

Réflexions et proposition autour des valeurs et questions posées par *La déclaration de Montréal pour un développement responsable de l'intelligence artificielle*

Auteurs

Annick Hernandez, directrice, hernandez.annick@gmail.com

Guillaume Hernandez, étudiant en génie physique, g.t.b.hernandez@gmail.com

Raphaël Hernandez, designer industriel, raphael.hernandez@hotmail.ca

Méthodologie

- 1- nous avons lu les questions et déclarations pour chacune des valeurs inscrites dans la déclaration
- 2- nous avons participé à la présentation *UQAM - Conférence éthique de l'IA* le 15 février dernier
- 3- nous nous sommes rencontrés deux fois 4 heures pour discuter
- 4- nous avons élaboré le présent mémoire.

Structure du document

Pour chacune des valeurs :

- 1- Un avant-propos si nécessaire
- 2- Des recommandations sous forme de déclarations
- 3- Des questions qui ont émergées lors de nos discussions et qu'il nous semble important de transmettre, car elles peuvent donner des pistes pour d'autres débats ou des besoins en information/formation.

Cadre à l'intérieur duquel nous avons placé notre réflexion :

Sachant que les IA ne sont pas autonomes (*Voir section DÉFINITIONS*) et que ce sont des humains qui les créent, elles doivent être considérées comme des outils ou des technologies.

À propos de l'AUTONOMIE

Préambule : Étant donné que chaque citoyen va être en contact avec les IA au cours de sa vie, nous pensons que les citoyens devraient être initiés à la problématique des IA en fonction du rôle qu'il joue face à elles – créateur, utilisateur, bénéficiaire (*Voir section DÉFINITIONS*) - et ce dès le plus jeune âge.

Recommandations :

- La **création des IA devrait être encadrée par une série de règles formant une norme** (de la même façon que la création du vivant est encadrée). Cette norme devrait prendre en compte tous les constituants et les parties prenantes de l'IA : les données, les algorithmes, ceux qui entraînent les IA, ceux qui financent la création des IA, les bénéficiaires, etc.
- Le jour où, collectivement, nous considérerons que les IA pourront être dangereuses pour la société, il serait recommandable d'**encadrer son utilisation avec un permis**, comme c'est le cas avec un permis de conduire.

Questions que nous nous sommes posée pendant la réflexion autour de l'AUTONOMIE :

- *Qu'y a-t-il de particulier aux IA qui les rendent si différentes des autres technologies*, car les machines ont toujours eu la capacité de remplacer les êtres humains.
Les IA nous font peur, car :
 - Elles pourraient avoir la capacité de nous surpasser. Cela pose la question de savoir si on peut imaginer un monde où les êtres humains ne seraient plus considérés comme des êtres singuliers, au sommet d'une hiérarchie. *Deviendrons-nous obsolètes?*
 - Elles pourraient devenir indiscernables. Au moment où les machines ne pourront plus être distinguées des êtres humains (enveloppe qui ressemble à un corps humain, simulation d'émotions, etc.) faudra-t-il les considérer comme des êtres humains?
- Au-delà de la question *Qu'est-ce qu'un être humain?* Cela pose la question de *Comment les humains se perçoivent-ils?*

À propos du BIEN ÊTRE

Recommandations :

1. Au même titre que toutes les technologies, **les IA devrait viser le bien-être de tous les êtres humains.**
2. Le jour où il ne sera plus possible de discerner les IA des humains, toutes les règles qui s'appliquent aux humains devront leur être appliquées.
3. Chaque citoyen qui reçoit une décision d'une IA devrait avoir le droit de demander que cette décision soit revue par un humain (comme dans la régulation européenne).

Questions que nous nous sommes posée pendant la réflexion autour du BIEN ÊTRE :

- *Qu'est-ce qu'un être humain?*
Nous pensons que l'on ne peut réduire l'être humain à une série de choix logiques et que l'incorporation (être dans un corps) de tous les mécanismes biologiques, logiques, etc. apportent au moins une autre dimension que, pour le moment, la science ne comprend pas encore (il y a bien des invisibles à découvrir dans ce domaine).
- *Qu'est-ce qu'avoir du cœur?*

À propos de la JUSTICE

Préambule : Compte tenu que les performances des IA sont pilotées par les données qui servent à les entraîner, celles-ci seront plus performantes face aux groupes qui produisent une plus grande quantité de données. Ainsi, les IA sont des outils qui vont exacerber les inégalités déjà présentes dans les sociétés et entre les sociétés.

Recommandations :

1. Nous préconisons **la création d'un fonds et de programmes** qui vont permettre de générer des données pour tous les groupes possibles (et pas uniquement dans les pays occidentaux), et ce, dans l'idée de diversifier l'origine des données et de tenir compte de la pluralité des groupes humains sur la planète.
2. Nous préconisons la création d'un **organisme régulateur** dont la tâche serait d'assurer au citoyen une bonne compréhension des décisions prises par les IA. Ce qui inclut de rendre explicite l'origine des données, le fonctionnement des algorithmes et l'intention de ceux qui créent les IA ou qui les commercialisent.
3. Il doit être indiqué explicitement quand une décision est prise par une IA.
4. Il doit être permis à un bénéficiaire d'une décision prise par une IA de demander qu'un être humain vérifie la décision, comme c'est le cas dans la législation européenne actuelle.

Question que nous nous sommes posée pendant la réflexion autour la JUSTICE :

- *Quelle forme pourrait prendre l'aide à la création d'IA qui tiennent compte de la diversité?*

À propos de la VIE PRIVÉE

Préambule : Étant donné qu'il est difficile, voire impossible, de se soustraire au consentement à la collecte des données des compagnies, pour pouvoir utiliser leurs outils, il nous semble crucial de mettre l'accent sur des mesures qui pourraient permettre aux utilisateurs des outils, de mieux comprendre à quoi ils consentent, comment les compagnies vont utiliser les données et la valeur des données recueillies.

Recommandations :

1. La politique de consentement à la collecte des données devrait être intelligible et sa présentation uniforme entre les compagnies, de façon à ce que l'utilisateur s'y retrouve. Elle pourrait se présenter sous la forme d'une série de pictogrammes comme les licences de propriétés intellectuelle Creative Commons. Elle devrait être émise et réévaluée régulièrement par un organisme indépendant des enjeux économiques des compagnies.
2. Il devrait y avoir une **transparence sur la valeur des données**. Les utilisateurs de services qui collectent des données devraient avoir le droit en tout temps de connaître la valeur monétaire de ces données. Ceci pourrait prendre la forme d'une loi d'accès à la valeur des données numériques ou sous la forme d'un **indice de valeur des données**, comme à la bourse.
3. Il doit être interdit aux compagnies qui collectent des données, de vendre les données nominatives (*Voir section DÉFINITIONS*) des citoyens. Cela devrait être illégal et passible de justice. Les compagnies conserveraient cependant le droit de vendre les tendances issues des données récoltées.
4. Seul un gouvernement, et sous l'ordonnance d'un mandat, pourrait accéder aux données nominatives que les compagnies collectent.

Questions que nous nous sommes posée pendant la réflexion autour la VIE PRIVÉE :

- *Comment créé un espace numérique privé ?*
Un pendant à une maison, qui est notre espace physique privé, à l'intérieur duquel nos enfants pourraient circuler en toute confiance; un espace où nous pourrions nous sentir protégés.
- *Comment dépasser l'antagonisme entre le modèle d'affaire des compagnies basé sur la collecte des données et l'idée qu'elles sont un bien commun?*
Ceci pourrait par exemple impliquer que les gouvernements devraient avoir accès gratuitement aux tendances des données pour pouvoir offrir de meilleurs services à leurs citoyens.

À propos de la CONNAISSANCE

Préambule : Les IA devraient aider à contrer le phénomène des fausses nouvelles et non y participer de façon incontrôlée.

Recommandations :

1. Il faudrait entraîner des IA à démasquer des fausses nouvelles.
2. Pour pallier aux fausses nouvelles, nous proposons de mettre en place
 - a. **Un indicateur de factualité** qui présenterait leur % de vérité factuelle.
 - b. **Un indicateur de vérifiabilité** qui indiquerait si la donnée est vérifiable ou non

Questions que nous nous sommes posée pendant la réflexion autour la CONNAISSANCE :

- Nous pensons qu'étant donné le potentiel disruptif des IA, il faudrait repenser le système d'éducation pour y intégrer l'éducation à l'IA. Cependant, nous ne sommes pas arrivés à nous entendre sur la forme que pourrait prendre cette éducation à l'IA : nous avons évoqué, entre autres, l'éthique, la philosophie, la logique, et la pensée critique.
- Nous pensons que l'encadrement éthique de la recherche sur l'IA devrait être étudié, ne serait-ce que pour vérifier si les mesures en place actuellement, les certificats d'éthiques par exemple, répondent adéquatement aux nouveaux enjeux.

DÉFINITIONS

Autonome:	Une technologie est considérée autonome si elle ne nécessite, d'aucune manière, l'aide de l'humain pour établir ses objectifs, ses procédés et sa subsistance.
Bénéficiaire d'une IA :	Le bénéficiaire d'une IA est une personne à qui est destiné un service rendu par une IA.
Données nominatives :	« Informations nominatives », donnant « une définition du profil ou de la personnalité de l'intéressé » (article 2), et « qui permettent sous quelque forme que ce soit, directement ou non, l'identification des personnes physiques auxquelles elles s'appliquent » (article 4). <i>Loi informatique et libertés originale du 6 janvier 1978.</i>
Créateur d'une IA :	Un créateur d'une IA est une personne – physique ou morale - impliquée dans le développement de l'IA : informaticiens, entraîneurs, analystes, etc.
Utilisateur d'une IA:	L'utilisateur d'une IA est une personne – physique ou morale – qui se sert d'une IA pour offrir un service, monétarisé ou non.