



tic&société

Vol. 15, N° 1 | 1er semestre 2021 | 2021

Logique algorithmique et reproduction sociale: les médiations sociales saisies par les algorithmes

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

Artificial Intelligence Ethics and the Social Ubiquity of Information and Communication Technologies: Rethinking Ethical Issues in our Information Societies

Ética de la inteligencia artificial y ubicuidad social de las tecnologías de la información y de la comunicación: ¿cómo pensar los retos éticos de la IA en nuestras sociedades de la información?

Frédéric BRUNEAULT et Andréane SABOURIN LAFLAMME



Édition électronique

URL : <https://journals.openedition.org/ticetsociete/5999>

DOI : 10.4000/ticetsociete.5999

Éditeur

Association ARTIC

Édition imprimée

Pagination : 159-189

Référence électronique

Frédéric BRUNEAULT et Andréane SABOURIN LAFLAMME, « Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ? », *tic&société* [En ligne], Vol. 15, N° 1 | 1er semestre 2021 | 2021, mis en ligne le 20 mai 2021, consulté le 02 juin 2021. URL : <http://journals.openedition.org/ticetsociete/5999> ; DOI : <https://doi.org/10.4000/ticetsociete.5999>

Licence Creative Commons

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

Frédéric BRUNEAULT

Andréane SABOURIN LAFLAMME

frederick.bruneault@clairendeau.qc.ca

andreane.sabourin-laflamme@clairendeau.qc.ca

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information?

Frédéric BRUNEAULT

Frédéric Bruneault enseigne la philosophie au Collège André-Laurendeau et est professeur associé à l'École des médias de l'Université du Québec à Montréal (UQAM), où il participe aux travaux du Groupe de recherche sur l'information et la surveillance au quotidien (GRISQ) dans le cadre des activités du Centre de recherche interuniversitaire sur la communication, l'information et la société (CRICIS). Ses recherches portent sur les implications philosophiques des innovations technologiques et les enjeux éthiques liés au développement technologique, notamment celui de l'éthique de l'intelligence artificielle. Spécialiste en éthique fondamentale, il s'intéresse aux théories éthiques actuelles qui traitent de ces questions.

Andréane SABOURIN LAFLAMME

Andréane Sabourin Laflamme enseigne la philosophie au Collège André-Laurendeau. Elle s'intéresse aux enjeux éthiques liés au développement technologique et au droit, notamment dans le contexte de la révolution numérique.

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information?

L'intelligence artificielle (IA) occupe une place importante dans les discussions actuelles sur les technologies de l'information et de la communication (TIC) et les nouveaux modes de gouvernance algorithmique. Elle est une dimension incontournable de ce que seront à l'avenir les médiations sociales et les modes de reproduction de nos sociétés de l'information. Si plusieurs travaux en éthique de l'intelligence artificielle (EIA) se penchent sur des questions éthiques propres à certains domaines d'expertise, ces réflexions éthiques demeurent souvent confinées à des domaines restreints d'application, sans prendre en compte les enjeux éthiques globaux dans lesquels elles s'inscrivent. Nous nous proposons donc de clarifier les principales approches en EIA, leurs présupposés philosophiques et les caractéristiques propres à chacune, afin d'identifier la démarche la plus prometteuse pour élaborer une réflexion éthique sur le déploiement de l'IA dans nos sociétés.

Éthique ; intelligence artificielle ; transhumanisme ; libéralisme ; information

Artificial intelligence ethics and the social ubiquity of information and communication technologies : rethinking ethical issues in our information societies

Abstract: Artificial intelligence (AI) plays an important role in current discussions on information and communication technologies (ICT) and new modes of algorithmic governance. It is an unavoidable dimension of how social mediations and modes of reproduction will evolve in our information societies. While several works in artificial intelligence ethics (AIE) address ethical issues specific to certain areas of expertise, these ethical reflections often remain confined to narrow areas of application, without considering the global ethical issues in which they fit. We therefore propose clarifying the main approaches to AIE, their philosophical assumptions and the specific characteristics of each, in order to identify the most promising approach to developing ethical reflections on the deployment of AI in our societies.

Keywords: Ethics; artificial intelligence; transhumanism; liberalism; information

Ética de la inteligencia artificial y ubicuidad social de las tecnologías de la información y de la comunicación: ¿cómo

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

pensar los retos éticos de la IA en nuestras sociedades de la información?

Resumen: La inteligencia artificial (IA) juega un papel importante en los debates actuales sobre las tecnologías de la información y la comunicación (TIC) y sobre los nuevos modos de gobernanza algorítmica. Constituye una dimensión inevitable de las mediaciones sociales y de los modos de reproducción de nuestras sociedades de la información en el futuro. Si bien varios trabajos en ética de la inteligencia artificial (EIA) abordan cuestiones éticas específicas de ciertas áreas de especialización, estas reflexiones éticas a menudo permanecen confinadas en áreas de aplicación limitadas, sin tener en cuenta las cuestiones éticas globales. Por lo tanto, proponemos aclarar los principales enfoques de la EIA, sus supuestos filosóficos y las características específicas de cada uno, a fin de identificar el enfoque más prometedor para desarrollar una reflexión ética sobre el despliegue de la IA en nuestras sociedades.

Palabras-clave: Ética; inteligencia artificial; transhumanismo; liberalismo; información

L'intelligence artificielle (IA) occupe une place importante dans les discussions actuelles sur les technologies de l'information et de la communication (TIC) et les nouveaux modes de gouvernance algorithmique. Étant à la fine pointe des innovations technologiques, l'IA est régulièrement présentée, en particulier l'apprentissage machine et l'apprentissage profond, comme la dimension incontournable de ce que seront à l'avenir les médiations sociales et les modes de reproduction de nos sociétés de l'information. Face à ces changements que plusieurs qualifient de *révolution numérique*, il apparaît essentiel de se demander quelles seront les conséquences de cette révolution et comment nous pouvons faire une utilisation responsable de l'IA. Des initiatives, comme la *Déclaration de Montréal pour un développement responsable de l'IA*, cherchent à réagir au fait que bon nombre d'intervenants et d'intervenantes, provenant des secteurs techniques, de la mise en marché ou du développement de l'IA, se penchent sur des questions éthiques propres à leurs domaines d'expertise, questions devenues incontournables dans toutes les sphères d'application de ces technologies.

Le problème est que ces réflexions éthiques portent souvent sur des domaines restreints d'application et peinent à prendre en compte les enjeux éthiques globaux dans lesquels elles s'inscrivent. Face à la polysémie du terme « éthique », bon nombre de travaux en éthique de l'intelligence artificielle (EIA) cherchent à traiter différents niveaux de réflexion éthique – l'éthique professionnelle, l'éthique appliquée, l'éthique normative et la méta-éthique –, la plupart du temps sans les distinguer. De plus, lorsqu'une approche en éthique normative n'est pas tout bonnement présumée, sans plus de discussion, on remarque bien souvent un amalgame de positions éthiques normatives, pouvant par ailleurs être incompatibles entre elles. Bien que louables devant l'urgence d'agir qu'impose le rythme accéléré du développement des systèmes d'IA et qu'elles soient préférables à l'absence de considérations éthiques qui règne encore trop souvent dans les milieux techniques et économiques, de telles démarches restent habituellement en périphérie des discussions éthiques fondamentales, sans remonter aux racines théoriques derrière de telles positions, ce qui favorise une certaine confusion. D'ailleurs, si les dimensions techniques du développement de l'IA doivent bien entendu être au cœur d'une réflexion éthique satisfaisante sur ces technologies, ces aspects techniques ne peuvent ni constituer à eux seuls une base suffisante à l'élaboration d'une EIA, ni se substituer à une démarche philosophique permettant de situer une telle EIA dans les débats qui ont jalonné l'histoire de la philosophie.

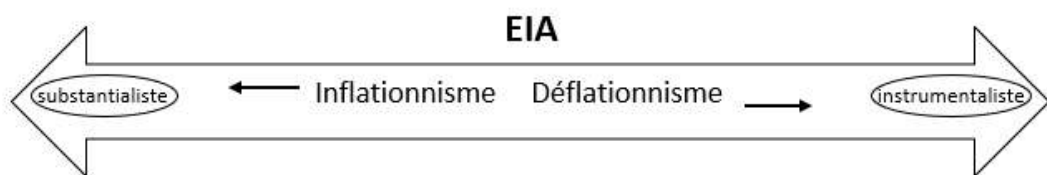
Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

Nous examinerons dans un premier temps les ressources argumentatives mobilisées dans le débat en EIA en vue de proposer un spectre argumentatif permettant de situer les principales positions qui y sont défendues, et ce en fonction de deux positions types aux extrémités du spectre, en montrant les présupposés philosophiques, les caractéristiques et les écueils propres à chacune de ces positions (1). Nous discuterons plus précisément de la perspective *inflationniste*, tendant vers ce que nous nommerons la position *substantialiste* (1.1), et ensuite de la perspective *déflationniste*, menant vers la position que nous qualifions d'*instrumentaliste* (1.2). Après avoir discuté des limites de cette opposition et des présupposés théoriques éthiques associés au spectre argumentatif (1.3), nous pourrions examiner la position *informationaliste* (2) qui cherche à élaborer une approche éthique complémentaire aux théories éthiques classiques. Nous serons alors en mesure de souligner les avantages de cette troisième voie et d'indiquer les pistes de réflexion qui émergent d'une telle perspective en vue du développement d'une réflexion poussée sur l'EIA.

1. Le débat en EIA

Nous reprenons la distinction proposée par Jocelyn Maclure (2019) entre perspectives inflationniste et déflationniste pour définir le spectre argumentatif permettant de situer les différentes positions défendues dans le débat en EIA. La perspective inflationniste met l'accent sur les conséquences à long terme du développement de l'IA en réfléchissant notamment aux enjeux et aux risques associés à l'émergence d'une IA générale (IAG). À l'opposé, la perspective déflationniste cherche plutôt à penser les enjeux éthiques spécifiques qui sont actuellement associés au développement des systèmes d'IA réels et à leur implantation dans nos sociétés. Loin d'être des catégories fixes permettant de classer les différentes positions dans le débat en EIA une fois pour toutes, il nous apparaît plutôt qu'il s'agit de concepts relationnels, en ce sens que, tout le long du spectre argumentatif en EIA, une position particulière sera inflationniste ou déflationniste comparativement à une autre position à partir de laquelle nous l'évaluons. Voilà pourquoi il nous semble nécessaire d'aller au-delà de cette distinction, bien qu'elle soit utile, pour qualifier les deux positions ultimes à chacun des bouts du spectre, définissant ainsi l'univers des possibles des positions en EIA. Nous soutiendrons donc qu'au summum de l'*inflationnisme* se trouve la position *substantialiste*, faisant du développement technologique en général

et de celui de l'IA en particulier une réalité substantielle, indépendante des décisions et des interactions humaines, bref la thèse du déterminisme technologique. À l'autre bout du spectre, au summum du *déflationnisme*, nous verrons la position *instrumentaliste*, qui stipule que le développement de l'IA, comme celui de la technologie en général, est une entreprise purement instrumentale, instituée et contrôlée par les humains qui peuvent donc l'infléchir comme bon leur semble, bref la thèse de la neutralité technologique. Si la vaste majorité des positions défendues en EIA se trouvent quelque part entre ces deux extrêmes (les défenseurs de ces positions sans nuances étant finalement peu nombreux), il est pourtant primordial de bien saisir les positions substantialiste et instrumentaliste, question de repérer leur influence dans les positions plus nuancées. Il est à noter que nous n'utiliserons donc pas, dans cette première partie de notre texte, les distinctions usuelles en théorie éthique entre les approches déontologique, conséquentialiste et celle fondée sur l'éthique des vertus. Nous pensons que ces distinctions sont inappropriées pour saisir les différences argumentatives en EIA, en ce sens qu'il est non seulement possible, mais dans les faits assez courant de retrouver des positions, bien qu'inspirées de théories différentes, à des endroits similaires sur le spectre argumentatif. Pour le dire de manière schématique, chaque position défendue dans le débat en EIA (P) est constituée d'une composante provenant des théories éthiques (E) et d'une composante associée à une certaine interprétation de la philosophie de la technologie (T), de telle sorte que $P = E + T$. Il nous apparaît que le facteur décisif qui détermine le positionnement d'une posture dans le spectre argumentatif en EIA est la composante T, bien plus que la composante E, de telle sorte qu'une position totalement différente en EIA (P') suppose une interprétation différente en philosophie de la technologie, mais elle pourrait très bien partir d'une même approche éthique : $P' = E + T'$. Nous reviendrons sur les théories éthiques classiques en 1.3.



1.1 Inflationnisme et position substantialiste

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

La perspective inflationniste en EIA mène à ce que nous nommons la position substantialiste. Cette perspective connaît depuis les deux dernières décennies une popularité grandissante et soutient que le développement de l'IA annonce le dépassement des conditions biologiques de l'existence humaine. Cet avènement du prochain stade de l'évolution humaine est en marche et les sociétés doivent s'y adapter. Nous verrons (1.1.1) quelles sont les positions éthiques sur l'IA liées à la perspective inflationniste, (1.1.2) quels sont les présupposés philosophiques de celle-ci et (1.1.3) quelles sont les principales caractéristiques de la position substantialiste en EIA.

1.1.1 Positions éthiques inflationnistes sur l'IA

La perspective inflationniste en EIA avance que l'apparition de l'IA est l'aboutissement nécessaire d'un développement technologique exponentiel, et ne constitue rien de moins que l'événement le plus significatif de l'évolution de l'espèce humaine. Force est de constater que le nombre de domaines où l'IA arrive à rivaliser avec l'intelligence humaine augmente à un rythme fulgurant, et que les progrès en apprentissage profond permettent de croire qu'une machine pourra être en mesure d'accomplir des tâches qui jusqu'ici s'étaient montrées trop intuitives pour être générées artificiellement – comme en font foi les progrès remarquables en traduction et en reconnaissance d'images, par exemple. Les partisans de la perspective inflationniste partagent la conviction que seul le temps nous sépare de l'apparition d'une *Human Level Machine Intelligence* (HLMI), c'est-à-dire d'une IA qui pourrait « exercer la plupart des professions humaines au moins aussi bien qu'un être humain moyen » (Bostrom, 2017, p. 37). Compte tenu des avantages incontestables de la machine sur le cerveau biologique (notamment l'espace mémoire et la rapidité de traitement de l'information), l'éducation de la première intelligence artificielle générale (IAG) s'effectuera à un rythme sans commune mesure avec l'éducation d'un cerveau biologique.

Plusieurs chercheurs observent une certaine prudence quant à ces prévisions, prudence probablement motivée par les nombreuses prédictions un peu trop enthousiastes faites, avant l'hiver de l'IA des années 1980, par leurs prédécesseurs, défenseurs du programme d'IA symbolique souvent nommé « *Good Old-Fashioned Artificial Intelligence* » (GOFAI). Les tenants de la perspective inflationniste en EIA prétendent malgré tout que le passage de l'IA faible à une IA forte ou une IAG est tout simplement

une nécessité. Les questions qui s'imposent dès lors ne sont pas de l'ordre de savoir si, mais plutôt par quels moyens technologiques, à quelle vitesse, sous quelle forme et quand une première IAG, dont le développement sera indépendant de toute intervention humaine, émergera

Qu'il participe de la singularité (Kurzweil, 2007) ou d'une explosion d'intelligence (Bostrom, 2014), cet événement constituera une révolution aussi importante que la sédentarisation ou la révolution industrielle, et catapultera l'humanité dans une phase de son évolution dans laquelle l'humain transcendera les limites de la biologie pour embrasser une réalité où la frontière entre le vivant et le non-vivant, aussi bien qu'entre le naturel et l'artificiel, sera à jamais effacée (Tegmark, 2017). Considérant le caractère à la fois inéluctable et radical de cette profonde mutation, certains envisagent la possibilité que cette « superintelligence » présente un risque existentiel pour l'espèce humaine (Bostrom, 2014 ; Russell, 2019). Ces préoccupations ont d'ailleurs fait l'objet d'une lettre ouverte signée par des acteurs parmi les plus influents dans le domaine de l'IA, qui considèrent que le risque d'une catastrophe planétaire associée à l'IA est significatif (<https://futureoflife.org/ai-open-letter/>). Au sein de la perspective inflationniste en EIA, les réflexions éthiques sont, à toutes fins pratiques, exclusivement orientées vers les conditions de l'émergence d'une IAG et les risques qui y sont associés.

1.1.2 Présupposés philosophiques inflationnistes

Bien qu'il n'y ait pas de consensus quant à l'issue de cette révolution et que s'écrivent à cet égard tant des récits dystopiques que des fables utopiques, il est manifeste que les postures que l'on peut associer à la perspective inflationniste en EIA reposent sur des présupposés philosophiques que partage le transhumanisme, une idéologie qui « reconnaît et anticipe les modifications radicales dans les conditions de notre existence provenant de différentes sciences et technologies » (More, 1990, p. 6, notre traduction¹). Il existe certes bien des oppositions entre les postures que l'on peut associer à la perspective inflationniste, notamment quant à la valeur morale du changement de paradigme qu'opérera l'émergence d'une IAG. En effet, alors que certains auteurs qui défendent une posture inflationniste appellent de leurs vœux le développement d'une IAG, d'autres le redoutent. Souvent présenté comme un humanisme ayant désormais les moyens de ses ambitions, le transhumanisme

¹ «recognizes and anticipates radical alterations in the conditions of our existence resulting from various sciences and technologies»

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

conçoit la technologie comme un outil qui permettra à l'humain d'augmenter considérablement ses capacités physiques et cognitives, le libérant des contraintes biologiques qui le limitent dans sa forme actuelle. Les avancées en robotique et en nanotechnologie nous permettent d'espérer pouvoir remplacer certaines parties de notre corps par des organes artificiels beaucoup plus performants, notre corps biologique étant très imparfait, décrit comme un « *désordre sanglant de matière organique*² » par nul autre que Marvin Minsky, tel que rapporté par Turkle (2005). Non sans rappeler une forme virulente d'idéalisme cartésien, le transhumanisme aspire ultimement à une entière libération du corps, si bien que la possibilité de télécharger le contenu de notre esprit sur un autre *hardware* que ce support biologique désuet qu'est le cerveau est sérieusement envisagée par plusieurs. L'humanité telle que nous la connaissons aujourd'hui ne serait que la première étape d'une évolution que nous avons largement subie jusqu'à présent, mais que nous aurions bientôt les moyens de contrôler. Dans cette perspective, l'IA constitue le moyen par excellence pour atteindre cet objectif, et l'apparition d'une IA forte pourrait très bien coïncider avec notre passage à l'humanité 2.0 (Kurzweil, 2007) ou à la vie 3.0 (Tegmark, 2017).

1.1.3 Caractéristiques de la position substantialiste

La perspective inflationniste en EIA culmine avec la position substantialiste, qui repose sur une conception déterministe de la technologie selon laquelle le développement technologique est nécessaire, inévitable et irréversible. La conjoncture du développement technologique implique que « si les efforts de développement scientifique et technologique ne cessent pas, tout ce qu'il est possible de faire en matière de développement technologique sera complètement fait » (Bostrom, 2017, p. 476). Spectateurs devant l'inéluctable, nous devrions donc nous résoudre à tenter de prévoir les risques d'une IAG pour l'humanité et à anticiper les moyens qui pourraient permettre de s'en prémunir. C'est néanmoins avec une certaine urgence qu'il faut s'attaquer au problème de contrôle (Bostrom, 2017), parce que si l'IA forte pourra sans doute résoudre, ou du moins nous aider à résoudre, plusieurs problèmes qui semblent insolubles aujourd'hui (changements climatiques, pauvreté, maladies, etc.), il reste que l'enjeu principal lié à l'apparition d'une intelligence non biologique, radicalement supérieure à l'intelligence humaine et pouvant s'améliorer elle-même, est de s'assurer que les actions de cette IA forte soient

² « *bloody mess of organic matter* »

conformes aux intérêts de l'humanité et correspondent aux valeurs humaines. Dans la mesure où il est admis que ces mutations pourraient nous mener tant à notre salut qu'à notre perte, la question qui occupe les substantialistes – question dont le caractère sensationnel peut expliquer l'attention qu'elle a reçue dans les médias au cours des dernières années – est de savoir s'il est légitime de céder à une certaine panique face aux dangers liés à l'apparition de l'IA forte, ou si elle ne constituerait pas, au contraire, le moyen par excellence pour nous libérer des contraintes qui jusqu'ici ont caractérisé la condition humaine. Le paradigme dans lequel s'inscrit l'argument substantialiste est conséquemment celui que nous proposons d'appeler « panique ou délivrance » (PoD). Considérant soit les risques existentiels, soit les bénéfices fabuleux de l'IA pour l'humanité, l'argument substantialiste de l'EIA évalue le déploiement de l'IA dans cette perspective PoD.

1.2 Déflationnisme et position instrumentaliste

La perspective déflationniste se caractérise, à première vue, par une approche diamétralement opposée à celle de la perspective inflationniste dans le débat en EIA. En effet, les déflationnistes soutiennent que les réelles préoccupations éthiques liées à l'IA ne relèvent que de l'utilisation qu'on en fait, le programme GOFAI ayant été abandonné et l'IA forte n'étant donc pas une possibilité technique (pour une réflexion récente autour du GOFAI, voir Levesque, 2017). Les innovations technologiques liées à l'IA, telles qu'elles se développent actuellement, relèvent plutôt de l'IA faible, ce qui fait que les craintes suscitées et les conséquences entrevues par la perspective inflationniste en EIA ne sont pas justifiées. La perspective déflationniste, qui tend vers la position instrumentaliste, soutient qu'il y a bel et bien des enjeux éthiques liés à ces innovations technologiques, mais d'une toute autre nature. Nous discuterons maintenant (1.2.1) des positions éthiques sur l'IA défendue à partir de la perspective déflationniste, (1.2.2) des présupposés philosophiques derrière cette perspective et (1.2.3) des caractéristiques de la position instrumentaliste en EIA.

1.2.1 Positions éthiques déflationnistes sur l'IA

Les positions éthiques sur l'IA associées à la perspective déflationniste reposent essentiellement sur le constat que l'IA n'existe pas et n'existera pas dans un avenir qu'il est possible d'entrevoir. Après l'hiver de l'IA des années 1980, une bonne partie

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

du renouveau et de la renaissance des programmes en IA repose essentiellement sur le fait que l'idée de reproduire l'esprit humain *in silico* a tout simplement été abandonnée (Domingos, 2015 ; Floridi, 1999). L'intelligence artificielle, telle qu'elle se développe aujourd'hui, ne se contente que d'automatiser des processus qui autrefois étaient réservés à la mise en œuvre de processus cognitifs humains, mais qui peuvent désormais être accomplis par des machines, étant donné le développement exponentiel des TIC. L'EIA doit donc, dans cette perspective, prendre acte de cette caractéristique du développement de l'IA et évaluer les enjeux éthiques liés au déploiement de l'IA faible.

Il s'agit d'une position qu'on retrouve très clairement exprimée, par exemple, dans la perspective déflationniste de l'EIA défendue par Jocelyn Maclure (Maclure, 2019 ; Maclure et Saint-Pierre, 2018). Si le renouveau de l'IA s'appuie sur le développement de l'IA faible, il s'ensuit que les spéculations liées à l'émergence d'une IAG n'ont pas lieu d'être, puisque, comme Maclure le dit, « la perspective que les machines dotées d'IA dominent le monde et anéantissent l'humanité nous apparaît donc trop improbable pour être centrale dans la réflexion éthique et juridique » (Maclure et Saint-Pierre, 2018, p. 752). C'est cette même position déflationniste sur l'EIA qui est sous-entendue dans les démarches qui visent à fournir les lignes directrices d'une éthique professionnelle pour encadrer le développement de l'IA (Boddington, 2017), dans la *Déclaration de Montréal pour un développement responsable de l'intelligence artificielle* (declarationmontreal-iaresponsable.com) et dans les 84 initiatives visant à établir des lignes directrices éthiques en IA répertoriées par Jobin, Ienca et Vayena (2019). Cette approche de l'EIA est également sous-entendue dans une démarche comme celle du *MIT Moral Machine*, qui proposait une plateforme « qui a compilé 40 millions de décisions, dans 10 langues, de la part de millions de personnes dans 233 pays et territoires » (Awad *et al.*, 2018, p. 59, notre traduction³) sur le comportement que devraient adopter des voitures autonomes dans certaines situations où une collision est inévitable, compte tenu que les personnes affectées par la collision présenteraient des caractéristiques socio-économiques variées (une version repensée du fameux dilemme du tramway). La position défendue en EIA par Yoshua Bengio et Sascha Luccioni (2019), celle de Joanna J. Bryson (2019), ainsi que celle de Martin Gibert (2019) font également écho à la perspective déflationniste. L'idée est donc qu'il faut mettre en place des conditions réglementaires qui permettent de limiter, voire d'empêcher, une

³ «gathered 40 million decisions, in 10 languages, from millions of people in 233 countries and territories»

mauvaise utilisation de l'IA, ce qui permettrait de régler les problèmes éthiques engendrés par son déploiement.

1.2.2 Présupposés philosophiques déflationnistes

Les présupposés philosophiques derrière la perspective déflationniste en EIA sont essentiellement ceux du libéralisme classique, c'est-à-dire la défense des droits individuels et des libertés fondamentales qui y sont associées. L'IA est ainsi conçue comme un élément supplémentaire dont il faut tenir compte dans l'application du cadre libéral. Une telle perspective vise à rendre l'IA conforme aux principes du libéralisme classique (Hunyadi, 2018), en mettant l'accent sur le développement de stratégies d'application des devoirs liés à l'éthique individuelle libérale, l'idée étant par exemple de s'assurer que les systèmes d'IA ne sont pas discriminatoires, qu'ils respectent la vie privée et la liberté individuelle. Ces enjeux sont d'une grande importance et il est impératif de s'assurer que le déploiement de l'IA s'effectue dans le respect des droits individuels et des valeurs qui sont au fondement de nos sociétés démocratiques. Le problème avec la perspective déflationniste de l'EIA, lorsqu'elle est poussée à l'extrémité du spectre, est qu'une telle position présuppose un individu libéral désincarné, utilisateur de moyens et d'instruments pour atteindre ses fins, complètement en contrôle de ses instruments et de ses propres pensées, et qui se définit indépendamment des conditions sociales et matérielles (notamment technologiques) qui caractérisent sa situation de fait.

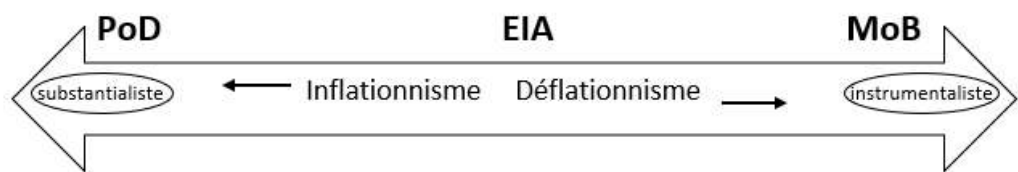
1.2.3 Caractéristiques de la position instrumentaliste

Poussée à l'extrême, la perspective déflationniste aboutit à une position que nous appelons instrumentaliste en EIA, c'est-à-dire à une position qui n'attribue de valeur au déploiement de l'IA qu'en fonction de l'utilisation qui en est faite. Nous proposons donc de nommer cette position « malveillance ou bienveillance » (MoB), en ce sens que l'IA n'est évaluée qu'en fonction des individus utilisateurs et de leurs fins. Une des difficultés des positions inspirées de cette perspective est par ailleurs de faire porter aux individus la responsabilité de se protéger eux-mêmes, en accordant par exemple une importance démesurée à la question du consentement. La position instrumentaliste en EIA s'appuie sur l'idée que la technologie, en elle-même, est foncièrement neutre et qu'elle n'est qu'indirectement chargée moralement, selon ce que les

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

individus utilisateurs décident d'en faire, le foyer de la valeur reposant exclusivement sur ces derniers. Pour résoudre les problèmes éthiques que suscite le développement des technologies de l'IA, il suffirait donc d'encadrer leur utilisation pour réduire le risque d'utilisations malveillantes.

Pour la position instrumentaliste au comble du déflationnisme, l'EIA est donc équivalente à une éthique appliquée de l'intelligence artificielle (EAIA). Cette EAIA doit mettre en lumière les sources de friction entre les nouvelles capacités technologiques offertes par l'IA et les cadres législatifs des sociétés libérales, par exemple les enjeux liés à la définition des sources de la responsabilité légale face au déploiement d'agents artificiels autonomes (pensons aux voitures autonomes, Awad *et al.*, 2018), les modifications importantes des relations de travail et de la gestion des ressources humaines (Maclure et Saint-Pierre, 2018), les problèmes générés en matière de protection de la vie privée (De Filippi, 2016) et de confidentialité (CRYPTO.QUÉBEC, 2018), les abus possibles que permettent ces technologies (Alter, 2018 ; Mayer-Schönberger, 2014), les biais qu'elles peuvent induire (Boutin, 2006), etc. La position instrumentaliste en EIA cherche d'abord à proposer des ajustements entre ces innovations technologiques et le cadre législatif et réglementaire propre aux droits individuels et aux libertés fondamentales, sans toutefois questionner ce cadre, ni le modifier, et sans chercher à évaluer les nouveaux rapports de force qui naissent de l'utilisation de l'IA dans nos sociétés.



1.3 Les théories éthiques dans le débat en EIA

Lorsqu'il est question des débats en EIA, il semble donc que les deux options soient la position PoD de la position substantialiste et celle MoB de la position instrumentaliste. Il nous apparaît toutefois que la différence entre ces deux approches ne réside pas dans leurs perspectives éthiques, mais bien dans leur évaluation du développement technologique. Comme mentionné plus haut, les

deux positions sont bien entendu différentes. Toutefois, ces deux positions (P) étant constituées de la conjonction d'une première composante liée à l'évaluation de l'IA en fonction des cadres de l'éthique normative (E) et d'une seconde composante liée à une interprétation de la philosophie de la technologie (T), chaque position étant ainsi la résultante de ces deux composantes ($P = E + T$), il semble que la différence provienne de la composante T de l'équation. Pour ces deux positions extrêmes en EIA, comme pour la plupart des positions intermédiaires sur le spectre argumentatif qu'elles délimitent, la composante E reste inchangée. En nous basant sur l'expression maintenant consacrée « *business as usual* », nous soutenons que nous avons ici affaire à des positions qui, malgré leurs divergences, défendent une même position éthique en EIA, à savoir celle que nous proposons de nommer « *ethics as usual* » (EaU), c'est-à-dire que le développement de l'IA ne demande fondamentalement pas de changement aux cadres de l'éthique normative classique. Tout se passe comme si les options de l'éthique normative classique (déontologisme, conséquentialisme et éthique des vertus) pouvaient tout simplement s'appliquer aux enjeux soulevés par cette nouveauté technologique qu'est l'IA, et ce même si les inflationnistes et les déflationnistes ne s'entendent pas sur la nature de ces enjeux. Or les limitations de ces approches théoriques en éthique sont bien connues et remontent bien au-delà des questions soulevées par le déploiement de l'IA. Il n'est pas possible de reprendre ces débats abondamment documentés depuis plusieurs décennies, voire quelques siècles. On peut toutefois affirmer qu'il nous apparaît très problématique de simplement reprendre une de ces approches théoriques pour l'appliquer directement aux questions éthiques liées au déploiement de l'IA, encore plus de compenser les limitations d'une approche en empruntant, de manière éclectique, indistinctement dans le bagage conceptuel d'autres théories. Il faut plutôt développer une approche cohérente basée sur des principes clairs pour penser convenablement la part de chacun des cadres de l'éthique normative classique dans l'élaboration d'une théorie normative qui puisse tenir compte des enjeux actuels liés au développement de l'IA et de l'ubiquité des TIC (pour une excellente synthèse de la question, voir van den Hoven 2010).

Nous pensons aussi que l'apparente opposition entre positions substantialiste et instrumentaliste aux extrêmes de l'inflationnisme et du déflationnisme s'érigent davantage en système de pensée qu'en réelle opposition, en ce sens que les positions extrêmes se renforcent mutuellement en fin de compte. Les instrumentalistes vont justifier leur position en reprochant aux substantialistes de se laisser distraire par des spéculations incertaines sur des scénarios

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

improbables, étant donné qu'ils partent d'une vision de l'IA (en particulier l'IAG et le GOFAI) qui n'est pas adéquate compte tenu du développement actuel de ces innovations technologiques (ce qui est tout à fait raisonnable). De leur côté, les substantialistes justifient leur démarche en reprochant (non sans raison) aux instrumentalistes de sous-évaluer l'aspect disruptif du développement de l'IA et de sous-évaluer les réflexions nécessaires à une approche satisfaisante des enjeux éthiques liés à l'IA en limitant l'EIA à une EAIA, ce qui amène les instrumentalistes à négliger les questions philosophiques fondamentales que ces innovations technologiques peuvent générer en termes de compréhension de soi, du monde et des relations avec les autres pour les individus qui les utilisent.

Voilà pourquoi, bien souvent, les positions éthiques entourant le déploiement de l'IA semblent être dans une impasse. D'un côté, nous avons une position qui présente des limitations importantes en raison d'une compréhension de l'IA qui ne correspond pas à la réalité du développement des systèmes d'IA. De l'autre côté, une position qui semble être contrainte, étant donné les présupposés philosophiques sur lesquels elle s'appuie, à ne traiter que de problèmes d'application de l'IA qui, bien qu'ils puissent être importants et pertinents dans l'élaboration d'une EIA, n'épuisent pas l'ensemble des questions fondamentales qui sont associées au déploiement de l'IA. Il faut donc privilégier une position mitoyenne dans l'évaluation des enjeux éthiques fondamentaux de l'IA qui évite cette dichotomie stérile. Force est de constater que bon nombre de positions récentes en EIA adoptent une perspective qui évite les antipodes. Il existe donc plusieurs exemples de telles positions mitoyennes. Toutefois, la plupart de ces positions mitoyennes dans le spectre argumentatif en EIA sont fondées sur une approche EaU, ce qu'il faut également remettre en question étant donné les apories bien connues des théories éthiques classiques, tant d'un point de vue théorique qu'en éthique de la technologie (voir 2.2). Cela n'est d'ailleurs pas sans rappeler le diagnostic posé sur la théorie éthique par Hans Jonas (1979), à savoir qu'il faut repenser l'éthique face au développement technologique (Bruneault, 2012), étant donné la transformation de l'agir humain et l'amplitude sans commune mesure des actions humaines générées par le développement technologique, ce qui se confirme avec le développement exponentiel du monde numérique.

Il faut également être en mesure de répondre à une autre série de critiques faites aux approches EaU de l'EIA. En effet, plusieurs défendent une perspective critique politique à propos du déploiement social de l'IA qui insiste sur les rapports de pouvoir impliqués par le développement du numérique et la gouvernance

algorithmique – questions qui demeurent trop souvent périphériques dans les approches EaU. Cette perspective critique politique sur l'IA (qui mériterait de faire l'objet d'une section complète dans ce texte, ce qui n'est malheureusement pas possible faute d'espace) constitue une composante essentielle et trop souvent négligée en EIA. Il faut donc penser une approche de l'EIA qui puisse se placer sur le terrain du débat éthique sans se limiter aux approches théoriques normatives classiques, tout en intégrant aussi les questionnements associés à une telle perspective critique politique, pour éviter de reproduire une tendance qu'ont les approches EaU en EIA, celle de privilégier l'élaboration de déclarations ou de guides d'action sous forme de liste de principes généraux, non hiérarchisés, peu problématisés et difficilement transférables tels quels dans les pratiques (Hagendorff, 2020 ; Mittelstadt, 2019). Une position satisfaisante en EIA doit donc rencontrer trois conditions : 1) éviter les positions substantialiste et instrumentaliste en développant une perspective en EIA qui soit *raisonnablement déflationniste*, tout en permettant de 2) situer les enjeux éthiques liés au déploiement de l'IA dans le cadre socio-politique associé à l'ubiquité des TIC dans nos sociétés et 3) repenser les cadres théoriques de l'éthique normative classique pour les mettre à jour et élaborer une explication cohérente des assises théoriques de la position défendue en EIA. Si plusieurs positions récentes en EIA remplissent la première condition, et un bon nombre d'entre elles la seconde, il nous apparaît que peu de travaux cherchent à répondre à la troisième. Nous pensons en ce sens que la position informationnaliste en EIA défendue par Luciano Floridi (2013) est particulièrement intéressante puisqu'elle remplit toutes ces exigences. Voilà pourquoi nous soutenons qu'elle permet de développer un cadre d'analyse qui évite certains écueils conceptuels courants en EIA.

2. Position informationnaliste en EIA

Face aux insuffisances des approches EaU de l'EIA et compte tenu des analyses de la perspective critique politique, la position informationnaliste de l'EIA cherche à développer un cadre éthique qui puisse poser adéquatement les questions philosophiques fondamentales soulevées par le déploiement de l'IA, en évitant de présupposer un développement de l'IA qui ne correspond pas à la réalité technique. Nous verrons maintenant (2.1) quelle est la position éthique sur l'IA défendue par la position informationnaliste,

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

(2.2) nous en examinerons les présupposés philosophiques et (2.3) nous verrons les applications que nous pouvons en espérer.

2.1 Position informationnaliste en EIA

La position informationnaliste sur l'IA, comme les autres perspectives déflationnistes, stipule que les systèmes actuels d'IA sont des systèmes d'IA faible et que les perspectives d'une IAG, comme celle présupposée par la position substantialiste, ne correspondent pas aux conditions technologiques actuelles et à celles qu'on peut entrevoir dans un avenir rapproché (Domingos, 2015 ; Floridi, 1999). Cependant, à la différence de la position instrumentaliste et d'autres perspectives déflationnistes, la position informationnaliste soutient que le déploiement social de l'IA implique beaucoup plus qu'un simple accommodement entre les droits individuels et ces innovations technologiques. Faisant écho à plusieurs analyses actuelles qui insistent sur les modifications que le déploiement social de l'IA implique dans nos rapports au monde, à nous-mêmes et aux autres (Broussard, 2018 ; Cheney-Lippold, 2017 ; Cohen, 2012 ; Dignum, 2019 ; Floridi, 2014 ; Freitag, 2018 ; Hansen, 2006 ; Rouvroy et Stiegler, 2015), la position informationnaliste soutient qu'il est nécessaire de traiter de ces questions philosophiques fondamentales pour espérer développer une approche satisfaisante de l'EIA. Par ailleurs, la position informationnaliste soutient, à la différence de la plupart de ces positions mitoyennes sur le spectre argumentatif en EIA, qu'une telle EIA ne peut pas être simplement une EaU. Il faut donc développer une nouvelle approche éthique qui permet, d'une part, de penser les enjeux éthiques soulevés par l'IA dans toute leur complexité et qui, d'autre part, fait place aux modifications que devront subir les cadres théoriques de l'éthique normative classique, afin de prendre en compte les aspects tout à fait inédits de la situation actuelle d'ubiquité des TIC et de l'IA.

La position informationnaliste s'appuie sur l'éthique de l'information, telle qu'elle est développée par Luciano Floridi (2013). Appuyant son éthique sur sa philosophie de l'information (Floridi, 2011), ce dernier soutient que la situation actuelle, liée au développement des TIC et à la révolution numérique qui y est associée, nous oblige à repenser l'idée même d'information, afin d'en faire le pivot central de la philosophie, en particulier d'une approche éthique satisfaisante de nos sociétés (Floridi, 2013, p. 1). Dans cette perspective, l'IA devra être évaluée d'un point de vue éthique en fonction des capacités et des problèmes qu'elle pourrait

générer en termes d'« épanouissement informationnel », notamment celui des individus, des groupes d'individus et des sociétés qui vivent dans les conditions matérielles et informationnelles découlant du déploiement de l'IA (Floridi, 2013, p. 169).

2.2 Présupposés philosophiques informationalistes

Les présupposés philosophiques derrière la position informationaliste de l'EIA sont clairement exposés, notamment sous forme synthétique par Floridi (2014). Le premier élément qu'il relève à propos des sociétés de l'information dans lesquelles nous vivons, modelées par le déploiement de l'IA, est ce qu'il appelle l'ubiquité des TIC. Le développement fulgurant et exponentiel des TIC nous oblige à repenser entièrement les présupposés philosophiques à partir desquels nous comprenons nos sociétés et nous nous comprenons nous-mêmes (Floridi, 2013, p. 8). Ce développement des TIC met en lumière la centralité et la prépondérance du concept d'information pour les individus et les sociétés. Floridi avance qu'il s'agit d'une quatrième révolution (Floridi, 2013, p. 13), en ce sens qu'après la révolution copernicienne qui a remis en question l'idée que la Terre est au centre de l'univers, après la révolution darwinienne qui a remis en question notre statut particulier dans le monde de la vie et après la révolution freudienne qui a remis en question l'idée cartésienne d'un individu totalement maître de lui-même et en plein contrôle de ses pensées, nous sommes aujourd'hui face à une quatrième révolution de ce type, à savoir la révolution de l'information, qui non seulement montre que ce concept d'information est au cœur même de l'identité humaine, mais qu'aujourd'hui nous ne sommes plus les seules entités capables de produire du sens et de produire des informations, puisque nous avons dorénavant nous-mêmes construit des machines qui sont capables de créer et manipuler cette information, à savoir les TIC en général et l'IA en particulier.

Selon Floridi, nous sommes donc entrés avec la révolution numérique dans ce qu'il appelle l'hyperhistoire (Floridi, 2013, p. 3). L'humanité est passée de la préhistoire à l'histoire au moment où les premières formes des TIC sont apparues, le passage de la préhistoire à l'histoire étant celui d'une société qui reprend continuellement les mêmes cycles de vie à une société qui est capable de conserver, notamment via l'écriture, certaines formes de mémoire collective qui permettent le développement des cultures complexes et élaborées que nous connaissons. Aujourd'hui, les TIC ont pris une nouvelle importance dans le développement de nos

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

sociétés. Dans l'histoire, les TIC ont servi à compiler et consigner les événements qui, par ailleurs, suivaient leur propre cours, qu'ils soient de nature sociale, économique ou politique. Aujourd'hui, les TIC sont devenues le moteur même des systèmes sociaux, économiques et politiques, de telle sorte qu'elles occupent le rôle central dans le développement et la reproduction des sociétés de l'information, du moins, en ce début de XXI^e siècle, dans les sociétés post-industrielles des pays du G7.

Au même titre que les êtres humains font partie du monde vivant, la biosphère, Floridi avance qu'ils évoluent également dans un univers informationnel qu'on peut appeler l'infosphère (Floridi, 2013, p. 8), univers qui était sous-entendu dans les phases précédentes du développement des sociétés humaines et qui en devient dorénavant une composante essentielle, car « ce n'est que très récemment que le progrès et le bien-être humain ont commencé à dépendre principalement du succès et de l'efficacité de la gestion du cycle de vie de l'information » (Floridi, 2013, p. 3, notre traduction⁴). Les êtres humains évoluent dans cet univers de signification et de sens, ce qui devient tout à fait apparent une fois que les TIC ont permis de libérer les capacités de gestion informationnelle au-delà des limitations qui existaient auparavant, une fois que nous entrons dans l'hyperhistoire. Il convient donc de réfléchir aux enjeux éthiques suscités par le déploiement de l'IA en situant ces questions dans cet univers de sens.

Floridi propose également un néologisme, l'expression « *onlife* », pour caractériser nos rapports actuels au monde numérique (Floridi, 2015). La différence qui était faite, il y a encore quelques années, entre la vie « *offline* » et la vie « *online* » devient de plus en plus floue, voire totalement effacée, si bien que notre vie est aujourd'hui constamment en lien avec le monde numérique en raison des TIC et de leur ubiquité (sur les questions liées à l'« intelligence ambiante », voir Costa, 2016). Il est indéniable que cet effacement a été accentué par la pandémie de COVID-19 et par les mesures de confinement qui ont imposé à plusieurs une vie personnelle, professionnelle et sociale essentiellement en ligne. Il faut donc réfléchir aux questions philosophiques et aux questions éthiques impliquées par le déploiement de l'IA, en prenant acte de cette vie « *onlife* » et de l'effacement que cette nouvelle dimension de notre existence suppose pour de nombreuses sphères de nos vies. Floridi (2015, p. 7) en identifie quatre : 1) l'effacement de la distinction entre monde réel et monde virtuel, 2) l'effacement des distinctions entre nature, personnes et créations artificielles, 3) le passage d'une situation de rareté informationnelle à ce que ce qu'il est dorénavant

⁴ «only very recently has human progress and welfare begun to depend mostly on the successful and efficient management of the life-cycle of information»

convenu d'appeler l'« infobésité », la surabondance d'information et 4) le passage d'une métaphysique des entités à une métaphysique des processus, puisque nous avons de plus en plus recours non pas aux entités matérielles et physiques pour décrire notre monde, mais bien aux relations entre ces entités. L'EIA doit donc nécessairement s'appuyer sur une théorie éthique qui permet de penser ces changements. L'éthique de l'information permet donc d'adopter une position en EIA qui est généralement déflationniste (comparativement aux transhumanistes, par exemple), tout en offrant un cadre conceptuel permettant de rendre compte de l'aspect disruptif de ces nouvelles technologies, ce qui remplit la première condition d'une position satisfaisante en EIA.

Cette éthique de l'information est par ailleurs « une macro-éthique non standard (parce qu'orientée vers le patient), ontocentrique et e-nvironnementale » (Floridi, 2013, p. 97, notre traduction⁵). D'abord, l'éthique de l'information est une éthique non standard (Floridi, 2013, p. 62) orientée vers le patient (*patient-oriented*). À la différence des cadres théoriques de l'éthique normative classique qui accordent une valeur morale en fonction 1) d'une évaluation de l'agent moral (comme dans l'éthique des vertus) ou 2) par une évaluation morale de l'action posée (comme dans les éthiques conséquentialiste et déontologique), l'éthique de l'information cherche plutôt à évaluer les actions ou les situations à la lumière de l'impact que celles-ci ont sur les entités qui subissent ces impacts, pour évaluer moralement les possibilités offertes à ces entités en termes d'« épanouissement informationnel » (Floridi, 2013, p. 74).

L'éthique de l'information est par ailleurs ontocentrique (Floridi, 2013, p. 65) ; elle accorde une valeur aux entités informationnelles, quelles qu'elles soient, en fonction du degré d'organisation informationnelle qu'elles incarnent. Or, puisque toutes les entités existantes sont des entités informationnelles, aussi simples soient-elles, il s'ensuit qu'elles devraient toutes être considérées, au moins minimalement, dans l'évaluation éthique des actions ou des situations. Contrairement aux éthiques anthropocentriques qui n'accordent de valeur qu'aux seuls êtres humains et à la différence des éthiques biocentriques qui n'accordent de valeur qu'aux êtres vivants, Floridi propose d'accorder une valeur non absolue à toute forme d'entité informationnelle et donc à toute entité existante, à tous les êtres.

L'éthique de l'information doit finalement être un « e-nvironnementalisme » (Floridi, 2013, p. 18 ; Floridi, 2014, p. 217) qui doit se distinguer de l'environnementalisme éthique tel qu'il s'est

⁵ « a non-standard (because patient-oriented), ontocentric and e-nvironmental macroethics »

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

développé au XX^e siècle, puisque l'« e-nvironmentalisme » de l'éthique de l'information ne doit pas opposer le monde artificiel créé par l'être humain à une nature intouchée et inviolable qui composerait l'environnement du monde humain. Au contraire, l'« e-nvironmentalisme » prend acte du fait que l'environnement, à partir duquel les êtres humains agissent et les sociétés humaines se développent, est de plus en plus constitué des objets artificiels créés par l'être humain. Il faut donc penser la dimension éthique de l'action humaine non seulement dans l'environnement naturel (au sens classique), mais également en tenant compte du fait que cet environnement est en bonne partie lui-même le résultat de l'action humaine et des décisions des générations précédentes. Ce faisant, les actions et les décisions actuelles doivent également être évaluées en fonction du rôle qu'elles joueront dans la construction de l'environnement à partir duquel les générations futures pourront et devront elles-mêmes agir. Cet aspect de l'éthique de l'information établit ainsi un pont aussi bien entre les individus d'une même époque, en tenant compte de la dimension sociale de l'existence humaine, qu'entre les générations. Il fait également référence à un élément central de l'approche de Floridi, à savoir sa méthode d'abstraction (*Method of Abstraction – MoA*) qui distingue différents niveaux d'abstraction (*Levels of Abstraction – LoA*) pour analyser une entité quelconque (Floridi, 2013, p. 29). Il devient possible, suivant cette MoA, d'évaluer l'impact éthique du développement de l'IA, non pas seulement sur les individus, mais bien à différents LoA – individus, relations interpersonnelles, groupes, sociétés, générations, etc. Cette dimension sociale, qui fait écho à la seconde condition d'une position satisfaisante en EIA, sera d'ailleurs développée plus amplement dans l'application politique de l'éthique de l'information (voir 2.3).

Reprenant le concept d'entropie, emprunté d'abord à la thermodynamique et ensuite à la cybernétique (Wiener, 1954), Floridi propose d'interpréter son éthique de l'information en termes d'« épanouissement informationnel » et d'« entropie informationnelle », c'est-à-dire ce qui favorise le développement informationnel et ce qui entraîne plutôt la désorganisation informationnelle dans l'infosphère. Il propose en ce sens quatre principes fondamentaux pour l'éthique de l'information (Floridi, 2013, p. 70), inspirés des célèbres lois de la robotique d'Isaac Asimov, à savoir :

- 0) qu'il ne faut pas causer l'entropie informationnelle ;
- 1) qu'il faut prévenir l'entropie informationnelle ;
- 2) qu'il faut éliminer l'entropie informationnelle ;
- 3) qu'il faut favoriser l'épanouissement informationnel.

Ces principes de l'éthique de l'information devraient donc s'appliquer aux actions et aux situations pour toutes les entités informationnelles, quelles qu'elles soient, de façon non absolue, tout en réservant la possibilité d'accorder un statut particulier, éventuellement absolu, aux entités informationnelles et sémantiques, créatrices de sens et capables de comprendre le sens de l'information, que sont les êtres humains. On peut donc envisager évaluer l'IA en fonction de sa capacité (pour chaque système d'IA séparément) à favoriser l'« épanouissement informationnel » ou à créer de l'« entropie informationnelle » dans l'infosphère, en considérant l'impact de ces systèmes d'IA sur toutes les entités informationnelles, notamment sur les êtres humains en tant qu'entités informationnelles et sémantiques. Ces particularités de l'éthique de l'information permettent de repenser les cadres théoriques de l'éthique normative classique, compte tenu de l'ubiquité des TIC et de l'IA dans nos sociétés, ce qui remplit la troisième condition d'une position satisfaisante en EIA.

2.3 Position informationnaliste : applications

La position informationnaliste de l'EIA peut être en discussion avec l'EAIA, puisqu'elle considère qu'il s'agit d'une composante incontournable de l'EIA. Elle ne vise pas non plus à supplanter les cadres théoriques en éthique normative classique, mais bien à entrer en discussion avec ces approches pour chercher à les repenser (Floridi, 2013, p. 77). Ainsi, l'apport de la position informationnaliste aux discussions d'éthique appliquée est considérable (par ex., voir le texte de Floridi sur les agents artificiels autonomes dans Anderson et Anderson, 2011, p. 184). À titre d'illustration, il est possible de souligner la particularité de l'interprétation des enjeux éthiques liés à la protection de la vie privée que propose l'éthique de l'information (Floridi, 2013, p. 228). Contrairement aux interprétations classiques associant la captation non souhaitée de données personnelles à un vol, l'éthique de l'information avance que, si les êtres humains sont des entités informationnelles et sémantiques, il s'ensuit que leurs données personnelles ne sont pas une propriété que les individus possèdent, au même titre que des biens matériels. Ces données personnelles sont plutôt une description même de ce que sont ces individus fondamentalement (Floridi, 2013, p. 242). La protection de la vie privée est donc une composante essentielle de l'évaluation des actions et des situations pour l'éthique de l'information, puisqu'une fois que nous comprenons la prépondérance du concept d'information dans l'identité personnelle, nous réalisons que les données personnelles sont une partie intrinsèque de ce que sont les

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

individus eux-mêmes. L'utilisation des données personnelles au détriment des individus s'apparente ainsi plus à une forme de kidnapping qu'à une forme de vol.

Soulignons également que l'interprétation politique de la position informationnaliste de l'EIA s'articule autour de quatre idées : une nouvelle définition du pouvoir, la déterritorialisation, la séparation de l'identité et de la cohérence ainsi que la remise en question continue du consentement (Floridi, 2015, p. 51). De telles considérations politiques, qu'il est impossible de développer ici davantage, font écho à l'approche des capacités technologiques (Brey, 2010 ; Costa, 2016 ; Oosterlaken et van den Hoven, 2012 ; Stahl, 2007 ; Zheng et Stahl, 2011), mais aussi à la perspective critique politique, notamment par les analyses critiques des nouveaux pouvoirs algorithmiques (Ananny et Crawford, 2016 ; Andrejevic, 2013 ; Bucher, 2018 ; De Grosbois, 2018 ; Dean, 2009 ; Freitag, 2003 ; Morozov, 2013 ; Nissenbaum, 2010 ; Rouvroy et Berns, 2013 ; Susskind, 2018 ; Taylor 2014 ; Whittaker, 2019), les approches en théorie critique de la technologie (Feenberg, 1991, 2010) et en théorie critique des TIC (George et Kane, 2015 ; Mondoux, 2011 ; Mondoux et Ménard, 2018 ; Ouellet, 2016 ; Ouellet *et al.*, 2015). La position informationnaliste sur les enjeux politiques de l'IA est complémentaire à la perspective critique politique. Elle ne vise pas à supplanter ou à remplacer celle-ci ; au contraire, elle cherche plutôt à offrir une interprétation satisfaisante des enjeux de l'EIA qui puisse être au diapason des analyses critiques des rapports de force, de la réification et des enjeux sociaux liés au déploiement de l'IA, ce qui remplit la seconde condition d'une position satisfaisante en EIA.

Il nous apparaît donc que la position informationnaliste est une perspective fort à propos pour développer une position en EIA qui ne soit ni catastrophiste, ni tributaire d'un solutionnisme technologique, ni démesurément déflationniste, tout en tenant compte des questions politiques qui émergent du déploiement social de l'IA et des limites des cadres théoriques en éthique normative classique. Une telle approche permet de couvrir l'ensemble des aspects de l'EIA, tant en éthique professionnelle qu'en éthique appliquée, en éthique normative et en méta-éthique. Voilà pourquoi il s'agit du point d'ancrage nécessaire au développement d'une EIA complète et appropriée au contexte social d'ubiquité des TIC. L'EIA en est encore à un stade précoce de son développement ; espérons que ce texte puisse être une contribution conséquente à son essor.

Références

Alter, A. (2018). *Irresistible, The Rise of Addictive Technology and the Business of Keeping Us Hooked*. New York : Penguin Books.

Ananny, M. et Crawford, K. (2016). Seeing Without Knowing: Limitations of the Transparency Ideal and its Application to Algorithmic Accountability. *New Media et Society*. 20(3), 973-989.

Anderson, M. et Anderson, S. L. (dir.). (2011). *Machine Ethics*. Cambridge : Cambridge University Press.

Andrejevic, M. (2013). *Infoglut, How Too Much Information Is Changing the Way We Think and Know*. New York : Routledge.

Awad, E. et al. (2018). The Moral Machine Experiment. *Nature*, 563, 59-77.

Bengio, Y. et Luccioni, S. (2019). On the Morality of Artificial Intelligence. Reperé à [arXiv:1912.11945](https://arxiv.org/abs/1912.11945)

Boddington, P. (2017). *Towards a Code of Ethics for Artificial Intelligence*. New York : Springer.

Bostrom, N. (2014). *Superintelligence, Paths, Dangers, Strategies*. Oxford : Oxford University Press. (Nous citons la traduction française : Bostrom, N. (2017). *Superintelligence*. Malakoff : Dunod.)

Boutin, E. (2006). Biais cognitifs et recherche d'information sur internet. Quelles perspectives pour les indicateurs de pertinence des moteurs de recherche. *VSST 2006*. https://archivesic.ccsd.cnrs.fr/sic_00827309/document

Brey, P. (2010). Values in Technology and Disclosive Computer Ethics. Dans L. Floridi (éd.), *The Cambridge Handbook of Information and Computer Ethics* (p. 41-58). Cambridge : Cambridge University Press.

Broussard, M. (2018). *Artificial Unintelligence, How Computers Misunderstand the World*. Cambridge, MA : MIT Press.

Bruneault, F. (2012). Comment définir une éthique pour notre civilisation technologique? L'apport d'une lecture conjointe des pensées de Karl-Otto Apel et Hans Jonas. *Laval théologique et philosophique*, 68(2), 335-357.

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

Bryson, J. J. (2019). The Past Decade and Future of AI's Impact on Society. Dans *Towards a New Enlightenment?* BBVA. bbvaopenmind.com/wp-content/uploads/2019/02/BBVA-OpenMind-book-2019-Towards-a-New-Enlightenment-A-Trascendent-Decade-3.pdf

Bucher, T. (2018). *If... Then, Algorithmic Power and Politics*. Oxford : Oxford University Press.

Cheney-Lippold, J. (2017). *We Are Data, Algorithms and the Making of Our Digital Selves*. New York : New York University Press.

Cohen, J. E. (2012). *Configuring the Networked Self, Law, Code, and the Play of Everyday Practice*. New Haven et London : Yale University Press.

Costa, L. (2016). *Virtuality and Capabilities in a World of Ambient Intelligence, New Challenges to Privacy and Data Protection*. New York : Springer.

CRYPTO.QUÉBEC. (2018). *On vous voit. Comment déjouer les malveillants sur Internet*. Montréal : Trécaré.

De Filippi, P. (2016). Gouvernance algorithmique : Vie privée et autonomie individuelle à l'ère des Big Data. Dans P. De Filippi et D. Bourcier (dir.), *Open Data et Data Protection : Nouveaux défis pour la vie privée*. Paris : Mare et Martin, 1-22.

De Grosbois, P. (2018). *Les Batailles d'Internet, Assauts et résistances à l'ère du capitalisme numérique*. Montréal : Écosociété.

Dean, J. (2009). *Democracy and Other Neoliberal Fantasies, Communicative Capitalism and Left Politics*. Durham, NC : Duke University Press.

Déclaration de Montréal pour un développement responsable de l'intelligence artificielle. (2018). Université de Montréal.

Dignum, V. (2019). *Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way*. Springer.

Domingos, P. (2015). *The Master Algorithm, How the Quest for the Ultimate Learning Machine Will Remake Our World*. New York : Basic Books.

Feenberg, A. (1991). *Critical Theory of Technology*. Oxford : Oxford University Press.

Feenberg, A. (2010). *Between Reason and Experience. Essays in Technology and Modernity*. Cambridge, MA : MIT Press.

Floridi, L. (1999). *Philosophy and Computing, An Introduction*. London : Routledge.

Floridi, L. (2011). *The Philosophy of Information*. Oxford : Oxford University Press.

Floridi, L. (2013). *The Ethics of Information*. Oxford : Oxford University Press.

Floridi, L. (2014). *The 4th Revolution, How the Infosphere Is Reshaping Human Reality*. Oxford : Oxford University Press.

Floridi, L. (dir.). (2015). *The Onlife Manifesto, Being Human in a Hyperconnected Era*. New York : Springer.

Freitag, M. (2003). La dissolution systémique du monde réel dans l'univers virtuel des nouvelles technologies de la communication informatique : une critique ontologique et anthropologique. Dans A. Mattelart et G. Tremblay (dir.), *2001 Bogues : Communication, démocratie et globalisation* (tome 4, p. 279-296). Québec : Presses de l'Université Laval.

Freitag, M. (2018). La société informatique et le respect des formes. *Le Naufrage de l'université et autres essais d'épistémologie politique*. Montréal : Alias.

George, É. et Kane, O. (2015). Les technologies numériques au prisme des approches critiques : éléments pour l'ébauche d'une rencontre. *Canadian Journal of Communication*, 40, 727-735.

Gibert, M. (2019). Faut-il avoir peur de la peur de l'IA? *La Quatrième Blessure : un blogue sur l'éthique de l'intelligence artificielle*. Repéré à : <https://medium.com/@martin.gibert/faut-il-avoir-peur-de-la-peur-de-lia-1687abc35342>

Jobin, A., Ienca, M. et Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature*, 389-399.

Jonas, H. (1979). *Das Prinzip Verantwortung, Versuch einer Ethik für die technologische Zivilisation*, Francfort sur le Main: Suhrkamp Verlag.

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, 30, 99-120.

Hansen, M. B. N. (2006). *Bodies in Code, Interfaces with Digital Media*. New York : Routledge.

Hunyadi, M. (2018). *Le Temps du posthumanisme, Un Diagnostic d'époque*. Paris : Les Belles Lettres.

Kurzweil, R. (2007). *Humanité 2.0. La bible du changement*. Paris : M21 Éditions.

Levesque, H. J. (2017). *Common Sense, the Turing Test, and the Quest for Real AI*. Cambridge, MA : MIT Press.

Maclure, J. (2019). The New AI Spring: A Deflationary View. *AI et Society. A Journal of Knowledge, Culture and Communication*, 35(3), 747-750.
<https://doi.org/10.1007/s00146-019-00912-z>

Maclure, J. et Saint-Pierre, M.-N. (2018). Le nouvel âge de l'intelligence artificielle : une synthèse des enjeux éthiques. *Les cahiers de propriété intellectuelle*, 30(3), 741-765.

Mayer-Schönberger, V. (2014). La Révolution Big Data, *Politique étrangère*, 4, 69-81.

Mittelstadt, B. (2019). Principles Alone Cannot Guarantee Ethical AI. *Nature*, 1, 501-507.

Mondoux, A. (2012). À propos du social dans les médias sociaux. *Terminal*, 111, 69-79.

Mondoux, A. (2011). Identité numérique et surveillance. *Les Cahiers du numérique*, 7(1), 49-59.

Mondoux, A. et Ménard, M. (2018). *Big Data et société, Industrialisation des médiations symboliques*. Québec : Presses de l'Université du Québec.

More, M. (1990). Transhumanism: Towards a Futurist Philosophy. *Extropy*, 6, 6-12. Repéré à <http://fenetic.net/irc/extropy/ext6.pdf>

Morozov, E. (2013). *To Save Everything, Click Here*. Philadelphia, PA : PublicAffairs Books.

Nissenbaum, H. (2010). *Privacy in Context: Technology, Policy and the Integrity of Social Life*. Stanford, CA : Stanford University Press.

Oosterlaken, I. et van den Hoven, J. (éd.). (2012). *The Capability Approach, Technology and Design*. New York : Springer.

Ouellet, M. (2016). *La Révolution culturelle du capital, le capitalisme cybernétique dans la société globale de l'information*. Montréal : Écosociété.

Ouellet, M., Ménard, M., Bonenfant, M. et Mondoux, A. (2015). Big Data et quantification de soi : la gouvernamentalité algorithmique dans le monde numériquement administré. *Canadian Journal of Communication*, 40, 597-613.

Rouvroy, A. (2018). Homo *juridicus* est-il soluble dans les données? Dans C. De Terwangne, É. Degrave et S. Dusollier (dir.), *Law, Norms and Freedoms in Cyberspace, Droit, normes et libertés dans le cybermonde*. Bruxelles : Larcier, 417-444.

Rouvroy, A. et Stiegler, B. (2015). Le régime de vérité numérique. *Socio*. 4. [En ligne] <http://journals.openedition.org/socio/1251> ; DOI : 10.4000/socio.1251

Rouvroy, A. et Berns, T. (2013). Gouvernamentalité algorithmique et perspectives d'émancipation. *Réseaux*. 177.

Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. New York, NY : Viking.

Russell, S., Dewey, D. et Tegmark, M. (2015). Research Priorities for Robust and Beneficial Artificial Intelligence. *AI Magazine*, 36(4), 105-114.

Stahl, B. C. (2007). Ontology, Life-World, and Responsibility in IS. Dans R. Sharman, R. Kishore et R. Ramesh (éd.), *Ontologies. A Handbook of Principles, Concepts and Applications in Information Systems*. New York, NY : Springer, 143-169.

Susskind, J. (2018). *Future Politics, Living Together in a World Transformed by Tech*. Oxford : Oxford University Press.

Taylor, A. (2014). *Démocratie.com, pouvoir, culture et résistance à l'ère des géants de la Silicon Valley*. Montréal : Lux.

Éthique de l'intelligence artificielle et ubiquité sociale des technologies de l'information et de la communication : comment penser les enjeux éthiques de l'IA dans nos sociétés de l'information ?

Tegmark, M. (2017). *Life 3.0, Being Human in the Age of Artificial Intelligence*. New York, NY : Vintage Books.

Turkle, S. (2005). *The second Self: Computers and the Human Spirit*. Cambridge : MIT Press.

van den Hoven, J. (2010). The Use of Normative Theories in Computer Ethics. *The Cambridge Handbook of Information and Computer Ethics*. Cambridge : Cambridge University Press.

Whittaker, M. (2019). Artificial Intelligence: Societal and Ethical Implications. *United States House of Representatives Committee on Science, Space, and Technology*.

<https://science.house.gov/imo/media/doc/Whittaker%20Testimony.pdf>

Wiener, N. (1954). *The Human Use of Human Beings, Cybernetics and Society*. New York : Doubleday.

Zheng, Y. et Stahl, B. C. (2011). Technology, Capabilities and Critical Perspectives: What Can Critical Theory Contribute to Sen's Capability Approach? *Ethics and Information Technology*, 13(2), 69-80.