjeux d'environnement et de développement durable. Partant de ce scénario en 2025, la discussion a débouché sur la formulation d'une initiative présentée en « Une » du Journal de l'IA responsable du 9 octobre 2020 : « Grand succès pour ConsoM'IA : 1 million d'abonnés en une semaine. Règlement général d'ouverture des data pour l'environnement ».

Quel a été le parcours délibératif de ce groupe pour mener à cette proposition originale ? Quels ont été ses moments marquants ? Comment se sont enrichies les idées à chaque étape ? Nous présentons, en les commentant, certains moments significatifs du parcours suivi par cette équipe.

## Premier moment délibératif : formulation d'enjeux éthiques en 2025

De nombreuses interrogations rédigées sur des *Post-it* ont été formulées par les participants en relation avec différents principes de la Déclaration de Montréal :

### LE PRINCIPE DE VIE PRIVÉE

« Pourra-t-on retracer tout l'historique de consommation de cette famille ? », « La gestion des bases de données sera-t-elle assez fiable pour protéger les données personnelles et donner confiance aux usagers ? », « Pourrait-il y avoir un droit à l'effacement ? ».

### LE PRINCIPE D'AUTONOMIE ET DE LIBERTÉ DE CHOIX

Ce dispositif conduit-il à un nouveau « pouvoir prescriptif ? », « Permet-il de conserver une autonomie dans la décision et le libre arbitre ? », « Comment avoir un jugement critique sur ces recommandations personnalisées ? », « Assiste-t-on au contrôle de la vie courante par une machine ? », « Y a-t-il un risque d'enfermement algorithmique, de bulles algorithmiques ? », « Comment un tel système peut-il prendre en compte le contexte singulier d'une décision d'achat (ex. une situation d'urgence) ? ».

### LE PRINCIPE DE RESPONSABILITÉ

Ce dispositif doit aider à « renforcer la responsabilité environnementale au quotidien », « à vivre une éthique personnelle comme ConsoActrice », mais avec le recours à des outils d'IA, « y a-t-il un risque d'externalisation de la responsabilité individuelle ? »

### LE PRINCIPE DE JUSTICE ET ÉQUITÉ

Ce dispositif, en demandant aux entreprises d'évaluer l'empreinte carbone de leurs produits et services préalablement à leur introduction sur le marché, permettra-t-il « d'assurer une libre

concurrence ? », « Risque-t-il de renforcer le pouvoir des grandes entreprises et de créer une barrière à l'entrée discriminante pour les PME, avec le coût de ces bilans environnementaux ? », « Comment le commerce équitable, qui apporte d'autres dimensions éthiques, sera-t-il évalué ? », « Certains producteurs seront-ils privilégiés ? ». « Les riches pourront se permettre de consommer plus et d'acheter des quotas de carbone pour compenser leurs émissions ! C'est une inégalité sociale ! » Par ailleurs, si « tout doit passer par les données et le marché », « les initiatives de réduction des GES hors marché (ex. projet d'habitants de quartier en mobilité active, en agriculture urbaine) seront-elles rendues invisibles et donc discriminées ? ». Enfin, « les styles de vie sont différents dans le monde, les régimes alimentaires sont variés (ex. végane, religieux), risque-t-on d'en favoriser certains et d'en discriminer d'autres ? », « de créer des discriminations culturelles ? ».

## LE PRINCIPE DE DÉMOCRATIE ET GOUVERNANCE

« Qui va réguler ce système ? Est-ce les Nations Unies ? Est-ce les pays riches ? Comment surveiller les abus ? », « Devrait-on avoir une autorité régulatrice sur les bilans de carbone ? », « Si on fait une économie de CO2, pourrait-on transmettre nos économies de carbone à nos proches ? », « Les recommandations à prioriser devraient-elles être discutées de manière démocratique ? ».

Plusieurs discussions approfondies ont ensuite eu lieu, les participants rebondissant sur les premières idées pour en générer d'autres. Puis, après près de 45 minutes de discussion, les participants ont sélectionné, à l'aide de pastilles colorées, des regroupements d'enjeux éthiques pour 2025 qui leur semblaient prioritaires. Deux principes de la Déclaration de Montréal sont ainsi ressortis : autonomie, conjugué à la liberté de choix, et justice, associé par les participants au principe d'équité.

Tableau 3 : Environnement, Premier moment délibératif : formulation d'enjeux éthiques en 2025

Enjeux éthiques 2025	1	2
Description	Risque d'enfermement algorithmique face à ce nouveau pouvoir de prescription et à la configuration des espaces de choix pour chacun. Comment garder une autonomie individuelle et collective ? Comment valoriser une initiative de réduction des émissions de carbone hors du système ?	Les plus riches pourraient être favorisés par la compensation carbone : quelle limite leur assigner ? Inversement, ceux qui consomment peu pourraient-ils redistribuer leur épargne de carbone ? Quelles relations entre pays du Nord et du Sud ? Y a-t-il des risques de discrimination culturelle ?
Principes associés	Autonomie et liberté de choix	Justice et équité

Cette sélection d'enjeux prioritaires par l'équipe conduit à développer plus précisément deux enjeux éthiques des SIA. Le premier est lié au principe d'autonomie, avec la possibilité d'actions visant la réduction des gaz à effet de serre hors marché

(ex. une initiative citoyenne sur la mobilité quotidienne). Le second a trait au principe de justice, avec la possibilité de compensation carbone pour les plus riches, ou de partage des émissions pour les citoyens ayant une consommation plus restreinte.

## Deuxième moment délibératif : propositions d'encadrement de l'IA pour 2018-2020

Pour répondre à ces enjeux, l'équipe a poursuivi ses discussions en essayant de réfléchir ensemble aux quatre principes associés. Plusieurs propositions d'encadrement de l'IA ont été formulées par les

participants. Nous en présentons ici trois, qui permettent de suivre le cheminement des idées jusqu'à la formulation de la « Une » du Journal.

Tableau 4 : Environnement, Deuxième moment délibératif : propositions d'encadrement pour 2018-2020

Propositions d'encadrement en 2018-2020	1	2	3
Description	<ul style="list-style-type: none"> <li>Établir un code déontologique pour les concepteurs, programmeurs et gestionnaires du système (ex. pour assurer l'égalité des prescriptions)</li> </ul>	<ul style="list-style-type: none"> <li>Le protecteur du consommateur, autorité administrative indépendante, auditée par l'Assemblée nationale.</li> <li>Audits du système, de la diversité de choix et des recommandations, et publication de rapports transparents.</li> <li>Accompagnement des citoyens pour leur autonomie.</li> </ul>	<ul style="list-style-type: none"> <li>Assurer un soutien financier significatif pour aider les personnes les plus modestes à s'adapter.</li> <li>Avoir la possibilité de dépasser la cible annuelle, mais avec un coût marginal croissant sur la tonne de carbone supplémentaire.</li> </ul>
Catégories d'instruments	Loi et règlement Code de déontologie	Acteur institutionnel	Mesures incitatives et d'accompagnement

Ces propositions, qui dénotent une véritable créativité institutionnelle (au-delà des exemples d'instruments très généraux donnés dans le livret du participant) se situent dans la lignée des enjeux identifiés à l'étape précédente. La proposition d'un protecteur du consommateur (assurant une évaluation régulière du système par des audits,

rendant publiquement des comptes et organisant un accompagnement des citoyens), présente ainsi un enrichissement des idées formulées à l'étape précédente. C'est à partir de cette proposition que les participants vont bâtir leur « Une » du Journal lors de l'étape suivante.

## Troisième moment délibératif : écriture de la « Une » du Journal en 2020

Ces mesures ont ensuite été mises en récit dans l’affiche. La « Une » du Journal de l’IA responsable du 9 octobre 2020 formulée par l’équipe était la suivante :

**Grand succès pour ConsomIA !  
Un million d’abonnés en une  
semaine**

**Règlement général  
d’ouverture des *datas* pour  
l’environnement**

Suite à la loi RGODE (Règlement général d’ouverture des *datas* pour l’environnement), obligeant à rendre publiques les données personnelles, l’association CONSOM’IA a mené une enquête d’envergure sur la liberté de choix des usagers d’ECOFIT et constaté de nombreuses limitations et bulles algorithmiques. La première recommandation du rapport de CONSOM’IA est de développer des moyens de contre-expertise, et la formation des usagers pour assurer un vrai pluralisme et la participation de tous à la réduction des gaz à effet de serre.

Ainsi, si l’utilisation de l’IA présente un certain potentiel pour gérer les enjeux environnementaux associés aux comportements de consommation, cette perspective soulève également de nombreux enjeux éthiques qui doivent être encadrés convenablement.

## 2.3.

### ABORDER LES ENJEUX DE LA TRANSITION NUMÉRIQUE DANS LE MONDE DU TRAVAIL

#### Résumé du scénario de départ

**Forage des données (*Data mining*) RH pour optimiser  
l’ambiance au travail**

Pierre-André a enfin décroché un emploi dans un bon bureau d’avocats. Après trois semaines de travail, il rencontre Marco aux ressources humaines pour une séance de mentorat personnalisée. Il lui explique que la firme utilise désormais AmbIA+, une IA d’analyse conversationnelle qui étudie les attitudes des salariés et aide à maintenir une ambiance de travail apaisante et productive (tous les courriels, appels téléphoniques et prises de parole en réunion d’équipe sont analysés). AmbIA+ fournit une assistance individualisée, elle conseille et entraîne, mais il n’y a pas de sanction. Tous les échanges que Pierre-André a eus jusqu’à présent au bureau se sont bien passés, sauf les 15 et 16 octobre derniers. « Vous avez à plusieurs reprises interrompu vos collègues en réunion pour répéter les mêmes idées, ce qui a créé de la tension chez eux. Apparemment, l’algorithme a aussi détecté des périodes d’inactivité sur le réseau de plusieurs heures, sans aucun échange avec vos collègues. Ce n’est pas grave en soi, mais c’est mieux de maintenir le contact avec l’équipe. Est-ce que vous vous souvenez de la raison de cette inactivité sur le réseau ? ». Pierre-André n’est plus seulement inquiet, il est embarrassé et s’interroge sur la pertinence de ces questions.

L'objectif de ce scénario était d'entamer une discussion sur les enjeux éthiques liés à l'utilisation de l'IA dans le cadre de la surveillance et de la gestion du personnel au sein des entreprises. Notamment, le système AmbIA+ imaginé ici est utilisé afin d'optimiser la performance et de contrôler l'ambiance au travail par le biais de techniques de *data mining*.

Le parcours délibératif présenté est issu d'une journée complète de discussion regroupant 10 ingénieurs, concepteurs d'IA, gestionnaires en stratégie numérique, chercheurs, étudiants et professeurs. Partant de ce scénario en 2025, la discussion a mené à la formulation d'une « Une » du Journal de l'IA responsable du 9 octobre 2025 : « Le premier employé renvoyé à cause de l'IA ».

## Premier moment délibératif : la formulation d'enjeux éthiques et sociaux en 2025

Lors de cette première partie, les participants ont identifié cinq catégories d'enjeux en lien avec le développement de l'IA dans le monde du travail.

### AUTONOMIE

D'abord, l'enjeu du respect de l'autonomie (notamment en ce qui a trait à la capacité d'agir des employés) a été souligné. Les participants ont dénoncé une certaine manipulation de « la manière dont les gens ressentent les choses », et une culture organisationnelle « forcée ». Les participants ont considéré que la sauvegarde d'informations, telles que les comportements et interactions entre employés dans le but de cultiver une culture organisationnelle, est problématique. En fonctionnant ainsi, l'entreprise réalise une certaine normalisation des employés, qui pourrait conduire à de fortes tensions (voire « un totalitarisme » si tout venait à être mesuré afin d'assurer un contrôle de l'entreprise sur les individus), par le biais de recommandations insistantes de l'IA. Le respect de l'autonomie est lié au respect d'un certain « libre arbitre » des employés

et du respect de leurs émotions qui, ici, ne sont pas suffisamment considérées selon les citoyens. Est-il bien de conserver toutes ces données (la moindre action étant observée) et de les utiliser dans ce but ? Les participants ont ainsi souligné un enjeu lié à la « surveillance », qui pourrait alors limiter le champ d'action et la liberté de parole des employés (très lié à l'enjeu du respect de la vie privée).

### « L'IA t'observe »

Un participant

Les citoyens ont mis de l'avant la nécessité de favoriser l'autonomie en permettant à chacun de travailler à sa façon pour le meilleur de l'entreprise. L'employé devrait être capable de garder le contrôle sur ses données et l'employeur devrait l'informer de la collecte et de l'utilisation de ces dernières avec précision, et chacun devrait être libre de pouvoir se « déconnecter », notamment afin de préserver la frontière entre ce qui relève du professionnel et du privé.

### VIE PRIVÉE

La place de cette frontière entre vie privée et professionnelle au sein de l'entreprise a été débattue par les participants et considérée comme floue. D'un côté, pour certains, le système d'IA en jeu porterait atteinte à la vie privée des employés, en écoutant leurs conversations :

### « Qu'en est-il des discussions personnelles ? »

Un participant

De l'autre, certains participants ont mentionné qu'habituellement, tout ce qui se rattache à la vie personnelle ne devrait pas se faire ni par le courriel ni par le téléphone du travail (et donc ne serait pas dans la mire de l'analyse du SIA présenté dans le scénario). Ces participants ont soulevé la question suivante : Y a-t-il une place pour la vie privée au sein de l'entreprise ?

Un consensus a cependant été observé concernant l'usage de l'IA : il doit s'agir d'un outil pour l'entreprise qui ne doit en aucun cas analyser ce qui relève de la sphère privée. L'enjeu est alors de

délimiter cette frontière afin d'identifier clairement quand l'IA peut être utilisée ou non (il s'agit de définir quelles sont les données qui relèvent uniquement de la sphère professionnelle et peuvent servir son analyse).

L'analyse comportementale réalisée par le SIA a également été dénoncée. Celle-ci pourrait porter atteinte tant à la vie privée qu'à l'autonomie (à savoir, les citoyens ont questionné par la suite s'il était éthique d'utiliser une IA pour retracer des conversations et comportements, même s'ils relèvent de la sphère professionnelle). Les citoyens dénoncent ici une forme d'intrusion et un bris de confidentialité issus de cette surveillance constante :

**« C'est comme le Big Brother qui nous surveille »**

Un participant

Certains des participants ont souligné que ces enjeux n'étaient pas spécifiques à l'IA, tandis que d'autres considèrent que la haute traçabilité permise par l'IA les potentialise.

## BIEN-ÊTRE

L'impact de ce genre de système sur le bien-être des employés a été considéré à la fois comme positif et négatif. En effet, si cette technologie peut être utilisée pour aider l'employé et améliorer la qualité de sa vie professionnelle (en participant à apaiser les relations, à déceler les cas de harcèlement ou d'intimidation ou encore en participant à la prévention du suicide), il semble qu'elle puisse tout autant lui nuire (le « déstabiliser », le rendre « inquiet », « normaliser » ses comportements). Les participants s'entendent sur un excès de sollicitation de l'employé dans le scénario, qui ne devrait pas avoir à justifier tous ses agissements. Les citoyens soulèvent ici un dilemme entre le bien-être et la liberté des employés. Jusqu'à quel point est-il possible de surveiller leurs gestes dans la perspective de protéger leur bien-être sans nuire indûment à leur liberté d'agir ? Qu'est-ce qu'une surveillance « bien utilisée » ? Les citoyens ont également souligné une corrélation entre le bien-être et les objectifs de performance : un employé

heureux est également plus productif. Protéger le bien-être semble donc aller dans le sens des intérêts de l'entreprise comme de l'individu.

## TRANSPARENCE

Les citoyens ont d'abord soulevé le non-respect du consentement de l'employé, qui ignore qu'il est « surveillé », et en appelle à plus de transparence de l'employeur qui devrait l'informer sur ce qui est enregistré ou non (notamment, dans le cadre de la formation des employés).

**« Le consentement de l'employé est important dans le recueil de données. Il faut être transparent envers ses employés. »**

Un participant

Cet enjeu de transparence a soulevé de nombreuses questions : Quelles sont les règles des rapports relationnels dans l'entreprise ? Quelle est la hiérarchie dans l'importance des données ? Qui peut les voir ? Un employé a-t-il le droit de voir les données de son patron ?

L'enjeu de transparence a également trait à l'IA ou au comment la rendre interprétable. Les citoyens ont soulevé ici la nécessité de communiquer autour du fonctionnement de l'algorithme afin notamment d'assurer une « adhésion » des individus à ces nouveaux systèmes. Ils ont également souligné l'importance de ne pas prendre une décision sur la base de la conclusion d'une technologie d'IA non explicable.

## PRODUCTIVITÉ ET PERFORMANCE

L'enjeu de la performance et de son évaluation a ensuite été soulevé à plusieurs reprises. À quel point l'IA doit intervenir dans le but d'augmenter la productivité de l'employé (ex. en l'empêchant de se répéter dans les discussions de groupe) ? Est-ce essentiel ? Il serait ici nécessaire de définir plus précisément comment l'IA pourrait mesurer et améliorer la performance. D'une part, le risque de valoriser la productivité au détriment de l'employé,

de son développement personnel et de son comportement au travail a été soulevé. L'employé a-t-il vraiment une chance de se perfectionner dans ce contexte ? Faut-il personnaliser les attentes et les objectifs selon les personnes ?

## « L'IA fait oublier la personne »

### Un participant

D'autre part, la définition même de ce qui doit être considéré comme productif (et contre-productif) s'avère ici essentielle pour les participants. Quels indicateurs utiliser ? Quelle est leur pertinence ? Qui peut et doit les définir ? Doivent-ils être en lien avec les objectifs de l'entreprise ? Ces objectifs doivent-ils être révisés à la lumière des analyses de l'IA ? Bien qu'elle soit utilisée dans l'optique de vérifier que les employés « performant », l'IA, en interprétant les résultats des employés, pourrait tout aussi bien s'avérer être un frein à l'innovation, notamment parce que certaines tâches sont plus faciles à mesurer que d'autres.

L'ensemble des débats concernant la définition de la performance, de la productivité et du respect du bien-être des employés a mené à la conclusion que ces enjeux sont très liés aux cultures des entreprises, qui peuvent être très différentes et refléter des objectifs, des intérêts et des valeurs variés.

Certains citoyens autour de la table s'inquiètent que l'instauration d'outils basés sur l'IA pousse les sociétés à imposer une performance systématiquement associée à un « score », qui pourrait se traduire par une normalisation. D'autres mentionnent que ces pratiques existent déjà, mais qu'elles seront seulement amplifiées, voire « industrialisées » par les systèmes d'IA qui peuvent traiter un plus grand nombre d'informations beaucoup plus rapidement. D'autres encore ont mentionné qu'une uniformisation des pratiques peut tout de même être utile, notamment pour répondre aux besoins de l'entreprise.

Un débat sur la prétendue objectivité de l'IA a ensuite eu lieu, soulevant des questions telles que : Comment connaître les critères de décision et d'automatisation ? De quelle manière l'IA va définir une absence d'activité de l'employé ? Comment équilibrer d'un côté la performance, l'objectivité et la standardisation par les technologies d'IA et de l'autre la subjectivité, la particularité et l'arbitrage humain ?

Après plus d'une heure de discussion, les participants ont sélectionné parmi ces cinq enjeux les trois qui semblaient prioritaires pour 2025. Ces enjeux font parfois écho aux principes de la version préliminaire de la Déclaration.

Tableau 5 : Les enjeux prioritaires

Enjeux éthiques 2025	1	2	3
Description	Comment encadrer l'évaluation de la performance en respectant à la fois les objectifs de l'entreprise (productivité) et l'individu (normalisation) ?	Comment faire en sorte qu'entreprise et employés comprennent l'IA (et garantir l'adhésion) ?	Comment l'IA peut préserver (contribuer, soutenir) le bien-être des employés.
Principes associés	Performance	Transparence (connaissance)	Bien-être

## Deuxième moment délibératif : proposition d'encadrement de l'IA pour 2018-2020

Lors de cette deuxième partie de l'activité, les participants ont été sollicités pour formuler des recommandations et imaginer quelles solutions pourraient être mises en place pour répondre à ces trois enjeux. Une grande diversité de mécanismes plus ou moins contraignants a été formulée pour chacun des enjeux identifiés, pour finalement converger en six principaux mécanismes à mettre en place pour tenir compte de l'ensemble des enjeux.

Pour répondre à l'enjeu de **performance**, les participants ont d'abord recommandé d'organiser des **formations continues** dans les entreprises afin d'accompagner les gens dans chaque étape de la « transformation numérique ». L'**éducation** au numérique permettrait d'encourager une ambiance d'apprentissage, mais également de diminuer la peur qui peut accompagner le développement de l'IA dans le monde du travail.

« Des formations continues à chaque étape de la transformation numérique d'une entreprise pour tous les salariés (encourager une ambiance d'apprentissage constant). »

Un participant

Également, la mise en place d'une **autorité administrative indépendante** (AAI) et d'un **correspondant** au sein de l'entreprise a été envisagée comme pistes de solution<sup>2</sup>, afin notamment de garantir le respect du dispositif RGPD (qui devrait être étendu à la traçabilité des données et l'explicabilité des décisions algorithmiques). Ce correspondant serait responsable de l'application des règlements, de soutenir le plaignant lors de problème et pourrait faire appel à l'AAI si besoin. Les participants ont également proposé la création d'**indicateurs** en lien direct avec non seulement les objectifs de l'entreprise (ex. résultats financiers), mais aussi avec le respect de certaines « valeurs

humaines » (ex. bien-être du salarié). Pour cela, la mise en place d'une loi est absolument nécessaire selon les citoyens, si l'on ne veut pas que ce score soit uniquement corrélé aux intérêts financiers de l'entreprise.

Il a également été recommandé que le gouvernement crée un **programme de recherche publique** sur l'*interprétabilité* et la transparence des algorithmes (pour rattraper le retard sur la recherche privée à ce niveau), afin de comprendre comment ces décisions sont prises et de limiter le monopole des grosses entreprises d'IA. Dans la même optique, la création de « **groupes pilotes** » (ou « groupes tests ») afin de mesurer les impacts (incluant les impacts psychologiques et sociaux) de l'utilisation de l'IA au sein des entreprises et en vérifier la pertinence et l'utilité devraient être mis en place.

Pour répondre à l'enjeu de **transparence**, certains participants ont défendu la mise en place de mesures moins contraignantes, comme une **communication obligatoire** sur les différentes règles suivies, les données des salariées qui sont utilisées, les objectifs de leur collecte (scoring individuel ou collectif ?) et ce qu'on est capable de déduire de ces données, ou encore les résultats des analyses faites sur les groupes pilotes mentionnés précédemment. Cette communication doit être réalisée auprès de tous les départements (RH, informatique, marketing, juridique), et inclure des notions sur l'IA, une formation sur ce qu'est un algorithme et comment il apprend à partir des données.

Pour protéger les enjeux de **bien-être**, les participants ont recommandé **une évaluation annuelle** de la perception des employés en lien avec l'usage des dispositifs d'IA. Cette évaluation pourrait éventuellement être consolidée par un comité (par exemple le CHSCT), responsable d'agir en cas de problèmes. Celle-ci devrait favoriser les bonnes relations entre salariés et employeurs, et assurer que le mode de travail de chacun est respecté même avec l'avènement de la technologie. La mise en place d'**une certification** (ou label) qui serait garante de bonnes pratiques au point de vue éthique, environnemental et sociétal devrait être développée et imposée par l'État. Cette certification permettrait de garantir le respect d'un minimum de critères au sein des entreprises, afin d'optimiser la productivité

<sup>2</sup> Règlement général sur la protection des données, réglementation européenne entrée en vigueur le 25 mai 2018.

et la performance. Les **indicateurs** de bien-être, au même titre que ceux de la performance, devraient être pris en compte.

En réponse aux enjeux de bien-être et de transparence, les citoyens ont proposé la mise en place d'une loi qui doit définir et imposer l'*interprétabilité* des IA (notamment, la justification de la décision et un accès garanti à des règles explicites). Cette loi devrait dicter les critères minimums pour protéger le bien-être des individus (incluant un droit à la déconnexion) afin de garantir la protection des droits fondamentaux qui pourraient être menacés par le développement de l'IA.

Pour tous ces enjeux, et dans l'optique de « réglementer sans pénaliser », devrait être créée une **charte**<sup>3</sup> officielle qui reprend les droits, les devoirs et les valeurs à défendre pour protéger l'individu au sein de l'entreprise.

Au final, **6 recommandations** reprenant les grandes lignes des propositions précédentes ont été formulées de manière consensuelle :

Tableau 6 : Les propositions retenues

Propositions d'encadrement	1	2	3	4	5	6
<b>Description</b>	Loi qui doit définir et imposer l' <i>interprétabilité</i> des IA et qui définit les critères minimums pour protéger le bien-être des individus.	Certification (ou label).	Formation pour les différents acteurs de l'entreprise.	La mise à jour du Code du travail afin de l'adapter à la réalité numérique.	La création d'une charte des droits et devoirs.	Le financement de recherches publiques et d'études pilotes sur l'IA et ses impacts sur le monde du travail.
<b>Enjeux associés</b>	Les 6 recommandations ont été formulées pour que chacune réponde aux 3 enjeux prioritaires					

<sup>3</sup> Les participants ont souligné qu'une charte n'a pas la même force contraignante en France qu'au Québec.

## Troisième moment délibératif : écriture de la « Une » d'un journal en 2020.

Cette étape est la mise en récit d'une des pistes de solution proposées en 2020. Les participants ont ici mis en scène un des risques de l'utilisation de l'IA au sein des entreprises et une des mesures envisagées pour y répondre.

### Le premier employé renvoyé à cause de l'IA

Un salarié se fait renvoyer après trois semaines de travail sur les recommandations d'une IA. Il a contesté son renvoi auprès des Prud'hommes. Le jugement a été rendu après le débat à l'Assemblée. Une loi encadrant cette pratique va être votée.

Les participants ont tenu à souligner qu'ils ont surtout discuté des impacts potentiellement négatifs de l'IA en milieu de travail. Cependant, ils ont reconnu que l'IA pourrait également apporter de nombreux avantages tant pour les employés que les entreprises, avantages qui pourraient faire l'objet d'une autre discussion. La mise en place d'un organisme interne appuyé par un mécanisme législatif semble être la solution indispensable au développement responsable de l'IA au sein des entreprises.

### 3. DISCUSSION AUTOUR DU THÈME DE LA CULTURE avec les membres de la Coalition de la diversité des expressions culturelles (CDEC)

Afin d'aborder les enjeux relatifs au développement de l'IA dans le domaine de l'art et de la culture, un atelier de discussion a été organisé avec la Coalition de la diversité des expressions culturelles (CDEC) le 25 septembre 2018, réunissant près de 11 experts et parties prenantes du domaine. La discussion avec les participants s'est déroulée successivement autour de trois thématiques :

1. Les droits d'auteur ;
2. L'enjeu de la diversité culturelle ;
3. La propagande et la manipulation.

Suite à ces discussions, la CDEC a produit un mémoire<sup>4</sup> particulièrement pertinent qui fait état des différents défis et opportunités liés au développement de l'IA en culture. Ce mémoire présente également les principes éthiques nécessaires au développement responsable de l'IA dans ce domaine et les principales recommandations qui ressortent de la rencontre du 25 septembre. La présente section résume ces discussions et les points principaux de ce mémoire.

#### 3.1.

### TROIS THÉMATIQUES PROPOSÉES PAR L'ÉQUIPE DE LA DÉCLARATION POUR ABORDER LES ENJEUX DU DÉVELOPPEMENT DE L'IA DANS LE DOMAINE DE LA CULTURE

#### LES DROITS D'AUTEUR DANS UN CONTEXTE DE COCRÉATION PAR LES IA

Le recours à des systèmes d'IA pour générer des œuvres à très bas prix (ex. par les *Generative Adversarial Network*) en musique, arts visuels, séries télévisuelles, ou pour la rédaction d'articles de journaux, va reposer de manière inédite la question des droits d'auteurs. Doivent-ils revenir aux auteurs qui ont créé les exemples sur lesquels apprennent les algorithmes, ou bien aux informaticiens programmeurs d'algorithme, ou au promoteur du projet ? Le pastiche ou le *remix* produit par un algorithme génératif est-il du plagiat ? Si l'IA remplace l'artiste, quel sera l'effet sur la diversité culturelle ? Et si une IA « interprète » une œuvre ? Quel est l'accès au patrimoine artistique par les algorithmes ?

Cette problématique concerne surtout les SIA générateurs de création (voir les applications de l'IA dans le domaine culturel, Mémoire de la CDEC, p. 2).

#### L'ENJEU DE LA DIVERSITÉ CULTURELLE FACE AUX NOUVEAUX DISPOSITIFS DE RECOMMANDATION PAR ALGORITHME

Face aux risques d'uniformisation des goûts et des conduites, de « bulle algorithmique », de capture de l'attention et de formatage des choix par les algorithmes de recommandation, comment maintenir la diversité de l'offre culturelle ? Quelles pratiques de réception libre, autonome et critique promouvoir auprès des usagers ? Quelles possibilités de déconnexion (objets connectés, robots

<sup>4</sup> Pour plus d'informations sur les enjeux du développement de l'IA sur la diversité culturelle, consulter le mémoire : <https://cdec-cdce.org/des-principes-ethiques-pour-un-developpement-de-lintelligence-artificielle-misant-sur-la-diversite-des-expressions-culturelles/>

domestiques) face aux stratégies de capture de l'attention ? Quelles transparence et explicabilité pour les usagers ? Une politique culturelle publique devrait-elle être proposée sur cet enjeu de la diversité ? Avec quels nouveaux mécanismes de financement de la diversité culturelle ?

Cette problématique concerne surtout les algorithmes de recommandation et la valorisation des données (voir les applications de l'IA dans le domaine culturel, Mémoire de la CDEC, p. 2).

### LA CENSURE ALGORITHMIQUE ET L'ART ACTUEL

Les algorithmes de reconnaissance sont utilisés par les médias sociaux pour exercer un pouvoir de censure parfois jugé excessif. Des artistes contemporains répondent à ces dispositifs par des interventions pour à la fois rendre visibles ces règles et les détourner. Quelle liberté critique pour les artistes contemporains de demain ? Quelle formation des artistes pour leur donner une connaissance critique ?

Cette problématique concerne surtout les algorithmes de recommandation et la valorisation des données (voir les applications de l'IA dans le domaine culturel, Mémoire de la CDEC, p. 2).

## 3.2

### LES ENJEUX DE LA PROMOTION DE LA DIVERSITÉ CULTURELLE À L'HEURE DE L'IA

#### DÉCOUVRABILITÉ ET HOMOGENÉISATION DES CONTENUS CULTURELS

*« Lorsque les recommandations sont basées, entre autres, sur la popularité des contenus, elles contribuent fortement à la concentration des écoutes (0.7 % des titres représentent 87 % des écoutes sur les services de musique en ligne au Canada), favorisant une minorité d'artistes »* (Mémoire de la CDEC, p. 4).

La **découvrabilité** a été identifiée comme un enjeu important de l'avènement de l'IA dans le domaine de la production culturelle. D'un côté, s'ils sont correctement paramétrés, les SIA pourraient devenir des outils de diversité culturelle, en élargissant par exemple les auditoires mondiaux. Ils pourraient proposer une diversité de contenu relativement élevée, en permettant par exemple à des artistes de se faire connaître sans nécessairement passer par des intermédiaires, voire sans coût de production. D'un autre côté, les participants s'inquiètent d'un risque d'uniformisation des contenus culturels et de la réelle neutralité du web. Par exemple, les algorithmes de recommandations proposent rarement du contenu québécois sur les différentes plateformes culturelles comme Netflix, Amazon, ou Spotify. Ces plateformes comportent un très grand nombre de contenus, si bien que les utilisateurs s'en remettent souvent à l'algorithme de recommandation pour faire leur choix. Les participants soulèvent le risque que les algorithmes finissent par enfermer les individus « dans un goût », empêchant de découvrir d'autres contenus que ceux recommandés sur la base de leur choix précédent. Ceci risque d'entraîner une homogénéisation des contenus culturels,

notamment parce que les créateurs, pour la plupart, adaptent leurs œuvres à ce mode de diffusion.

Les participants soulignent qu'il est nécessaire que le milieu audiovisuel fasse des efforts de découvrabilité dans les plateformes numériques. Le problème ne vient pas seulement des SIA, c'est aussi un enjeu lié aux objectifs de l'industrie. Si les gouvernements ont une responsabilité concernant la diversité culturelle, les multinationales viennent toutefois défier leur pouvoir. Le Canada a par exemple été le premier signataire de la Convention de l'UNESCO pour la protection et la promotion de la diversité des expressions culturelles. Bien que certains contenus soient financés publiquement, le modèle dominant reste un modèle commercial, ce qui peut être problématique, comme lorsqu'on constate que les livres québécois ne sont pas mis de l'avant par l'algorithme d'Amazon (sur Amazon.ca), et ce malgré la volonté provinciale.

Les participants ont ainsi insisté sur le développement d'une démocratisation du pouvoir de créer et d'utiliser les SIA en ce sens, qui deviennent alors des agents importants de l'écosystème culturel en ce qui a trait à l'émergence des créations.

## LES CONSÉQUENCES SUR L'EMPLOI

**« En mai 2018, des chercheurs révélaient les résultats d'une enquête menée auprès de plus de 350 chercheurs en IA. En moyenne, ils prédisent que l'IA pourra dépasser les humains pour produire des essais de niveau scolaire en 2026, des chansons populaires en 2028 et des best-sellers en 2049. »**

(Rapport de la CDEC, p. 3)

La possibilité pour une IA de générer une œuvre à part entière sans impliquer d'artiste est un sujet d'inquiétude. Par qui ces œuvres seront-elles développées ? Uniquement par des entreprises qui possèdent déjà un certain capital ? Est-ce compatible avec le développement d'une société

composée de plus d'artistes et plus de diversité ? Les participants ont ici dénoncé l'existence d'un risque de boucle fermée (notamment en lien avec les travaux sur la directive européenne à ce sujet), où seulement une infime minorité d'artistes serait favorisée par les algorithmes. Ils craignent que le nombre d'artistes décline de façon importante.

Comment les créateurs humains pourraient-ils se démarquer face aux créations logicielles ? Comment cela va-t-il affecter les capacités individuelles et collectives de création et d'innovation ? Dans une expérimentation conduite chez l'ACTRA (Alliance of Canadian Cinema, Television and Radio Artists), il n'a pas été possible de faire la différence entre les personnages créés par une IA et ceux de vrais humains. Cela semble aller au-delà des enjeux du *DeepFake*, puisqu'il s'agit de la création de nouveaux visages. Or c'est un des rôles d'un grand nombre de leurs membres, qui travaillent dans le jeu vidéo. L'enjeu soulevé ici est celui de la perte d'emploi que ces technologies pourraient engendrer.

La question de la rémunération a également été soulevée, ouvrant la voie à une discussion sur les droits d'auteur. Qui rémunérer et comment, si des algorithmes sont à l'origine de la création ? La rémunération des acteurs se calcule selon des journées de travail et les droits de suite. Au niveau des droits de suite, des compromis sont envisageables, comme des normes pour réguler la réutilisation des œuvres. Les participants ont également souhaité tempérer l'impact potentiel de l'IA sur la reconnaissance de l'artiste, qui pourrait être plus limité que prévu, comme pour le livre numérique qui n'a pas tué le livre papier.

## REPENSER LES DROITS D'AUTEUR

**« Évidemment, la dématérialisation des contenus culturels, les changements technologiques et l'arrivée de nouveaux joueurs qui ont transformé les modèles d'affaires ont un impact majeur sur la rémunération des artistes,**

## et le paiement de redevances de droits d'auteur. »

(Mémoire de la CDEC, p. 5)

Différentes questions ont été soulevées à ce sujet : Quelles seront les redevances ? Quelles seront les réglementations concernant les droits d'auteur ? Quelle portion d'un manuscrit est nécessaire à un algorithme pour en produire un autre ? De quelle manière l'IA va pouvoir se baser sur d'autres livres pour en commercialiser de nouveaux ? L'avènement de SIA créateurs soulèvent ainsi d'importants enjeux de propriété des contenus culturels. Concernant la musique, la grande puissance de calcul et le plus grand nombre de données disponibles, combinées aux avancées majeures concernant l'efficacité des algorithmes, a permis la création d'outils de « génération » artistique. En musique, généralement, l'apprentissage se fait cependant déjà en reprenant ce qui a été fait. Si l'IA fait de même, c'est cependant à une vitesse incomparable.

Les participants ont ainsi reconnu que l'IA bafoue déjà les droits d'auteurs et ont remis en question la possibilité d'accorder ce genre de droits à une machine. Considérant que les satires et les parodies font déjà partie d'exceptions au droit d'auteur, serait-il possible de faire également exception pour l'IA ? La loi concernant les droits d'auteur est actuellement en révision. Il semble ici essentiel pour les participants qu'elle intègre les évolutions liées aux SIA dans sa nouvelle forme.

## 3.3

### LES PRINCIPES ÉTHIQUES DE LA CDEC

Dans son mémoire, la CDEC recommande l'adoption de quatre principes éthiques pour prévenir les dérives du développement de l'IA dans le milieu culturel. Le respect de ces principes devrait permettre, selon eux, que l'IA intègre mieux les enjeux culturels en général et ceux de la diversité en particulier. Ces quatre principes sont cohérents avec ceux développés dans la version finale de la Déclaration de Montréal, mais intègrent les spécificités du domaine culturel et du domaine des arts.

La CDEC propose **un principe de diversité des expressions culturelles** qui fait directement écho au **principe d'inclusion de la diversité**. Cependant, la CDEC précise ici que ce principe devrait faire en sorte que les SIA « valorisent les contenus culturels et linguistiques locaux au sein des populations dont ils sont issus, favorisant ainsi la cohésion sociale tout comme le tissu économique local ; incitent les usagers à faire des découvertes à l'extérieur de leur univers ; facilitent le passage entre les familles technologiques (ex. Apple), plutôt que de les y enfermer ; favorisent les interactions et le partage des contenus. » (Mémoire de la CDEC, p. 6)

La CDEC propose également **un principe de valorisation de la culture, des artistes et des producteurs** de contenus culturels, soit que la contribution de l'IA permette d'éviter l'actuelle dévalorisation des contenus culturels et qu'ils ne « favorisent l'appropriation démesurée des revenus qui devraient être destinés aux écosystèmes culturels. » (Mémoire de la CDEC, p. 6). Si ce principe est en lien avec plusieurs des principes de la Déclaration, il appelle surtout à respecter le 6<sup>e</sup> principe d'équité et les sous-principes associés : *Le développement et l'utilisation des SIA doivent contribuer à la réalisation d'une société juste et équitable.*

La CDEC propose ensuite un **principe de transparence** (au niveau du code des algorithmes, mais aussi des données prises en compte) **et de dialogue** (notamment, avec les utilisateurs). Ce principe appelle à respecter le 5<sup>e</sup> principe de la Déclaration, celui de participation démocratique : Les SIA doivent satisfaire les critères d’intelligibilité, de justifiabilité et d’accessibilité, et doivent pouvoir être soumis à un examen, un débat et un contrôle démocratiques. De même pour le 2<sup>e</sup> principe de la Déclaration, celui d’autonomie : *Les SIA doivent être développés et utilisés dans le respect de l’autonomie des personnes et dans le but d’accroître le contrôle des individus sur leur vie et leur environnement.*

Enfin, la CDEC propose un principe de **primauté de l’intérêt public**, défini comme suit : *Toutes les innovations technologiques ne sont pas désirables. Le développement de l’IA devrait toujours privilégier*

*l’amélioration de la qualité de vie des populations, de la cohésion sociale et des pratiques démocratiques. Les gouvernements doivent défendre l’intérêt public face à des développements qui pourraient avoir des impacts plutôt négatifs sur la société.* (Mémoire de la CDEC, p .7). Le respect de ce principe s’accorde avec le respect du 1<sup>er</sup> principe de la Déclaration, le principe de **bien-être** : *Le développement et l’utilisation des systèmes d’intelligence artificielle (SIA) doivent permettre d’accroître le bien-être de tous les êtres sensibles.* Et également avec le 8<sup>e</sup> principe de la Déclaration, celui de **prudence** : *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.*

Principes éthiques de la CDEC	Principes de la Déclaration de Montréal
Principe de diversité des expressions culturelles	Principe 7 d’inclusion de la diversité
Principe de valorisation de la culture, des artistes et des producteurs de contenus culturels	Principe 6 d’équité
Principe de transparence et dialogue	Principe 2 du respect de l’autonomie Principe 5 de participation démocratique
Principe de primauté de l’intérêt public	Principe 1 de bien-être durable Principe 8 de prudence

## 3.4

### QUELQUES-UNES DES RECOMMANDATIONS FORMULÉES

Différentes recommandations sont ressorties des discussions du 25 septembre. Elles ont été formulées dans le but de promouvoir le contenu culturel québécois et conscientiser les citoyens aux impacts du développement de l'IA sur la culture. D'abord, pour favoriser la diversité de l'expression culturelle dans l'univers numérique, les participants recommandent la définition d'exigences minimales de représentation de contenus culturels canadiens dans les recommandations issues d'algorithmes, comme c'est déjà le cas pour la télévision ou la radio québécoise. Ces exigences doivent être formulées dans des lois et réglementations, car elles ne pourront, selon les participants, émaner du libre marché.

Lors des discussions, les participants ont reconnu que la maîtrise de la littératie relative au développement de l'IA est essentielle. Il est nécessaire d'outiller les individus de manière à ce que chacun puisse comprendre où les recommandations des algorithmes les amènent. Les participants recommandent la mise en place d'une politique d'éducation des usagers, qui aurait pour but de contrer la fausse impression de choix, en les incitant à varier leur exploration et rester conscients de l'influence des algorithmes. Cette politique irait dans le sens d'une éducation au choix critique. Ainsi, une forme d'autodéfense intellectuelle devrait être développée pour tous, et ce dès l'enfance. Les participants recommandent également de sensibiliser les développeurs informatiques aux impacts des SIA sur la culture.

Lors de ces discussions est également ressortie une recommandation concernant la transparence et l'explicabilité des recommandations algorithmiques. Les usagers devraient être systématiquement informés lorsqu'une recommandation émane d'un SIA et devraient avoir accès facilement

à de l'information explicative tant sur le fonctionnement des algorithmes que sur l'existence d'autres contenus culturels.

Les participants ont également recommandé que les entreprises qui développent des SIA ayant un impact sur la culture consacrent une part de leur chiffre d'affaires à la promotion de la diversité culturelle, en finançant par exemple, certaines bibliothèques, événements culturels, ou médias. Les participants prônent également la mise en place d'une surveillance sur le profilage de goût et sur la protection des renseignements personnels. Ils suggèrent aussi le développement d'un « laboratoire de l'IA en culture » afin d'observer les algorithmes, d'apprendre à interagir avec eux et éventuellement d'influencer leur développement.

Deux principales recommandations issues de ces discussions sont particulièrement développées dans le mémoire de la CDEC :

1. **Éducation et formation**
2. **Révision des lois touchant le milieu culturel**

Ces recommandations s'inscrivent dans la lignée de celles formulées pour les autres secteurs lors de la coconstruction de l'hiver, à savoir 26 % des recommandations comptabilisées sont des dispositions légales et juridiques et 19 % des formations (*cf. Rapport des résultats des ateliers de coconstruction de l'hiver*).

## 4. FAIRE LE PONT ENTRE LES DÉLIBÉRATIONS CITOYENNES ET LA RELÈVE EN RECHERCHE : SIMULATION DE LA RÉDACTION DE BRÈVES POLITIQUES

### 4.1.

#### DESCRIPTION DE L'ACTIVITÉ

Afin de faire le pont entre la relève en recherche et les citoyens, la Déclaration a participé à l'organisation d'une activité de simulation en partenariat avec le Comité intersectoriel étudiant (CIÉ) des Fonds de recherche du Québec (FRQ) et l'École de politique appliquée (EPA) de l'Université de Sherbrooke dans le cadre des Journées de la relève en recherche (J2R) organisées par l'Acfas. L'activité de simulation « Politique et intelligence artificielle » avait pour but de réunir des étudiants de la relève en recherche dans l'optique de produire trois brèves politiques sur l'IA. L'objectif de cette activité était de permettre à la relève de prendre part aux discussions relatives à l'IA et aux enjeux éthiques et sociaux de son développement. Ce thème s'est avéré fédérateur pour le CIÉ :

« L'IA a été choisie car elle présente des dimensions intersectorielles et englobe des enjeux intéressants d'un point de vue de l'alliage entre science, société et développement de politiques publiques. Dans le

cadre de cette simulation, l'IA permet à la relève de prendre part aux réflexions et au positionnement du Québec comme un leader en cette matière. »

(Guide à l'intention des participants, p. 5).

Dans cette optique, la Déclaration de Montréal a fourni trois problématiques issues des discussions de la coconstruction citoyenne réalisée lors de l'hiver 2018. Il nous a semblé pertinent, tant pour l'élaboration des travaux de la Déclaration, que pour le travail des jeunes chercheurs sélectionnés pour l'activité, de débattre et émettre des recommandations autour de ces thèmes portés par les délibérations. Ces trois problématiques touchent en effet à des enjeux particulièrement sensibles du développement de l'IA sur lesquels il est urgent de délibérer :

1. La sécurité et l'intégrité des systèmes d'IA : Comment maximiser les impacts positifs tout en minimisant les effets néfastes du développement de l'IA ?
2. L'IA, les médias et la manipulation de l'information : Comment lutter contre la diffusion et l'amplification de fausses nouvelles et de campagnes de désinformation ? Comment favoriser la démocratisation de l'accès à l'information tout en encourageant la formation d'une pensée critique et la prise de décision éclairée ?
3. La gouvernance publique, privée et participative ; les communs numériques : Laquelle de ces gouvernances serait la plus appropriée ? Quelles balises mettre en place ?

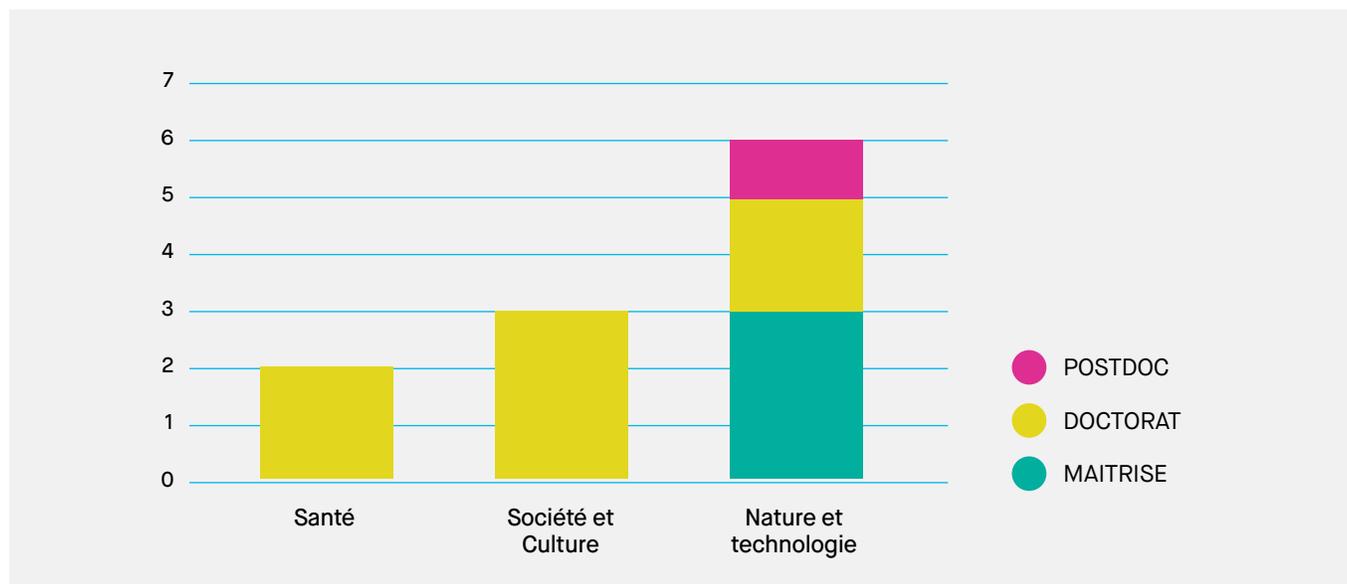
La simulation s'est déroulée en vue de trois objectifs : « 1. Familiariser les participants à l'écriture et à la présentation d'une brève politique ; 2. Faciliter l'acquisition de compétences liées au développement de politiques scientifiques ; 3. Analyser une problématique sociale avec un regard scientifique. » (Guide à l'intention des participants, CIÉ, p. 6). Les brèves politiques ont ainsi été élaborées dans le cadre d'un exercice se voulant avant tout pédagogique : la diffusion

auprès de décideurs ou parties prenantes n'était pas l'objectif de l'activité. Toutefois, les recommandations présentées dans ces brèves nous donnent accès à la perspective des étudiants de la relève ayant participé, et elles ont été particulièrement pertinentes. Les brèves sont ainsi annexées au présent rapport bien qu'elles ne soient pas porteuses de recommandations concrètes en vue de politiques publiques. Nous avons décidé de

présenter les brèves telles que rédigées par les étudiants, sans révision, afin de ne pas travestir leurs productions et de représenter au mieux leur contribution (cf. Annexe 2).

L'activité s'est déroulée les 18 et 19 octobre 2018 à l'Université de Sherbrooke et a réuni 11 étudiants aux expertises et aux niveaux variés.

**Graphique 5 : Profils des étudiants participants à l'activité en fonction de leur domaine d'étude (selon les secteurs des trois Fonds)**



Les étudiants ont été répartis en trois groupes, chacun traitant d'une des problématiques et accompagné d'un facilitateur (indépendant de l'équipe de la Déclaration afin de ne pas orienter les recommandations). Après quatre présentations introductives sur l'IA et la rédaction de brèves politiques ; les étudiants ont disposé de six heures

seulement pour rédiger les brèves et préparer une présentation orale de leur travail. Les trois groupes ont présenté leurs brèves devant un jury<sup>5</sup> le 19 octobre au matin. L'équipe 2 (traitant de la problématique L'IA, les médias et la manipulation de l'information) a remporté le concours.

<sup>5</sup> Le jury était composé de trois membres :

**Claude Asselin**, professeur, Département d'anatomie et de biologie cellulaire, Faculté de médecine et des sciences de la santé, Université de Sherbrooke. [Représentant de l'Acfas].

**Benoit Sévigny**, directeur du Service des communications et de la mobilisation des connaissances. [Représentant FRQ].

**Nathalie Voarino**, candidate au doctorat en bioéthique, coordonatrice scientifique de la Déclaration de Montréal. [Représentante de la Déclaration]

## 4.2

# LES PROBLÉMATIQUES ISSUES DES PRÉOCCUPATIONS CITOYENNES

## Problématique 1 : Sécurité publique et intégrité des systèmes

Lors des consultations, les citoyens ont reconnu que l'essor de l'IA pourrait contribuer à rendre nos environnements physique et numérique plus sécuritaires. Par exemple, dans une ville intelligente, l'autonomisation des transports pourrait réduire le taux d'accidents routiers ; en santé publique des modèles épidémiologiques pourraient permettre aux autorités de mieux prévoir la propagation des maladies ; en matière de cybersécurité, des spécialistes en sécurité informatique se tournent vers l'intelligence artificielle pour reconnaître les agressions.

Cependant, les citoyens ont reconnu que certaines conditions sont nécessaires pour s'assurer que les avancées de l'IA soient bénéfiques pour la sécurité publique. Garantir un « bon » usage de l'IA par le biais de l'intégrité et de la sécurité des systèmes sont des enjeux fondamentaux du développement responsable de ces technologies.

Les impacts négatifs de l'IA sur la sécurité publique peuvent prendre quatre formes :

1. **Une IA conçue dans le but de menacer la sécurité publique<sup>6</sup>.** Par exemple, l'utilisation de l'IA dans des buts de cybercriminalité (vol d'identité, *hacking* de centrales nucléaires, etc.), déstabilisation politique (propagande ciblée, création de fausses vidéos, etc.) ou automation d'équipement militaire (drones, soldats robots, etc.)<sup>7</sup>
2. **L'utilisation des informations collectées pour des fins autres que celles initialement prévues.** Ici, les citoyens craignent l'utilisation de dossiers médicaux complets par des compagnies d'assurance, celle de dossiers scolaires pour automatiser le marché de l'emploi, ou celle de systèmes d'optimisation de la circulation pour suivre et surveiller les usagers de la route.
3. **Un détournement volontaire de systèmes d'IA.** Quelqu'un de mal intentionné pourrait cibler directement le mode de fonctionnement de l'algorithme<sup>8</sup>, par exemple, en trompant un système de reconnaissance faciale pour avoir accès à des données protégées. On pourrait aussi profiter des défis de sécurité liés à la prolifération des objets connectés<sup>9</sup>, par exemple, pour prendre le contrôle d'une voiture autonome, ou pour paralyser un réseau avec une attaque massive de déni de service<sup>10</sup>.
4. **Une mauvaise évaluation d'une IA.** Une intelligence artificielle dont la fiabilité ou la robustesse a été surestimée et qui se retrouve à l'origine d'un accident<sup>11</sup> [6]. Par exemple, les citoyens ont mentionné qu'un accident de camion autonome ou une erreur systématique d'un logiciel de diagnostic médical peut avoir de lourdes conséquences.

<sup>6</sup> Brundage M, Avin S, Clark J, Toner H, Eckersley P, Garfinkel B, et al. The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. 2018;(February 2018). Disponible sur: <http://arxiv.org/abs/1802.07228>

<sup>7</sup> Bouvet M, Chiva E. Un regard (décalé ?) sur Intelligence Artificielle et Défense/Sécurité - CGE [Internet]. Conférence des Grandes Écoles. 2016 [cité le 3 sep 2018]. Disponible sur: <http://www.cge.asso.fr/liste-actualites/un-regard-decale-sur-intelligence-artificielle-et-defensesecurite/>

<sup>8</sup> Kurakin A, Goodfellow I, Bengio S, Dong Y, Liao F, Liang M, et al. Adversarial Attacks and Defences Competition [Internet]. 2018 [cité le 3 sep 2018]. Disponible sur: <https://arxiv.org/pdf/1804.00097.pdf>

<sup>9</sup> Zhang Z-K, Cho M-C, Wang C-W, Hsu C-W, Chen C-K, Shieh S. IoT Security: Ongoing Challenges and Research Opportunities. In: 2014 IEEE 7th International Conference on Service-Oriented Computing and Applications [Internet]. IEEE; 2014 [cité le 3 Sep 2018]. p. 230-4. Disponible sur: <http://ieeexplore.ieee.org/document/6978614/>

<sup>10</sup> Franceschi-Bicchierai Lorenzo. How 1.5 Million Connected Cameras Were Hijacked to Make an Unprecedented Botnet [Internet]. Vice Motherboard. 2016 [cité le 3 sep 2018]. Disponible sur: [https://motherboard.vice.com/en\\_us/article/8q8dab/15-million-connected-cameras-ddos-botnet-brian-krebs](https://motherboard.vice.com/en_us/article/8q8dab/15-million-connected-cameras-ddos-botnet-brian-krebs)

<sup>11</sup> Amodei D, Olah C, Brain G, Steinhardt J, Christiano P, Schulman J, et al. Concrete Problems in AI Safety [Internet]. [cité le 3 sep 2018]. Disponible sur: <http://arxiv.org/abs/1606.06565.pdf>

Les citoyens se sont donc demandé comment limiter les risques d'impact négatif de l'IA sur la sécurité (publique). Ils ont soulevé différents dilemmes potentiels en rapport avec la sécurité et l'intégrité des systèmes :

- > **Le respect de la transparence (un impératif souvent repris) pourrait-il nuire à la sécurité en facilitant le piratage ?**
- > **Est-ce qu'assurer la plus grande sécurité possible implique forcément un compromis avec l'efficacité du système (qui doit être sécuritaire sans pour autant devenir inopérant) ?**
- > **Et plus généralement, comment s'assurer de maximiser les impacts positifs du développement de l'IA tout en prévenant les effets néfastes ?**

Autres références pertinentes :

Asilomar AI principes

Vidéo : Adversarial ML

## Problématique 2 : L'IA, les médias et la manipulation de l'information

Les citoyens se sont inquiétés d'un risque de manipulation qui plane sur les utilisateurs à mesure que leurs actes sont de plus en plus influencés par des mécanismes d'IA qui influencent leurs décisions, souvent à leur insu ou via des incitatifs. Cela pose un problème de confiance en ces dispositifs puisqu'il y a une forme d'atteinte à l'autonomie et un risque d'orientation des actions (par exemple, en fonction d'intérêts privés). Les citoyens se demandent par exemple si les nouvelles technologies relevant de l'IA pourraient créer une nouvelle classe de

lobby, qui parfois risquerait d'avoir trop de pouvoir. Pour entretenir une certaine liberté dans les choix orientés par l'IA et éviter d'accorder une confiance aveugle à ces dispositifs, il est donc important de cultiver une pensée critique chez tout citoyen et professionnel interagissant avec l'IA.

Si la propagande n'est pas un phénomène nouveau, la facilité et la vitesse avec lesquelles celle-ci peut être créée et diffusée, via des fausses nouvelles ou des campagnes de désinformation, sont sans précédent. Cela est notamment permis grâce aux plateformes de création et diffusion de contenu en ligne (via les réseaux sociaux, les blogues et sites internet, les forums de discussion) qui sont structurées sur des modèles de rétention de l'attention, de publicité et de recommandations<sup>12-13-14-15</sup>. Ce phénomène est également amplifié par la capacité de cibler très précisément des individus selon la collecte et l'analyse de leurs données personnelles comme l'a mis en lumière le scandale Cambridge Analytica<sup>16</sup>, par exemple. La diversité des contenus à laquelle chaque individu est exposé est alors réduite à l'ensemble de ce qui se rapproche le plus de ce qu'il a déjà aimé, partagé, commenté. En se trouvant principalement confronté aux idées avec lesquelles il est en accord, l'individu est pris dans un effet de « bulle de filtre »<sup>17</sup> qui remet en question la capacité actuelle de tout citoyen à se forger une pensée critique.

Les grandes entreprises de médias sociaux annoncent des ensembles de mesures pour limiter le propagandisme potentiel de leurs outils (cf. les rapports de transparence de Facebook, Google et Twitter<sup>18</sup>). Mais est-ce suffisant ? Comment s'assurer que ces outils, qui ont démocratisé l'accès à l'information et la mise en relation, ne constituent

<sup>12</sup> Ingram M. Fake news is part of a bigger problem: automated propaganda [Internet]. Columbia Journalism Review. 2018 [cité 3 sept 2018]. Disponible à : <https://www.cjr.org/analysis/algorithm-russia-facebook.php>

<sup>13</sup> Lewis P. « Fiction is outperforming reality » : how YouTube's algorithm distorts truth. The Guardian [Internet]. 2 févr 2018 [cité 3 sept 2018]; Disponible à : <https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth>

<sup>14</sup> Marwick A, Lewis R. Media Manipulation and Disinformation Online [Internet]. Data & Society Research Institute; 2017 mai [cité 3 sept 2018]. Disponible à : <https://datasociety.net/output/media-manipulation-and-disinfo-online/>

<sup>15</sup> Tusikov N. Regulate social media platforms before it's too late [Internet]. The Conversation. 2017 [cité 3 sept 2018]. Disponible à : <http://theconversation.com/regulate-social-media-platforms-before-its-too-late-86984>

<sup>16</sup> Cadwalladr C, Graham-Harrison E. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian [Internet]. 17 mars 2018 [cité 3 sept 2018]; Disponible à : <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>

<sup>17</sup> Pariser E. The filter bubble: what the Internet is hiding from you. London: Penguin Books; 2012.

<sup>18</sup> Rapport préliminaire de Facebook : <https://transparency.facebook.com/community-standards-enforcement/>  
Rapport de transparence Google : <https://transparencyreport.google.com/about>  
Rapport de transparence Twitter : <https://transparency.twitter.com/fr.html>

pas des outils de démocratisation de la propagande ?  
Comment lutter contre la diffusion et l'amplification de fausses nouvelles et de campagnes de désinformation afin de préserver la démocratie ?  
Comment favoriser la démocratisation de l'accès à l'information tout en encourageant la formation d'une pensée critique et la prise de décision éclairée ?

Autres références pertinentes :

Caplan R, Hanson L, Donovan J. Dead Reckoning, Navigating Content Moderation After Fake News. Data & Society Research Institute; 2018 févr [cité 3 sept 2018]. Disponible à: <https://datasociety.net/output/dead-reckoning/>

Foisy P-V. Facebook veut s'attaquer aux fausses nouvelles au Canada [Internet]. Radio-Canada.ca. [cité 3 sept 2018]. Disponible à: <https://ici.radio-canada.ca/nouvelle/1109432/fake-news-facebook-fausses-nouvelles-canada-verification-faits>

Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, et al. The science of fake news. Science. 9 mars 2018;359(6380):1094-6.

Jeangène Vilmer J-B, Escorcía A, Guillaume, M, Herrera J. Les manipulations de l'information, un défi pour nos démocraties [Internet]. Paris, France: CAPS et IRSEM; 2018 août. Disponible à: <https://www.defense.gouv.fr/irsem/page-d-accueil/nos-evenements/lancement-du-rapport-conjoint-caps-irsem-les-manipulations-de-l-information>

Sites internet à consulter :

The Computational Propaganda Project : <http://comprop.oii.ox.ac.uk>

Observatory on Social Media : <https://truthy.indiana.edu>

Conversation AI : <https://conversationalai.github.io>

"A Citizen's Guide to Fake News" sur le site du Center for Information Technology & Society, UC Santa Barbara : <http://cits.ucsb.edu/fake-news>

## Problématique 3 : Gouvernance publique, privée ou participative : les communs numériques

À de nombreuses reprises, les citoyens ont soulevé des enjeux relatifs au partage de la gestion du développement de l'IA entre institutions publiques et privées, et les risques qui accompagnent ce partage, par exemple les conflits d'intérêts, la protection de l'indépendance des acteurs institutionnels ou des institutions publiques, la valeur marchande des données et la protection de la vie privée. Également, le risque de l'apparition d'un monopole privé dans la gouvernance du développement de l'IA a été mentionné à plusieurs reprises. En effet, en ce qui concerne notamment les compagnies propriétaires de données massives (qui sont à la base du fonctionnement de l'IA), certains s'inquiètent de l'apparition de monopoles, renforcés par les fusions de nouveaux fournisseurs de services plus petits<sup>19</sup>.

En ce qui a trait à une gouvernance étatique, un encadrement légal et juridique de l'IA s'accompagne de différents risques et défis<sup>20</sup>, par exemple, en se centrant trop sur les capacités des dispositifs aux dépens de la protection des valeurs humaines<sup>21</sup>, si bien qu'on peut se demander s'il est possible de réglementer l'IA et remettre en question le pouvoir réel de l'État<sup>22</sup>.

Si la discussion autour de la gouvernance oppose souvent les institutions publiques aux compagnies privées, une alternative a été proposée : celle d'une gouvernance participative qui donne directement la main aux citoyens en proposant, par exemple, la mise en place d'une grande consultation publique ou d'un espace permanent de concertation.

Dans cette perspective de gouvernance participative, la possibilité d'accorder de l'importance à la contribution des usagers dans la conception des outils d'IA et de leur gestion a ainsi

<sup>19</sup> IBig data: Bringing competition policy to the digital era - OECD [Internet]. [cited 2018 Sep 3]. Available from: <http://www.oecd.org/competition/big-data-bringing-competition-policy-to-the-digital-era.htm>

<sup>20</sup> Scherer MU. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. Harvard Journal of Law & Technology, Vol. 29, No. 2, Spring 2016. <http://dx.doi.org/10.2139/ssrn.2609777>

<sup>21</sup> Ambrose ML. Regulating the loop: ironies of automation law. 2014;38.

<sup>22</sup> Danaher J. Philosophical Disquisitions: Is effective regulation of AI possible? Eight potential regulatory problems [Internet]. Philosophical Disquisitions. 2015 [cited 2018 Sep 3]. Available from: <http://philosophicaldisquisitions.blogspot.com/2015/07/is-effective-regulation-of-ai-possible.html>

été proposée comme alternative. Cette participation pourrait prendre la forme d'une expérimentation collective (*design thinking*) par le biais de matériel en accès libre (*open source*). Ce matériel accessible à tous renvoie à la notion de (biens) communs numériques, soit l'ensemble des ressources et savoirs partagé et cocréé disponible en accès libre (par exemple, les logiciels libres). Plus qu'une simple forme de propriété, il s'agit ici d'un mode d'organisation coopératif garantissant l'horizontalité (échanges entre pairs) et la liberté d'expression<sup>23</sup>. Cette organisation dépend des formes de régulation décidées par les acteurs eux-mêmes.

« Le déploiement du numérique se caractérise par la création de biens publics par les communautés sur internet. Ce processus a supposé l'émergence de formes organisationnelles significativement nouvelles supportées par les technologies de l'information, en particulier les mouvements *open-source* puis Web 2.0. »<sup>24</sup>

Ce mode de gouvernance n'est pas lui non plus sans défis, il est notamment fragile à différentes formes d'*enclosure* (réduction des usages communs) par l'État comme par les compagnies<sup>25</sup>.

Ces enjeux soulèvent différents questionnements : quel serait le meilleur partage entre gestion publique, privée et participative pour la gouvernance de l'IA ? Ces différents modes entrent-ils forcément en tension et lequel serait le plus approprié ? Est-il nécessaire de baliser ces types de gestion et si oui, quelles balises mettre en place ?

Autres références pertinentes :

Chessen M. Encoded laws, policies, and virtues: the offspring of artificial intelligence and public-policy... [Internet]. Medium. 2017 [cited 2018 Sep 3]. Available from: <https://medium.com/artificial-intelligence-policy-laws-and-ethics/encoded-laws-policies-and-virtues-the-offspring-of-artificial-intelligence-and-public-policy-3dfb357faf9>

Shafto, P. Why Big Tech Companies Are Open-Sourcing Their AI Systems [Internet]. IFLScience. [cited 2018 Sep 3]. Available from: <https://www.iflscience.com/technology/why-big-tech-companies-are-open-sourcing-their-ai-systems/>

Sites internet à consulter :

[Le site internet de la FACIL](#)

[La Déclaration des communs numériques de la FACIL](#)

### 4.3.

## LES RECOMMANDATIONS DE LA RELÈVE EN RECHERCHE

L'exercice s'est avéré particulièrement fructueux et les trois brèves sont porteuses de pistes de recommandations pertinentes en vue du développement responsable de l'IA.

**La première brève** politique traite des conséquences engendrées par une mauvaise évaluation des capacités de l'IA, un élément qui s'intègre dans la première problématique proposée sur la sécurité et l'intégrité des systèmes. Cette brève vise à promouvoir la sécurité et la protection des Canadiens, plus particulièrement en ce qui concerne les systèmes embarqués. Les recommandations énoncées se rapportent à la création d'un organisme pancanadien de certification des systèmes embarqués utilisant l'IA (ARBIA) et le développement d'un régime de responsabilité sans faute. Cette brève s'est démarquée par la pertinence de ses recommandations, d'une part en faisant référence à des mécanismes déjà existants (le MEI<sup>26</sup> et l'intégration des recommandations dans la Stratégie numérique du Québec) et, d'autre part, en proposant

<sup>23</sup> Crosnier HL. Communs numériques et communs de la connaissance. Introduction. *tic&société*. 2018 May 31;(Vol. 12, N° 1):1–12.

<sup>24</sup> Ruzé E. La constitution et la gouvernance des biens communs numériques ancillaires dans les communautés de l'Internet. Le cas du wiki de la communauté open-source WordPress. *Management & Avenir*. 2013;(65):189–205.

<sup>25</sup> Crosnier HL. Une bonne nouvelle pour la théorie des biens communs. *Vacarme*. 2011;(56):92–4.

<sup>26</sup> Anciennement le MESI (Ministère de l'économie, de la science et de l'innovation), renommé MEI (Ministère de l'économie et de l'innovation) le jour même de l'écriture de la brève.

un mécanisme original de protection des citoyens (la responsabilité des dommages liés à un accident étant soit celle de l'État soit celle des entreprises qui développent lesdits systèmes embarqués).

**La seconde brève**, intitulée « Fausses nouvelles, vrais enjeux : s'éduquer pour y faire face » aborde le problème de la manipulation de l'information (en particulier, la création de fausses nouvelles) sur internet par le biais de l'IA. Les étudiants ont ici mis de l'avant la nécessité de favoriser le développement d'un esprit critique. Leurs recommandations vont dans le sens de l'importance de l'éducation face aux médias et à l'information au Québec. Deux principales recommandations ressortent de leur analyse : 1) Alerter la population québécoise sur la dissémination de fausses nouvelles (vigilance) notamment par le biais de SIA spécialisés et 2) Offrir des outils au corps enseignant québécois pour intégrer la conscientisation face aux fausses nouvelles à l'éducation dès l'école primaire (éducation). Afin d'assurer la mise en place de ces mécanismes, ils recommandent la création d'un Comité de vigilance numérique du Québec, annexé à l'*Observatoire international sur les impacts sociétaux de l'intelligence artificielle et du numérique et d'intégrer les processus éducatifs dans le Plan d'action numérique* du Québec.

**La troisième brève** traite de la problématique de la gouvernance de l'IA, explorant les enjeux liés aux brèches dans les politiques et les réglementations actuelles qui concernent le secteur de l'IA. Le défi est ici de trouver une forme de gouvernance de l'IA qui répond au mieux aux besoins des différents acteurs concernés (entreprises, citoyens, institutions publiques). L'objectif principal de cette brève politique est de *fournir une méthode de travail et de réflexion* afin de répondre adéquatement aux enjeux soulevés : politiques actuelles mal adaptées, inquiétudes des citoyens, flou relatif face à la responsabilité lors d'incidents et absence de méthodes pour gérer les problèmes liés à l'IA. Plusieurs pistes de solutions ont été proposées. Les étudiants recommandent la création d'un organisme (provincial) indépendant de régulation de l'utilisation des données communes qui serait responsable, entre autres, de la mise en place de mécanismes éducatifs ou, encore, d'ajuster les lois

et réglementations en vigueur afin de les adapter aux nouvelles réalités technologiques. L'intégration d'un volet à la mission de l'organisme afin que celui-ci soit constamment en phase avec le marché est également proposée. L'organisme en question, nommé « Educ'IA » lors de la présentation orale, s'organiserait sous la forme d'un *think tank*.

En résumé, les étudiants recommandent la mise en place d'un organisme indépendant de gestion de l'IA aux responsabilités variées (impliquant la création d'une certification ou d'un système de vigilance) et intégrant des processus éducatifs. Ces recommandations vont dans le sens de celles formulées lors des autres activités de coconstruction. Si la mise en place d'un organisme indépendant ou de mécanismes éducatifs s'aligne avec les recommandations de la Déclaration en vue de politiques publiques, celle de la création d'un système de responsabilité mériterait une analyse plus approfondie. Faisant écho aux recommandations citoyennes de l'hiver 2018 sur la mise en place de mécanismes assurantiels définissant les paramètres du partage des responsabilités en cas de faute (cf. *Rapport des résultats des ateliers de coconstruction de l'hiver*), cette recommandation invite à considérer une analyse à part entière portant sur la responsabilité civile et pénale face aux impacts de l'IA.

La pertinence de cette activité nous conduit enfin à appuyer le CIÉ dans sa recommandation de créer plus d'opportunités pour les étudiants aux cycles supérieurs de se former aux activités professionnelles extra-académiques, et plus particulièrement à la participation politique.

Ces trois activités nous ont ainsi permis d'explorer les enjeux liés au développement de l'IA à partir de nouvelles thématiques (ex. propagande) et également, d'expérimenter de nouvelles manières de procéder (ex. activité de simulation).

Quelle que soit l'activité, les recommandations des participants viennent appuyer la nécessité de mettre en place des formations adaptées pour tous et de mettre à jour le cadre légal et réglementaire lié à la progression de nouvelles connaissances sur le développement de l'IA et de ses impacts. Ces recommandations encouragent également la

promotion d'une gouvernance participative, mettant de l'avant l'importance d'impliquer les parties prenantes à différents moments clés de la gestion du développement de l'IA et de la décision politique.

En ouvrant la discussion sur des pistes de solution nouvelles ou des spécificités sectorielles, ces activités nous indiquent que la coconstruction mériterait d'être poursuivie au-delà des travaux de la Déclaration de Montréal, et renforcent la pertinence d'une consultation citoyenne sur le développement responsable de l'IA.

## 5. CONCLUSION

Ces trois activités nous ont ainsi permis d'explorer les enjeux liés au développement de l'IA à partir de nouvelles thématiques (ex. propagande) et également, d'expérimenter de nouvelles manières de procéder (ex. activité de simulation).

Quelle que soit l'activité, les recommandations des participants viennent appuyer la nécessité de mettre en place des formations adaptées pour tous et de mettre à jour le cadre légal et réglementaire lié à la progression de nouvelles connaissances sur le développement de l'IA et de ses impacts. Ces recommandations encouragent également la promotion d'une gouvernance participative, mettant de l'avant l'importance d'impliquer les parties prenantes à différents moments clés de la gestion du développement de l'IA et de la décision politique.

En ouvrant la discussion sur des pistes de solution nouvelles ou des spécificités sectorielles, ces activités nous indiquent que la coconstruction mériterait d'être poursuivie au-delà des travaux de la Déclaration de Montréal, et renforcent la pertinence d'une consultation citoyenne sur le développement responsable de l'IA.

# ANNEXE 1

## Les scénarios

### DÉMOCRATIE

#### Fausse nouvelle dans la campagne électorale

**23 mars 2022.** Ce matin, Dominique B. se rend à la réunion de crise de l'Agence sur l'intégrité de l'information (All), mise en place dans le cadre de la Loi contre la manipulation de l'information. Le président de la République sortant, candidat à sa réélection, vient de perdre 7 points dans les sondages d'intention de vote en trois semaines et la tendance à la baisse semble se confirmer. Alors qu'il était assuré de l'emporter deux mois auparavant, il est désormais dépassé par la candidate populiste de droite qui a pris la tête de la course électorale. Le tournant se situe le 2 mars, avec la diffusion sur internet d'une vidéo montrant le président de la République discuter avec le président du Mouvement des entreprises de France, en marge de son école d'été. Le président de la République assurait qu'il comprenait la situation des entreprises qui employaient des travailleurs immigrés sans papiers, qu'il était important de maintenir des bas salaires pour garantir la vitalité des petites et moyennes entreprises, et qu'il veillerait à ce que ces entreprises ne soient pas pénalisées.

La vidéo s'était vite répandue dans les réseaux sociaux et les propos du Président avaient été relayés dans les premières heures par deux grands médias, la chaîne d'information TBT et le site [lefutureur.com](http://lefutureur.com). Le porte-parole de l'Élysée avait immédiatement démenti les propos attribués au Président et avait fait savoir que la vidéo était un faux créé par une agence étrangère qui tentait d'interférer dans les élections françaises. La technique utilisée pour créer la vidéo avait été mise au point par l'entreprise américaine Monkeypaw Productions qui avait tiré parti des algorithmes GAN (*generative adversarial networks*), élaborés par des chercheurs de l'Université de Montréal en 2014.

Contre toute attente, les images créées grâce à l'IA avaient atteint un degré de réalisme stupéfiant en moins de dix ans, si bien qu'une fausse vidéo ne pouvait plus être détectée à l'œil nu.

Ni le démenti de l'Élysée, ni le mea culpa de TBT et du Futureur, ni encore l'interdiction de diffusion de la vidéo n'avaient eu l'effet espéré. La vidéo était encore consultable sur différents sites étrangers comme le site [rassvet.io](http://rassvet.io). Un député du parti populiste de droite en avait profité pour accuser le Président de faire le jeu de l'immigration clandestine et de nuire aux intérêts des Français. Le nombre de gazouillis avec le mot-clic *#Presidentclandestin* avait passé la barre des 300 000 en une semaine. À un mois du premier tour des élections présidentielles, Dominique B., directrice de l'All, doit présenter un plan pour enrayer les effets dévastateurs de cette fausse information et rétablir les conditions d'une campagne électorale saine. Mais ce matin, le sentiment d'avoir déjà épuisé toutes les solutions l'emporte à l'All.

## ENVIRONNEMENT

### La cote environnementale basée sur l'empreinte de carbone

**1<sup>er</sup> février 2025.** Pour la cinquième année de suite les températures battent des records de chaleur dans le monde entier. La majorité des pays ayant signé l'Accord de Paris en décembre 2015 n'ont pas tenu leurs engagements en raison des impératifs économiques de court terme, malgré les mises en garde du Groupe d'experts intergouvernemental sur l'évolution du climat (GIEC). En conséquence, les villes européennes du C40, le réseau des villes engagées dans la transition écologique, ont accéléré leur coopération pour proposer à leurs habitants un système de permis carbone individuel fortement incitatif, le système ÉcoFit, connecté à leur compte bancaire et aux différentes applications d'achat en ligne : dans ces villes, le prix des biens et services est affiché en euros et en carbone, et chaque citoyen doit viser 4 tonnes d'émission de carbone par an pour l'ensemble de sa consommation. Les personnes qui atteignent cet objectif augmentent leur cote environnementale calculée par l'algorithme ÉcoFit, à partir de leurs données personnelles de consommation. Cette cote leur donne un accès gratuit à de multiples services écoresponsables en transport, éducation, formation et culture.

**15 juin 2025.** Au moment de passer leur commande de coquilles Saint-Jacques grâce à leur réfrigérateur FrigoMax connecté, Ive et Charles, habitants du 20<sup>e</sup> arr. à Paris, découvrent ce nouveau système de points auquel ils viennent d'adhérer : Coquilles Saint-Jacques (Provenance : Pérou) : 12 € / 22 kg éq. CO<sub>2</sub>/kg\*<sup>27</sup>. Un message d'avertissement s'affiche : « Cet achat doit rester exceptionnel. Vous ne pourrez pas tenir votre objectif annuel si vous le reproduisez souvent. » Et l'algorithme de recommandation de FrigoMax leur propose alors des coquilles Saint-Jacques de Saint-Brieuc, fraîches, qui coûtent 22,5€ mais seulement 0,25 kg éq. CO<sub>2</sub>/kg.

**15 octobre 2025.** Après quelques écarts, et suite aux nombreux messages d'avertissement, Ive et Charles ont fait un effort pour consommer plus sobrement grâce aux recommandations d'ÉcoFit : régime presque végétarien, nouvelle isolation de leur logement, transport en commun et en vélo, contrat d'électricité verte, choix exclusif d'applications avec *data-centers* carbone neutre : c'est qu'au bureau, tout le monde compare maintenant sa cote environnementale !

**1<sup>er</sup> décembre 2025.** Grâce à leurs comportements de plus en plus vertueux, Ive et Charles ont réussi à rester juste en-dessous du plafond visé : après 6 mois, ils sont chacun à 1,95 tonne de carbone pour leur consommation globale. De plus, ils ont moins dépensé monétairement, ce qui leur procure une épargne inattendue. Le couple considère alors de réaliser son projet de séjour à Cuba pour Noël et commence à consulter les sites des agences de voyage. Un message leur parvient sur leur téléphone : « Attention à l'*effet rebond* : dépenser vos économies dans un voyage annulerait tous vos efforts ! Pensez à voyager local ! »

<sup>27</sup> kg éq. CO<sub>2</sub>/kg = kilogramme équivalent carbone ; exprimé ici par kg de produit importé par avion.

## MONDE DU TRAVAIL

### Forage des données (*data mining*) RH pour optimiser l'ambiance au travail

30 octobre 2025. Pierre-André a enfin décroché un emploi dans un bon bureau d'avocats qui traite notamment du droit de l'environnement, l'un de ses domaines de prédilection.

Après trois semaines de travail, il rencontre Marco aux ressources humaines pour une séance de mentorat personnalisée. Marco fait le point sur l'intégration de Pierre-André, sur ses attentes initiales, ses difficultés, etc. Il lui explique aussi que la firme utilise désormais AmbIA+, une IA d'analyse conversationnelle qui étudie les attitudes des salariés et aide à maintenir une ambiance de travail apaisante et productive. C'est une question d'efficacité. Ainsi, tous les courriels, appels téléphoniques et prises de parole en réunion d'équipe sont analysés pour extraire un historique des humeurs et des émotions des salariés. Ces données sont ensuite rapportées à un laboratoire de recherche en psychologie.

Pierre-André est déstabilisé et même un peu inquiet, mais Marco essaie de le rassurer :

- > AmbIA+ fournit une assistance individualisée, elle conseille et entraîne, mais il n'y a pas de sanction. D'ailleurs, AmbIA+ ne mémorise que la forme des interactions, et tous les échanges que vous avez eus jusqu'à présent au bureau se sont bien passés.

Tous, sauf pour le 15 et le 16 octobre derniers. Pierre-André travaillait alors sur le dossier de la nouvelle station d'épuration des eaux usées de la ville de Lille. « Selon AmbIA+, rapporte Marco, vous avez à plusieurs reprises interrompu vos collègues en réunion pour répéter les mêmes idées, ce qui a créé de la tension chez eux. Il faudrait essayer d'exposer vos arguments en une fois, lors du tour de table, pour ne pas perdre de temps. »

Mais ce n'est pas tout :

- > Apparemment, l'algorithme a aussi détecté des périodes d'inactivité sur le réseau de plusieurs heures, sans aucun échange avec vos collègues. Ce n'est pas grave en soi, mais c'est mieux de maintenir le contact avec l'équipe. Est-ce que vous vous souvenez de la raison de cette inactivité ?

Pierre-André n'est plus seulement inquiet, il est embarrassé et s'interroge sur la pertinence de ces questions :

- > Oui, c'est vrai, j'aime bien travailler avec un crayon sur un rapport papier et je préfère ne pas rédiger directement sur le document collaboratif en ligne... et en effet lors de la réunion du 15, j'apportais une idée nouvelle qui ne me semblait pas bien comprise et je craignais qu'on ne l'oublie. Mais est-ce vraiment un problème ?

Compréhensif, Marco répond qu'il n'y a vraiment aucun problème : « Mais ne vous déconnectez pas de l'équipe, c'est mieux pour la performance collective. Allez, on se revoit dans deux mois. Et bonne chance pour la réunion de demain ! »

## ANNEXE 2

### Les brèves étudiantes

#### Simulation 2018, CIÉ-FRQ

#### Brève politique sur l'intelligence artificielle

*Le présent document est le résultat d'un exercice de simulation, dont l'objectif était d'acquérir des compétences en rédaction et en communication publique. Étant donné le contexte pédagogique dans lequel cette note a été produite, elle n'a pas la vocation, dans les faits, d'être adressée à des décideurs ou à des acteurs de la fonction publique.*

*La Déclaration de Montréal a choisi de publier ces brèves afin de représenter fidèlement le résultat d'un travail réalisé en 6 heures par les étudiants de la relève et montrer la pertinence d'un tel exercice.*

#### Problématique 1 : Sécurité publique et intégrité des systèmes

#### Sous-problématique 4 : Les conséquences engendrées par une mauvaise évaluation des capacités de l'intelligence artificielle



**Document rédigé par :**

**Joël Simoneau**

**Jérôme Gélinas Bélanger**

**Fidele Ndjoulou**

**Moumouni Ouiminga**

**À l'intention du Gouvernement du Canada**

## **Titre de la brève**

Pour l'établissement d'un système de responsabilité de l'intelligence artificielle dans les biens de consommation

Cette brève politique expose une démarche qui vise à promouvoir une intelligence artificielle responsable pour la sécurité et la protection des Canadiennes et Canadiens. Elle porte spécifiquement sur les systèmes embarqués. Les recommandations énoncées sont :

1. La mise sur pied d'un organisme pancanadien de certification des systèmes embarqués<sup>1</sup> utilisant l'intelligence artificielle
2. Le développement d'un régime de responsabilité sans faute

De manière concrète, le gouvernement devrait s'atteler dans un premier temps à la création d'un organisme fédéral responsable de la certification obligatoire des systèmes embarqués utilisant l'IA. Dans un deuxième temps, il est impératif de mettre sur place un régime propriétaire sans faute.

Une telle politique permettra d'assurer une meilleure santé et sécurité ainsi qu'une protection légale à tous les Canadiennes et les Canadiens dans leur interaction avec des objets utilisant l'IA, tout en impliquant les entreprises privées dans le processus de la saine utilisation de l'IA.

---

<sup>1</sup> On qualifie de « système embarqué » un système électronique et informatique autonome dédié à une tâche précise, souvent en temps réel, possédant une taille limitée et ayant une consommation énergétique restreinte. [www.futura-sciences.com/tech/definitions/technologie-systeme-embarque-15282/](http://www.futura-sciences.com/tech/definitions/technologie-systeme-embarque-15282/)

## La société canadienne à l'ère du développement de l'intelligence artificielle

Notre société est en train de vivre une transformation globale basée sur l'évolution du numérique. Autant celui-ci véhicule des informations à une vitesse précédemment inimaginable, qu'il transforme notre rapport avec les objets. Les avancées technologiques récentes en intelligence artificielle (IA) permettent d'imaginer un futur imminent où certaines tâches avec prises de décision redondantes seraient attribuées à des logiciels conçus expressément pour cette fonction. Les véhicules autonomes sont déjà au coin de la rue, les dispositifs médicaux intelligents sont derrière les portes des universités. Une mauvaise médication ou des accidents automobiles sont des dangers qui tendent à être réglés par l'utilisation intelligente et sécuritaire de l'IA, mais il faut aussi s'assurer qu'elle n'en devient pas la cause. Cela représente des inquiétudes énoncées par les Canadiennes et les Canadiens à travers les travaux de la Déclaration de Montréal pour un développement responsable de l'IA.

La régularisation des systèmes embarqués, soit un appareil physique contenant un logiciel utilisant une IA, devrait

être un projet d'importance pour le gouvernement canadien.

Ceux-ci représentent une

implémentation physique et commercialisable d'un produit d'IA, et il serait important d'en assurer une réglementation en amont de leur arrivée prochaine sur le marché canadien. Une prise de décision proactive et l'installation d'un cadre réglementaire permettrait l'encadrement des IA pouvant avoir un impact physique direct sur le peuple canadien.

Ce document propose l'instauration d'un organisme réglementaire de certification des systèmes embarqués utilisant l'IA et d'un régime de responsabilité basé sur le propriétaire sans faute. L'organisme permettrait d'encadrer les normes de sécurité de conception et d'utilisation des systèmes, et le régime permettrait de définir exactement le rapport de responsabilité dans le but de protéger les Canadiennes et les Canadiens, autant légalement qu'au niveau de leur santé et bien-être. La combinaison de ces deux mesures encadrera les systèmes embarqués, de leur commercialisation jusqu'à leur utilisation, ce qui maximisera les impacts positifs du développement de l'IA, en réduisant ses effets néfastes.

## **Constats et pistes d'action sur le développement de l'intelligence artificielle au Canada**

### **1. Organisme de certification**

À l'heure présente, aucun cadre législatif n'existe quant à l'utilisation de l'intelligence artificielle intégrée à des systèmes embarqués au Canada. Ce flou juridique pose un certain nombre de défis pour les différents paliers de gouvernement, notamment le gouvernement fédéral, relativement à leur capacité de structurer la mise en marché et la régulation de ces objets au pays. De façon plus générale, ce manque de structure à ce niveau engendre des complexités juridiques en termes d'évaluation du risque que présentent ces technologies pour le public, mais aussi en termes de l'attribution du poids de la responsabilité advenant un incident découlant de l'utilisation d'une technologie basée sur l'IA.

Face à ces défis, il apparaît nécessaire pour l'État canadien de créer un organisme réglementaire de certification des systèmes embarqués utilisant l'IA, l'office de

réglementation nommé l'Agence de Réglementation sur les Biens utilisant l'Intelligence Artificielle (ARBIA), à vocation interdisciplinaire et agissant comme pilier décisionnel. Cet organisme possédera trois principaux axes d'action afin de parvenir à structurer la réglementation de l'IA à l'échelle canadienne: 1) l'investissement dans la recherche et l'innovation permettant le développement de balises législatives basées sur des connaissances techniques, 2) l'instauration de comités experts possédant une bi-spécialisation reposant sur l'IA et leur propre champ d'expertise à l'intérieur des différents ministères pouvant être éventuellement affectés par le développement de l'IA et 3) le développement d'une plateforme réglementaire encadrant la mise en marché et le régime propriétaire sans faute.

L'implication directe du gouvernement canadien dans les cas de problématique de bien utilisant l'IA permettra d'assurer une veille scientifique et sécuritaire proactive, et de protéger légalement les consommateurs canadiens, qui n'auront pas à subir des procès-bâillons.

## **2. Régime de responsabilité sans faute**

On entend par responsabilité l'obligation de répondre d'un dommage devant la justice et d'en assumer les conséquences notamment civiles et pénales envers la victime et/ou la société. Dans un régime de responsabilité sans faute, le gouvernement du Canada sera responsable des accidents physiques ou matériels causés par un bien matériel utilisant l'IA. Dans le cas d'un bien non conforme au processus de certification, le gouvernement canadien peut tenter des actions contre le fabricant.

L'implication directe du gouvernement canadien dans les cas de problématique de bien utilisant l'IA permettra d'assurer une veille scientifique et sécuritaire proactive, et de protéger légalement les consommateurs canadiens, qui n'auront pas à subir des procès-bâillons.

### **2.1 Secteurs d'activités concernés**

L'intelligence artificielle s'applique à plusieurs secteurs d'activité notamment la santé, l'éducation, la sécurité, l'agriculture. Cependant, cette brève politique touche de manière spécifique l'automobile autonome, les dispositifs médicaux et la domotique.

### **2.1.1 Automobiles autonomes**

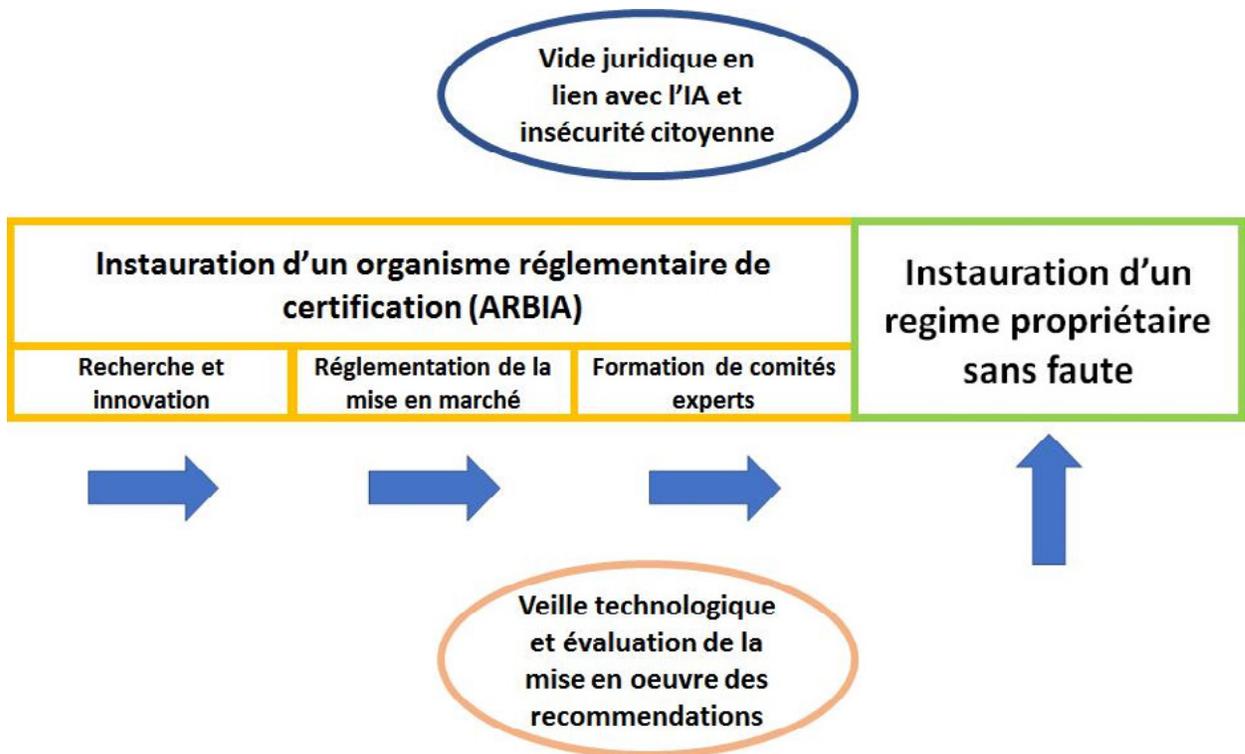
Avec le développement de l'intelligence artificielle, le secteur de l'automobile a connu une transformation radicale. Une nouvelle catégorie d'automobile dite automobile autonome est mise sur le marché. Néanmoins, ces voitures ont déjà causé des accidents aux États-Unis, par exemple l'accident mortel causé par une voiture autonome en Floride en mars 2018. Compte tenu de l'absence des règles spécifiques à la circulation des automobiles autonomes, et dans le souci d'apporter une meilleure protection aux citoyens, il est impératif de définir un cadre réglementaire au niveau fédéral. Ce cadre va fixer la responsabilité des parties prenantes, à savoir l'État et les compagnies propriétaires des voitures autonomes. Une faute liée aux défaillances est une faute de la compagnie responsable et propriétaire de la voiture, tandis qu'une utilisation faite par l'individu utilisateur est une faute de sa part.

### **2.1.2 Dispositifs médicaux**

Pour faciliter la vie des personnes vivant avec le diabète, une pompe à insuline a été développée. C'est un système intégré composé de trois parties (boîtier,

composantes électroniques, cathéter) qui libère automatiquement de l'insuline. Le fonctionnement de ce système nécessite impérativement une formation du patient, une autosurveillance glycémique et un suivi

médical rapproché. Cela engendra une responsabilité du patient en cas de mauvaise utilisation de sa part. Les conséquences pourraient être énormes au point d'engendrer d'éventuels cas de décès.



## Retombés et recommandations

Dans l'objectif d'assurer la santé et la sécurité des Canadiennes et Canadiens face aux systèmes embarqués utilisant l'IA et de maximiser les impacts positifs du développement de l'IA, nous émettons les recommandations suivantes.

### Recommandation 1

#### Création de l'Agence de Réglementation sur les Biens utilisant l'Intelligence Artificielle (ARBIA)

Retombées :

- Coût nul pour le gouvernement canadien  
Le financement de l'organisme de certification et des actions légales sera couvert par une licence de fabrication des systèmes.
- Fiabilité des systèmes d'IA pour la santé et sécurité des Canadiennes et Canadiens.  
Par le respect de normes définies par des comités experts, normes qui seront mises à jour selon les cas vécus.

## Recommandation 2

### Instauration d'un régime propriétaire sans faute

Retombées :

- Accessibilité judiciaire améliorée dans le cas de faute des fabricants.  
Le système sans faute donne la responsabilité de la poursuite judiciaire au gouvernement canadien, qui a plus de ressources que les citoyennes et citoyens individuellement.
- Promotion de l'implication sociale des entreprises.  
Considérant leur responsabilité directement impliquée, les fabricants vont être encouragés à développer des mécanismes d'utilisation sécuritaire de leurs produits.
- Veille technologique et sécuritaire du gouvernement canadien  
Considérant l'implication directe du gouvernement canadien dans les processus judiciaires, celui-ci assure une veille permanente dans la gestion saine des IA.

Bibliographie :

<sup>1</sup> Pour aller plus loin voir Palmer, Vernon. « Trois principes de la responsabilité sans faute » (1987) 39:4 Revue internationale de droit comparé; Mémeteau, Gérard. « Un point sur la responsabilité civile du fait des prothèses » (2013) 2013:123 Médecine & Droit 175-180; Jacob, Julien. « Prévention des risques technologiques à l'aide de la responsabilité civile en présence d'une innovation à double impact » (2013) 202:1 Économie & amp; prévision 1-18.

<sup>2</sup> <https://diabetnutrition.ch/les-traitements/la-pompe-a-insuline-quest-ce-que-cest/>

<sup>3</sup> <https://ici.radio-canada.ca/info/videos/media-7560667/premier-accident-mortel-impliquant-une-voiture-autonome> source consultée le 18 octobre 2018

## *Simulation 2018 dans le cadre des J2R*

### *« Politique et intelligence artificielle »*

*Le présent document est le résultat d'un exercice de simulation, dont l'objectif était d'acquérir des compétences en rédaction et en communication publique. Étant donné le contexte pédagogique dans lequel elle a été produite cette note, elle n'a pas la vocation, dans les faits, d'être adressée à des décideurs ou à des acteurs de la fonction publique.*

*La Déclaration de Montréal a choisi de publier ces brèves afin de représenter fidèlement le résultat d'un travail réalisé en 6 heures par les étudiants de la relève et montrer la pertinence d'un tel exercice.*

Fausse nouvelles, vrais enjeux : s'éduquer pour y faire face

Présenté aux membres du jury

Par

Jean Clairemond César, étudiant au doctorat en éducation, Université de Sherbrooke  
Isabelle Dufour, inf., candidate au doctorat, Université de Sherbrooke  
Gaël Grissonnanche, post-doctorant en physique, Université de Sherbrooke  
Philippe Lebel, doctorant en microbiologie, Université de Montréal

18 octobre 2018

## Tables des matières

Tables des matières .....	2
But de la brève .....	3
Couverture (1 page) .....	4
Introduction (1 page) .....	5
Données probantes et analyse (3 pages).....	6
Répercussions sur les politiques et recommandations (1 page) .....	9
Tableau 1. Grille d'évaluation des brèves politiques.....	10

## But de la brève

Émettre des recommandations auprès d'un décideur public en vue d'offrir une ou plusieurs pistes de solution à un problème spécifique découlant d'une des trois problématiques décrites dans le document élaboré par l'équipe de la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle. Une problématique sera attribuée par équipe et les participants devront identifier les éléments suivants :

- Le **problème**
- La ou les **solution(s)** recommandée(s) et les **répercussions** sur la population visée et non visée
- Le **décideur public** impliqué
- Les **facteurs environnementaux** pouvant faire obstacle à la mise en œuvre de la ou des solution(s) recommandée(s).

## Couverture (1 page)

La première page présente une synthèse de la brève politique. Elle présente la pertinence de la brève et ses grandes lignes, les conclusions clés et la marche à suivre.

Cette brève politique présente le problème des fausses nouvelles sur l'internet. Aujourd'hui la proportion de Canadiens qui consomme de l'information en ligne a dépassé celle des médias traditionnels. L'efficacité de cette technologie repose sur l'intelligence artificielle (IA) en offrant des contenus filtrés selon le comportement et l'intérêt de l'utilisateur. De nos jours, plusieurs acteurs sociopolitiques ont levé le drapeau rouge sur cet enjeu de société. D'un autre côté, les grandes entreprises de médias sociaux telles que Google, Facebook, Amazon et tant d'autres proposent déjà des mesures pour limiter le potentiel propagandiste de leur algorithme, et la définition d'une fausse nouvelle ne fait pas consensus. La pertinence de notre brève se situe dans la nécessité d'augmenter l'esprit critique au sein de la population québécoise. L'absence d'esprit critique peut occasionner plusieurs problèmes en éducation et en santé et dans d'autres domaines. Destiné aux ministres concernés par la stratégie numérique, ce document présente plusieurs recommandations sur l'importance de l'éducation aux médias et à l'information au Québec.

## Introduction (1 page)

Cette section décrit l'objectif principal de la brève et le problème politique. Elle établit un lien entre les données probantes et le problème.

L'avènement de l'internet apporte aux citoyens la démocratisation de l'accès à l'information à travers des moteurs de recherches intelligents et des médias sociaux. Aujourd'hui, la proportion de Canadiens consommant de l'information en ligne a dépassé celle des médias traditionnels. L'efficacité de cette technologie repose sur l'intelligence artificielle (IA) en offrant des contenus filtrés selon le comportement et l'intérêt de l'utilisateur. Tandis que certaines études montrent que cette exposition partielle à l'information tend à engendrer chez l'utilisateur une confirmation systématique de sa pensée, d'autres en revanche arguent que celui-ci n'a jamais été exposé à une telle diversité de sources lorsque comparé à la presse écrite, à la télévision, à la radio, etc.

C'est dans ce contexte qu'émerge sur la scène internationale la notion de fausse nouvelle comme un enjeu de désinformation massive dans une société démocratique. L'usage d'IA comme en a fait la firme Cambridge Analytica aux États-Unis a montré au monde le niveau de déstabilisation sociétale que cette technologie peut engendrer. Alors que les grandes entreprises de médias sociaux telles que Google, Facebook, Amazon et tant d'autres proposent déjà aujourd'hui des mesures pour limiter le potentiel propagandiste de leur algorithme, la définition d'une fausse nouvelle ne fait pas consensus. En effet, selon le Global News, près de 58% des Canadiens définissent celle-ci comme une histoire pour laquelle les faits sont faux. Cependant, 46% l'emploient pour désigner les nouvelles de journaux et les discours de personnalités politiques n'exprimant qu'un unique côté des faits. Encore, ce même chiffre désigne le pourcentage pensant que ce terme est uniquement utilisé par les politiciens pour discréditer les médias qui les critiquent. À l'autre bout du spectre, des actions pour valoriser l'esprit critique de l'utilisateur demeurent une avenue qui doit être envisagée.

La problématique amenée par l'essor des fausses nouvelles dans les médias est importante et est susceptible d'avoir un impact important sur la population. À cet égard, l'objectif de cette brève est d'augmenter la sécurité de la population et leur éducation face aux fausses nouvelles.

### Données probantes et analyse (3 pages)

Cette section représente le cœur de la brève politique. La qualité de cette section est jugée par la pertinence des données présentées, des interprétations tirées de ces données, ainsi que de leurs apports et de leurs limites. Elle peut contenir des graphiques, des tableaux et des schémas.

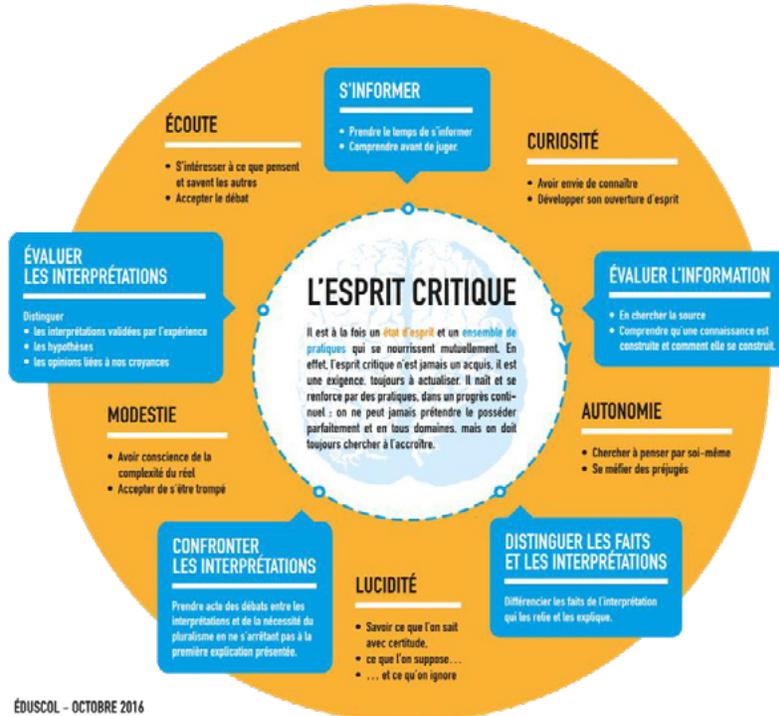
Selon Jeff Yates, expert québécois de la question, une fausse nouvelle se définit comme « une information soit carrément fausse, détournée, exagérée ou dénaturée à un point tel qu'elle n'est plus véridique, et présentée comme une vraie nouvelle dans le but de tromper les gens. Cela peut être fait pour générer des clics et des partages sur les réseaux sociaux, pour atteindre des objectifs quelconques (politiques, idéologiques, économiques, etc.) ou simplement pour se moquer de la crédulité des lecteurs ». Sujet de débats socio-économiques, les fausses nouvelles dans les médias ont connu un essor marqué durant les dernières années, et principalement avec le développement de l'IA. En effet, des méthodes associées à l'IA sont utilisées par les sites de médias sociaux et peuvent procéder de façon automatique à la diffusion de fausses nouvelles.

En Amérique du Nord, environ 60% de la population croit que la dispersion de telles informations dans les médias cause de la confusion. Les enfants et les adolescents sont particulièrement à risque d'attribuer du crédit et de participer à leur diffusion. Le manque de consensus dans la définition d'une fausse nouvelle contribue à l'incertitude vécue par la population, 45% des Canadiens en ayant une vision erronée. À l'ère numérique, la nécessité de développer une pensée critique concernant les informations transmises par les médias se positionne donc comme un enjeu central.

Selon Vallerand, la pensée critique est une pensée responsable qui s'appuie sur des critères et qui est sensible au contexte et aux autres (Vallerand, 2016). L'IA pourrait être utilisée pour développer l'esprit critique des jeunes et les former au doute constructif. Les jeunes du Québec apprendront comment mettre en perspective une information diffusée sur le web. En ce sens, les décideurs politiques donneront les moyens nécessaires pour y arriver.

En France par exemple, le développement de l'esprit critique est au centre de la mission assignée au système éducatif français, comme le présente le modèle de l'esprit critique d'Eduscol. Il est renforcé par l'attention désormais portée à l'éducation aux médias et à l'information. Le travail de formation des élèves au décryptage du réel et à la construction, progressive, d'un esprit éclairé, autonome et critique est essentiel. L'esprit critique est une compétence essentielle du citoyen et de la citoyenne du 21<sup>e</sup> siècle. Analyser une source, mettre en perspective une image ou une information, en extraire l'essentiel, critiquer le contenu, se questionner sont autant de savoirs numériques nécessaires à l'exercice d'une citoyenneté avisée. Une étude faite en 2017 au Royaume-Uni a montré que seulement 4 % de la population testée avait été capable d'identifier correctement les vraies des fausses nouvelles. Ce résultat est inquiétant, notamment en termes de sécurité publique. Pensons par exemple au mouvement anti-vaccination qui

cause un retour en force de maladies mortelles, tel que la coqueluche aux États-Unis, malgré qu'il a été démontré depuis longtemps que les causes de ce mouvement sont fausses.



Selon les experts de la pensée critique, Christopher DiCarlo et l'auteur du *Petit cours d'autodéfense intellectuelle*, Normand Baillargeon, la solution passe par l'éducation. En effet, il est préférable que l'école forme une jeunesse plus critique de ce qu'elle consulte plutôt que de faire confiance aux grandes entreprises privées du web pour autocensurer leur contenu.

D'ailleurs on retrouve plusieurs initiatives, ailleurs comme chez nous, qui vont dans ce sens. En France, plus précisément en Haute-Savoie, des enseignantes ont créé une habitude locale où elles prennent une heure par semaine pour sensibiliser leurs élèves à détecter les fausses nouvelles. Cette initiative, accueillie avec enthousiasme par les élèves, semble rapidement porter fruit puisque ces jeunes de 10 ans ont déjà développé les réflexes de vérifier d'où proviennent des images-chocs qui publicisent de fausses nouvelles sensationnalistes, par exemple.

Plus près de chez nous, depuis mai 2018, un nouveau programme s'implante dans les écoles ontariennes : Actufuté. Ce programme se veut une collaboration entre la Fondation pour le journalisme canadien et l'organisme CIVIX qui est responsable du programme Vote étudiant. Ce dernier prend vie autour des périodes d'élections et encourage la participation citoyenne des 9 à 19 ans. C'est dans ces périodes riches en nouvelles qu'Actufuté viendra aider les élèves à démystifier le vrai du faux.

Au Québec, « École branchée », un organisme sans but lucratif (OSBL) propose des outils aux enseignantes et aux enseignants pour intégrer ces considérations dans leur programme de tous les jours. Malheureusement, à ce jour, seulement 15 % à 20 % du corps enseignant est rejoint par l'organisme. Démonstration qu'une intervention gouvernementale est nécessaire pour offrir une protection équitable à tous nos jeunes contre ce fléau. Ce faisant, la jeunesse pourra aussi transmettre cette information et conscientiser ses proches à la problématique.

Pour y arriver, le Plan d'action numérique en éducation et en enseignement supérieur, annoncé à l'été 2018, prévoit quelque 900 millions de dollars pour, justement, préparer la génération de demain à ce nouvel environnement numérique. Le gouvernement du Québec pourrait ainsi soutenir les services d'« École branchée », voire même intégrer son contenu au cursus normal de l'éducation primaire et secondaire.

L'intégration de cette notion d'éducation directement au cursus scolaire vient contrer l'obstacle environnemental principal. Le désir des professeurs d'assurer le développement de leurs étudiants pourra aussi agir à titre de facilitateurs.

## Répercussions sur les politiques et recommandations (1 page)

Cette section présente les recommandations proposées et les répercussions anticipées. Ces recommandations et ces répercussions peuvent s'organiser autour de thèmes, de parties intéressées ou d'échéancier.

L'argumentaire soulevé met en lumière plusieurs défis soulevés par l'IA. Entre autres, elle contribue à la diffusion de masse de fausses nouvelles. Cela cause de la confusion au sein de la population, une perte de confiance envers les sources d'information. Les jeunes et les adolescents sont particulièrement sensibles aux fausses nouvelles, leur capacité de raisonnement et leur esprit critique étant en construction. L'implication des instances gouvernementales est donc primordiale pour assurer la protection et l'éducation des populations, et particulièrement des jeunes, sur la problématique des fausses nouvelles. Les recommandations adressées font appel à des notions de vigilance et d'éducation.

### Vigilance

Notre recommandation :

- Alerter la population québécoise sur la dissémination de fausses nouvelles

L'objectif étant de favoriser le développement et l'utilisation d'IA spécialisée pour détecter les fausses nouvelles diffusées sur les médias sociaux.

Cette initiative s'inscrit également en parallèle avec le Détecteur de rumeurs, où les alertes du CVMQ, en cas de détection d'une fausse nouvelle de grande importance, pourront être diffusées.

La création du Comité de vigilance numérique du Québec sera annexée à l'Observatoire international sur les impacts sociétaux de l'intelligence artificielle et du numérique.

### Éducation

Notre recommandation :

- Offrir des outils au corps enseignant québécois pour intégrer la conscientisation face aux fausses nouvelles dans l'éducation, dès l'école primaire.

Cette mesure aura deux buts : préparer directement cette génération à affronter le fléau des fausses nouvelles et les inciter à répandre ces bonnes pratiques auprès de leurs proches.

Grâce au financement déjà prévu pour le Plan d'action numérique en éducation et en enseignement supérieur ainsi qu'aux initiatives déjà en place, il sera possible de protéger la population québécoise sans investissement supplémentaire et sans réinventer la roue. À court terme, la promotion de ces outils auprès des enseignantes et des enseignants aura déjà un impact et il sera possible de penser intégrer ces enseignements au cursus normal à moyen terme.

Tableau 1. Grille d'évaluation des brèves politiques

Critères	Tous les points	- 1 p	- 2 p	- 3 p
<b>Couverture</b>	La synthèse présente la pertinence de la brève et ses grandes lignes, les conclusions clefs et la marche à suivre.	Des éléments sont manquants	La synthèse est manquante	
<b>Introduction</b>	Cette section décrit <u>très bien</u> l'objectif principal de la brève et le problème politique. Elle établit un lien entre les données probantes et le problème.	Cette section décrit <u>bien</u> l'objectif principal de la brève et le problème politique. Elle établit un lien entre les données probantes et le problème.	Cette section décrit <u>convenablement</u> l'objectif principal de la brève et le problème politique. Elle établit un lien entre les données probantes et le problème.	Cette section est absente
<b>Données probantes et analyse</b>	Cette section est <u>très pertinente</u> au regard du problème; Les interprétations sont <u>justes et convaincantes</u> ; Les facteurs environnementaux (socio-politico-économico-culturels) sont <u>très bien pris en compte</u> dans la possible intégration des recommandations; Les apports et les limites sont <u>très bien identifiés</u> .	Cette section est <u>pertinente</u> au regard du problème; Les interprétations sont <u>justes</u> ; Les facteurs environnementaux (socio-politico-économico-culturels) sont <u>bien pris en compte</u> dans la possible intégration des recommandations; Les apports et les limites sont <u>bien identifiés</u> .	Cette section est <u>plutôt pertinente</u> au regard du problème; Les interprétations sont <u>plutôt justes</u> ; Les facteurs environnementaux (socio-politico-économico-culturels) sont <u>pris en compte</u> dans la possible intégration des recommandations; Les apports et les limites sont <u>identifiés</u> .	Cette section <u>n'est pas pertinente</u> au regard du problème; Les interprétations sont <u>erronées</u> ; Les facteurs environnementaux (socio-politico-économico-culturels) <u>ne sont pas pris en compte</u> dans la possible intégration des recommandations; Les apports et les limites <u>ne sont pas correctement identifiés</u> .
<b>Répercussions et recommandations</b>	Les recommandations proposées sont <u>très pertinentes</u> et les répercussions anticipées <u>très bien identifiées</u> .	Les recommandations proposées sont <u>pertinentes</u> et les répercussions anticipées sont <u>bien identifiées</u> .	Les recommandations proposées sont <u>plus ou moins pertinentes</u> et les répercussions anticipées <u>plus ou moins bien identifiées</u> .	Les recommandations proposées <u>ne sont pas pertinentes</u> et les répercussions anticipées <u>ne sont pas bien identifiées</u> .
<b>Qualité de la présentation orale</b>	La présentation de la brève est très convaincante.	La présentation de la brève est convaincante.	La présentation de la brève est peu convaincante.	La présentation de la brève n'est pas convaincante.
<b>Total des points</b>	/15			

## *Simulation 2018 dans le cadre des J2R*

### *« Politique et intelligence artificielle »*

*Le présent document est le résultat d'un exercice de simulation, dont l'objectif était d'acquérir des compétences en rédaction et en communication publique. Étant donné le contexte pédagogique dans lequel a été produite cette note, elle n'a pas la vocation, dans les faits, d'être adressée à des décideurs ou à des acteurs de la fonction publique.*

*La Déclaration de Montréal a choisi de publier ces brèves afin de représenter fidèlement le résultat d'un travail réalisé en 6 heures par les étudiants de la relève et montrer la pertinence d'un tel exercice.*

Gouvernance publique, privée ou participative : les communs numériques

Présenté aux membres du jury

Par

Thomas Bousquet  
Alexandre Côté, PhD(c)  
Christian Kouakou, PhD(c)

## Tables des matières

<b>Simulation 2018 dans le cadre des J2R</b> .....	1
<b>« Politique et intelligence artificielle »</b> .....	1
Tables des matières .....	2
Couverture .....	3
Introduction .....	4
Données probantes et analyse .....	5
Le problème de l’immigration discriminante .....	5
La confidentialité des données .....	5
La stratégie du Québec en matière d’intelligence artificielle.....	6
Répercussions sur les politiques et recommandations .....	7

## Couverture

### Vue d'ensemble

L'évolution rapide de la technologie et la science entourant l'intelligence artificielle (IA) exposent certaines brèches dans les politiques et les réglementations actuelles qui concernent ce secteur de développement. Le gouvernement doit se pencher sur ce problème qui soulève de vives inquiétudes pour la population qui s'interroge sur la protection de sa vie privée. L'inquiétude reste présente du côté des entreprises privées qui elles, ont inlassablement besoin d'alimenter leur système d'IA avec des données de plus en plus complexes et précises. Le défi du gouvernement est de trouver une forme de gouvernance de l'IA qui répond au mieux aux besoins des différents acteurs concernés.

L'objectif principal de cette brève politique proposée par notre Groupe de travail ponctuel sur l'utilisation de l'IA et des données communes est donc de *fournir une méthode de travail et de réflexion* afin de répondre adéquatement à ces interrogations, en s'assurant de considérer les intérêts distincts de la population et du secteur privé.

#### Intérêts des acteurs impliqués

Population	<ul style="list-style-type: none"><li>- Protéger ses données personnelles</li><li>- Être rassuré par l'indépendance des instances faisant usage de ses données</li></ul>
Gouvernement	<ul style="list-style-type: none"><li>- Assurer la protection du public</li><li>- Stimuler la croissance économique du secteur technologique</li></ul>
Privé	<ul style="list-style-type: none"><li>- Connaître une croissance économique stable</li><li>- Développer et améliorer les connaissances concernant l'IA</li></ul>

Problèmes soulevés par ces intérêts distincts :

- Politique actuelle mal adaptée
- Inquiétudes des citoyens
- Flou dans les responsabilités lors d'incidents
- Absence de méthodes pour gérer les problèmes liés à l'IA

#### Plusieurs pistes de solutions

- Création d'un organisme de régulation indépendant
- Favorisation d'une responsabilité partagée relativement aux données communes
- Ajustement des lois et réglementations en vigueur afin de les adapter aux nouvelles réalités technologiques



## Introduction

L'évolution rapide de la technologie et la science entourant l'intelligence artificielle (IA) exposent certaines brèches dans les politiques et les réglementations actuelles qui concernent ce secteur de développement. Bien que certaines lois soient déjà en place, la vitesse de l'appareil public peut difficilement rattraper celle de la croissance technologique, et les règles en place deviennent rapidement inadaptées.

### Les inquiétudes de la population

Cette inadéquation des politiques publiques en matière d'encadrement de l'IA et de l'utilisation des données servant à sa croissance inquiète la population québécoise. Une consultation récente initiée par un groupe d'experts composé, entre autres, de gens de l'Université de Montréal, de l'Université McGill et de l'Institut de valorisation des données (IVADO), a permis d'identifier certaines préoccupations clés des citoyens vis-à-vis les enjeux actuels concernant notamment :

- La responsabilité face aux données et à l'IA ;
- La protection de la vie privée des individus ;
- La valeur marchande des données partagées ;
- Les risques de mise en place d'un monopole, et de conflits d'intérêts entre les différents acteurs touchés ;
- Ainsi que l'indépendance des différents acteurs qui interviennent dans le domaine.

### Les considérations face au secteur privé

La croissance économique québécoise étant de plus en plus liée aux nouvelles technologies, à l'exploitation des données et au développement de l'IA, il est important pour le gouvernement – malgré les inquiétudes soulevées – de ne pas laisser le secteur privé au dépourvu. L'accès aux données de la population est le carburant de cet important moteur économique qui doit manifestement être régulé, mais pour qui une marge de manœuvre doit être maintenue.

### Le défi de la gouvernance

Les trois acteurs généraux qui sont touchés par la problématique – le gouvernement, la population et le secteur privé – ressentent déjà les impacts du manque d'ajustement des politiques actuelles. L'exemple récent des piratages de données des grands services technologiques comme Facebook et Google – utilisés par des centaines de milliers de Québécois – expose bien ce phénomène : la population est ultimement la victime, blâmant à la fois le secteur privé et le gouvernement pour les dommages encourus.

Dans cette optique, il est donc primordial pour le député délégué à la transformation numérique gouvernementale de se positionner et de répondre aux questions et inquiétudes soulevées par la société civile et le secteur privé :

1. Qui est responsable face aux données communes ?
2. Quel appareil assure la protection du public ? Fonctionne-t-il adéquatement ?
3. Quel niveau de transparence est optimal ?
4. Comment peut-on stimuler la croissance économique dans le secteur des nouvelles technologies, tout en maintenant la confiance de la population face au processus ?

L'objectif principal de cette brève politique proposée par notre Groupe de travail ponctuel sur l'utilisation de l'IA et des données communes est donc de *fournir une méthode de travail et de réflexion* afin de répondre adéquatement à ces interrogations, en s'assurant de considérer les intérêts distincts de la population et du secteur privé.

## Données probantes et analyse

### Le problème de l'immigration discriminante

Le gouvernement canadien serait en expérimentation de l'utilisation de l'IA pour le tri des demandes de visa et d'immigration. Cette information a été publiée dans un rapport du Citizen Lab et relayée dans les médias canadiens.

L'une des auteurs du rapport a souligné que « sans garanties et mécanismes de surveillance appropriés, utiliser l'IA pour déterminer l'immigration et le statut de réfugié est très risqué ». Il est clair que l'utilisation de l'IA pourrait aider grandement à accélérer le tri et traitement de données d'immigration. Cela ne saurait cependant et en aucun cas outrepasser la décision discrétionnaire liée au droit à l'immigration qui ne saurait être laissé à une machine ou un algorithme ; algorithme qui plus est, n'est pas sans biais car fortement dépendant des considérations des personnes qui programment. Un arbitrage est à faire entre « rapidité dans le traitement des demandes » et « sélection discrétionnaire des dossiers ».

En outre, le rapport à l'immigration de l'équipe responsable de l'algorithme pourrait fortement déteindre sur le résultat final de sélection. D'une part, l'on pourrait avoir un processus de sélection moins rigoureux, voire laxiste, qui laisserait entrer sur le territoire québécois des personnes ne remplissant pas les conditions requises pour l'immigration. D'autre part, un processus plus strict pourrait ôter la possibilité aux personnes remplissant les conditions d'y accéder. Car rappelons-le, la sélection initiale aura été faite non pas par choix discrétionnaire, mais par une machine intelligente.

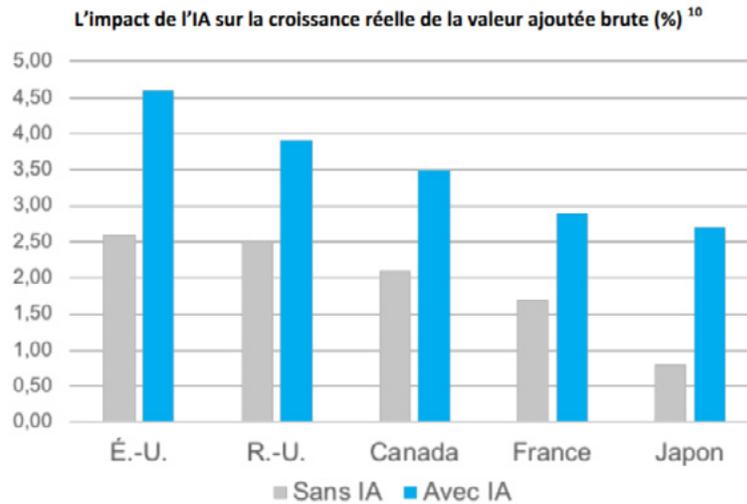
### La confidentialité des données

Les plus gros acteurs en matière de la gestion des données d'utilisateurs, dont font partie Google et Facebook, mettent en place des actions correctives dans leur gestion des données personnelles : ces données ont une valeur monétaire importante et ne doivent être utilisées que dans le cadre où elles ont été fournies par l'utilisateur du service. Les piratages sont récurrents comme l'ont montré les événements récents (500 000 comptes Google et 29 millions d'utilisateurs Facebook piratés) et la population est donc inquiète de l'utilisation qui peut être faite de ses données.

Les *Big Data* ont également un intérêt dans la santé des populations : des intelligences artificielles sont en développement pour aider au diagnostic médical et nécessitent donc des données sensibles, très personnelles et donc qui font partie d'un haut niveau de confidentialité. La société québécoise a donc évoqué ses inquiétudes quant à la gestion de ces données lors de l'étude qui a été menée sur une partie de la population lors de la rédaction de la Déclaration de Montréal pour l'IA responsable.

Les données personnelles sont donc classifiées selon leur confidentialité et sont donc disponibles à plusieurs niveaux. L'utilisateur sait à qui et dans quel but chaque organisme collecte des données à son sujet. Il devient donc important qu'un comité multisectoriel neutre se charge de veiller à ce que les possesseurs de ces grandes quantités de données les utilisent de manière éthique pour éviter que des centaines de milliers d'utilisateurs québécois de ces services voient leurs données diffusées sur le web.

## La stratégie du Québec en matière d'intelligence artificielle



De nombreux pays font du développement de l'IA une priorité majeure en investissant dans des organismes de recherche visant à progresser dans ce domaine, comme le montre ce graphique issu de la communication de « Économie, science et innovation Québec » à propos de l'essor de l'écosystème québécois en intelligence artificielle publié en mai 2018.

Le Québec ressent également ce besoin et prévoit des fonds à la recherche dans ce domaine, c'est pourquoi cette volonté doit alors s'étendre à la protection des populations et de leurs données sans pour autant empêcher les acteurs privés et universitaires de progresser dans leurs recherches et leurs innovations. D'après le projet de mettre le numérique au service du bien commun au Québec (disponible sur le site [economie.gouv.qc.ca](http://economie.gouv.qc.ca)) d'ici 5 ans, le Québec prévoit mettre à la disposition de la population une transformation numérique des municipalités qui se traduit par une collecte de données continue.

De plus, le Québec prévoit également que les citoyens pourront interagir de façon numérique avec les services de santé et sociaux d'ici les prochaines années. Toutes les données nécessaires à ces services nécessitent des données sensibles sur les citoyens et il est donc très important pour la province de se mobiliser pour protéger les utilisateurs contre une mauvaise utilisation de ces données dans tous les domaines confondus, que ce soit les services publics ou bien les entreprises privées de services.

Le développement du numérique, prisé par le Québec, doit donc faire évoluer les règlements relatifs à la protection des données des citoyens en faisant travailler les entreprises leaders de l'IA conjointement avec les pouvoirs publics et les intérêts des utilisateurs.

## Répercussions sur les politiques et recommandations

Nous constatons au final un manque d'adaptation des politiques actuelles face à l'IA et la gestion des données qui alimentent sa croissance. Cette problématique engendre un flou qui touche non seulement le gouvernement et la classe politique, mais également la population et le secteur privé.

Il existe actuellement un manque de clarté quant à la responsabilité de ces trois acteurs face aux données communes (1). Non seulement la *Loi sur la protection des renseignements personnels et des documents électroniques* (LPRPDE ; fédérale) et la *Loi sur la protection des renseignements personnels dans le secteur privé* (provinciale) ne répondent pas à ce manque, elles ne se sont pas non plus adaptées assez rapidement aux nouvelles réalités technologiques, ce qui risque de créer un doute dans la population quant à sa protection et la protection de ses données personnelles (2). Un certain degré de transparence (3) est évidemment requis afin de pallier cet aspect de la problématique, tout en gardant à l'esprit l'importance de ne pas entraver le développement économique du secteur technologique au Québec.

### Recommandations

Dans cette optique :

1. Nous emboîtons le pas du *Citizen Lab* de l'Université de Toronto et recommandons la création d'un organisme indépendant de régulation de l'utilisation des données communes. À la différence des chercheurs torontois, nous recommandons cependant que cet organisme soit de juridiction provinciale afin de prendre en considération les particularités de la population québécoise.
2. Nous favorisons une responsabilité partagée des données communes, alimentée par ce nouvel organisme. Ce dernier serait chargé, entre autres, de l'éducation de la population en matière de protection des données personnelles et de la surveillance du secteur privé quant à l'utilisation de ces données.
3. La création de cet organisme viendrait également répondre à la deuxième question soulevée par notre analyse de la situation, soit l'identité de l'appareil veillant à la protection du public. Il serait maintenant clair aux yeux de la population qu'une entité veille à ses intérêts, entre autres en s'assurant – à travers des recommandations émises à l'endroit du gouvernement – de l'adéquation des lois et réglementations concernant l'IA et la gestion des données.
4. Nous suggérons fortement que le nouvel organisme s'assure qu'un certain niveau de transparence soit respecté, autant par le gouvernement que par le secteur privé. Le type de données utilisées, leurs sources, les buts de leur utilisation et leur portée devraient être de nature publique.
5. Nous recommandons qu'un volet économique soit intégré à la mission de l'organisme afin que celui-ci soit constamment en phase avec le marché, s'assurant que les politiques mises en place permettent d'atteindre le parfait équilibre entre croissance et respect du milieu.
6. Finalement, nous recommandons que le nouvel organisme développe une méthode de réflexion permettant d'adapter les lois et réglementations concernant l'IA et les données communes aux changements rapides et fréquents inhérents au secteur des hautes technologies.

# CRÉDITS DU RAPPORT FINAL

## Le rapport de la Déclaration de Montréal IA responsable a été rédigé sous la direction de :

**Marc-Antoine Dilhac**, instigateur du projet et responsable du Comité d'élaboration de la Déclaration ; codirecteur scientifique de la coconstruction ; professeur au Département de philosophie de l'Université de Montréal ; chaire de recherche du Canada en Éthique publique et théorie politique ; directeur de l'axe Éthique et politique, Centre de recherche en éthique

**Christophe Abrassart**, codirecteur scientifique de la coconstruction, professeur à l'École de design et codirecteur du Lab Ville Prospective à la Faculté de l'aménagement de l'Université de Montréal, membre du Centre de recherche en éthique

**Nathalie Voarino**, coordonnatrice scientifique de l'équipe de la Déclaration, candidate au doctorat en bioéthique, Université de Montréal

### Coordination

**Anne-Marie Savoie**, conseillère, vice-rectorat à la recherche, à la découverte, à la création et à l'innovation de l'Université de Montréal

### Collaboration aux contenus

**Camille Vézy**, candidate au doctorat en communication, Université de Montréal

### Révision et édition

**Chantal Berthiaume**, gestionnaire de contenu et rédactrice

**Anne-Marie Savoie**, conseillère, vice-rectorat à la recherche, à la découverte, à la création et à l'innovation de l'Université de Montréal

**Joliane Grandmont-Benoit**, coordonnatrice de projets, vice-rectorat aux affaires étudiantes et aux études, Université de Montréal

### Traduction

**Rachel Anne Normand et François Girard**, Services linguistiques Révidaction

### Graphisme

**Stéphanie Hauschild**, directrice artistique

La rédaction de ce rapport n'aurait pu être possible sans les réflexions des citoyens, des professionnels et des experts ayant participé aux ateliers.

# NOS PARTENAIRES

Université   
de Montréal



CENTRE DE RECHERCHE EN ETHIQUE



ICRA  
Programme  
IA et  
société



Québec   
Fonds de recherche – Nature et technologies  
Fonds de recherche – Santé  
Fonds de recherche – Société et culture



